National Park Service
U.S. Department of the Interior

Natural Resource Stewardship and Science

Sierra Nevada Network Climate Reporting Protocol

Version 1.0

Natural Resource Report NPS/SIEN/NRR—2012/543

ON THE COVER
Air monitoring technician Alysia Schmidt skiing to the Upper Kaweah air quality monitoring station in SEQU
Photograph by: Jennie Skancke - SIEN

Sierra Nevada Network Climate Reporting Protocol

Version 1.0

Natural Resource Report NPS/SIEN/NRR—2012/543

Jennie R. Skancke[1]
Alice L. Chung-MacCoubrey[1]
Leslie S. Chow[1]
Jeffrey D. Balmat[2]

[1]National Park Service, Sierra Nevada Network
Sequoia and Kings Canyon National Parks
47050 Generals Highway
Three Rivers, CA 93271

[2]US Geological Survey
2255 N. Gemini Dr.
Flagstaff, AZ 86001

June 2012

National Park Service
Natural Resource Stewardship and
Science
Fort Collins, Colorado

The National Park Service, Natural Resource Stewardship and Science office in Fort Collins, Colorado publishes a range of reports that address natural resource topics of interest and applicability to a broad audience in the National Park Service and others in natural resource management, including scientists, conservation and environmental constituencies, and the public.

The Natural Resource Report Series is used to disseminate high-priority, current natural resource management information with managerial application. The series targets a general, diverse audience, and may contain NPS policy considerations or address sensitive issues of management applicability.

All manuscripts in the series receive the appropriate level of peer review to ensure that the information is scientifically credible, technically accurate, appropriately written for the intended audience, and designed and published in a professional manner. This report received formal peer review by subject-matter experts who were not directly involved in the collection, analysis, or reporting of the data, and whose background and expertise put them on par technically and scientifically with the authors of the information.

This report is available from the Sierra Nevada Network (http://science.nature.nps.gov/im/units/sien/) and the Natural Resource Publications Management Website (http://www.nature.nps.gov/publications/rprpm/).

Please cite this publication as:

Skancke, J. R., A. L. Chung-MacCoubrey, L. S. Chow, and J. D. Balmat. 2012. Sierra Nevada Network climate reporting protocol: Version 1.0. Natural Resource Report NPS/SIEN/NRR—2012/543. National Park Service, Fort Collins, Colorado.

NPS 963/115248, June 2012

Contents

Contents (continued)

Contents (continued)

Figures

Tables

Page

Appendices

Page

Standard Operating Procedures (SOPs)

This protocol relies upon SOPs and other documents developed, maintained, and updated by the National Inventory & Monitoring Division (IMD). A description of the IMD Climate Database Project and links to the latest versions of SOPs may be found at https://irma.nps.gov/App/Reference/GroupedProducts?parentReferenceCode=2167254. Below we list IMD SOPs used by this protocol and their citations. The first title is the abbreviated name used within our protocol. The second title is the full SOP name.

IMD SOP: Requesting Essential Parameter Values. Refers to Requesting Essential Parameter Values from COOP, SNOTEL, Stream Gage, and Snow Course Data for Acquisition, Compilation, Basic Quality Control, and Distribution by the NPS Natural Resource Program Center (Frakes 2012a).

IMD SOP: Submitting Data for Upload. Refers to Submitting Data for Upload to the I&M Enterprise Climate Database (Frakes 2012b).

IMD SOP: Connecting to the I&M Climate Database. Refers to Connecting to the I&M Enterprise Climate Database for the Purpose of Data Retrieval, Summary and/or Analysis (Frakes and Kingston 2011a).

IMD SOP: Climate Database Schema. Refers to Climate Database Schema, Table and Field Definitions (Frakes and Kingston 2011b).

IMD SOP: Climate Grid Analysis Tools (CGAT). Refers to Climate Grid Analysis Toolset - Tools for Assessing Regional Climatological Trends (Sherrill and Frakes 2011).

Acronyms and Abbreviations

CDEC	California Data Exchange Center
CDWR	California Department of Water Resources
COOP	Cooperative Observer Network
ENSO	El Niño-Southern Oscillation
ESRI	Environmental Systems Research Institute, Inc.
GIS	Geographic Information System
I&M	Inventory & Monitoring
IMD	National Inventory & Monitoring Division
NCDC	National Climatic Data Center
NOAA	National Oceanic and Atmospheric Administration
NPS	National Park Service
NRCS	Natural Resources Conservation Service
NRSS	Natural Resource Stewardship and Science Directorate
NWS	National Weather Service
PDO	Pacific Decadal Oscillation
PDSI	Palmer Drought Severity Index
PRISM	Parameter-Elevation Regressions on Independent Slopes Model
Protocol	Sierra Nevada Network Weather and Climate Monitoring Protocol
RAWS	Remote Automated Weather Station
Scripps	Scripps Institution of Oceanography
SIEN	Sierra Nevada Inventory & Monitoring Network
SOP	Standard Operating Procedure
SWE	Snow water equivalent
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey

Executive Summary

Weather and climate are primary drivers of the biological, physical, and ecological processes that determine the distribution, structure, and function of ecosystems. Climate also influences park operations and visitor experience, thus creating a need by park management for information on current and future climate and climate trends. Evidence suggests that climate has changed in the past century and will continue to change. For these reasons, the Sierra Nevada Network Inventory and Monitoring Program (SIEN) identified climate as a high priority vital sign. Here, we present the SIEN approach to weather and climate reporting within the four national park units in the Sierra Nevada Network: Devils Postpile National Monument, Sequoia and Kings Canyon National Parks, and Yosemite National Park. The goals of this climate protocol are (1) to characterize and report variations and changes in key climate variables and (2) to compile and present climate data and products to support park management and facilitate analyses of other vital signs. Specifically, our objectives are to:

1. Report the annual status and long-term trends of the following:
    a. **Temperature**, long-term averages and extremes, at the scale of points (individual monitoring locations), parks, and the region.
    b. **Precipitation**, long-term averages and extremes, at the scale of points, parks, and the region.
    c. **Snow water equivalent** and **snow depth** at the scale of points and major watersheds.
    d. **Drought** at the regional scale.
    e. **Streamflow** at the scale of points (individual streamgages).

2. Archive data from weather stations and snow courses to facilitate access and covariate analyses.

Our approach is to rely upon existing climate monitoring programs with stations that provide consistent, long-term, and high-quality climate records within or near SIEN parks, rather than to establish new monitoring stations. Specifically, we will use data from existing weather stations, snow courses, and streamgages currently operated within SIEN parks. Data from many of these stations are quality checked and managed by the administering program, such as the National Weather Service. Status reports will be prepared annually, while trend reports will be produced on a 5 year cycle and will present rigorous analyses of inter-annual variability, long-term historical trends, and teleconnections with hemispheric climate mechanisms. In the process of preparing these reports, we will create high-quality historical climate datasets that may support resource management, park interpretation programs, and research linking resource dynamics to climate.

This protocol provides a basic summary of climate in the Sierra Nevada, describes climate measures selected for monitoring, articulates our monitoring objectives and goals, summarizes our approach to station selection and data acquisition, analysis, management, and reporting, and outlines personnel requirements and operational costs. This protocol also includes standard operating procedures (SOPs) for implementing the protocol and appendices which describe existing climate monitoring programs and provide comprehensive information on weather stations within SIEN. This protocol and SOPs are patterned after the Rocky Mountain Climate Protocol Narrative and its associated SOPs, which were published in 2010.

Acknowledgments

We thank the professionals who came together for the development of the Rocky Mountain/Yellowstone Climate Protocol. The SIEN protocol relies heavily on, and in many cases duplicates, the draft protocols of the Northern Coast and Cascades I&M Network, Northern Colorado Plateau Network, and the Rocky Mountain & Greater Yellowstone I&M Networks which provided structure and technical expertise developed over years of protocol development and scoping meetings. By adopting a protocol which duplicates or relies on the approach and SOPs used in other protocols, SIEN is able to take advantage of the resources and tools developed by professionals with extensive climate data expertise. Furthermore, coordination between networks is easier when protocols are similar, which allows for greater efficiency within our network because we are able to learn from the experience of those professionals who have been implementing protocols over the last several years.

The authors would also like to acknowledge the following people:

The climate working group, including Annie Esperanza, Jim Roche, Lee Tarnay, Andi Heard, David Scott, and Linda Mutch, provided guidance and feedback throughout the development of the protocol.

Kelly Redmond and Laura Edwards (Western Regional Climate Center) provided significant background material and advice.

The staff at the Inventory and Monitoring Division in Ft. Collins, particularly Brent Frakes, Kirk Sherrill, and Simon Kingston provided extensive assistance with data acquisition, management, and analysis and overall support with the understanding of climate data. We thank the IMD staff for recognizing the networks' need for climate tools including a climate database and for all their hard work developing these resources for our use. We believe their work will result in a much more solid and consistent approach to handling and reporting data within the Inventory and Monitoring Networks.

Isabel Ashton at the Rocky Mountain Network and Rob Daly at the Yellowstone Network offered valuable assistance regarding the climate reporting approach of the Rocky Mountain and Yellowstone Networks.

Revision History Log

Previous Version #	Revision date	Author	Changes made	Section and paragraph	Reason for change	New Version #

1.0 Background and Objectives

1.1 The Sierra Nevada Network Inventory & Monitoring Program

The National Park Service (NPS) Inventory & Monitoring (I&M) Program was established in 2000 as part of the Natural Resource Challenge, a long-term strategy to improve park management by increasing access to and reliance on high-quality scientific information. The Sierra Nevada Network (SIEN) is one of 32 I&M networks that will develop and provide scientifically credible information on the status and long-term trends in selected Vital Signs, or indicators of ecosystem condition.

The five goals of the I&M Program (Fancy et al. 2008) are to:

1. Inventory the natural resources and park ecosystems under NPS stewardship to determine their nature and status
2. Monitor park ecosystems to better understand their dynamic nature and condition and to provide reference points for comparisons with other, altered environments
3. Establish natural resource inventory and monitoring as a standard practice throughout the NPS system that transcends traditional program, activity, and funding boundaries
4. Integrate natural resource inventory and monitoring information into NPS planning, management, and decision making
5. Share NPS accomplishments and information with other natural resource organizations and form partnerships for attaining common goals and objectives

To establish the I&M Program, the NPS created networks of parks that are linked by geography and shared natural resource characteristics. The network approach improves efficiency because parks are able to share budgets, staffing, and other resources to plan and implement an integrated program. Each network will accomplish I&M goals by conducting park-wide inventories and establishing a long-term Vital Sign monitoring program. Vital Signs were selected by each network and are physical, chemical, and biological elements and processes of park ecosystems that have been selected "to represent the overall health or condition of park resources, known or hypothesized effects of stressors, or elements that have important human values" (NPS 2008b).

The Sierra Nevada Network is comprised of four units: Devils Postpile National Monument (DEPO), Sequoia & Kings Canyon National Park (jointly administered units that are referred to as SEKI, or individually as SEQU & KICA), and Yosemite National Park (YOSE). Collectively, these parks comprise over 650,000 hectares, and all are located (in their entirety) within California (Table 1). SIEN parks occur within the Sierra Nevada mountain range, which runs roughly along a NNW–SSE axis (Figure 1). For its selected set of Vital Signs, SIEN will collect, organize, and make available related natural resource data and results of analyses, syntheses, and modeling to better inform park managers and contribute to overall NPS institutional knowledge. The SIEN Vital Signs Monitoring Plan (Mutch et al. 2008) describes the rationale, basis, and overall foundation for this network's long-term ecological monitoring program. This plan resulted from a three-year planning effort that included park staff, network staff, and numerous scientific partners from other organizations.

Table 1. Size and visitation statistics for parks within Sierra Nevada Network.

	DEPO	SEKI	YOSE
Size (hectares)	324	349,581	308,075
Percent Wilderness (%)	75	>96	94
Elevation Range (m)	2200–2500	418–4417	610–3998
Number of Visitors (2009)	110,212	965,170 (SEQU) 609,296 (KICA)	3,737,472

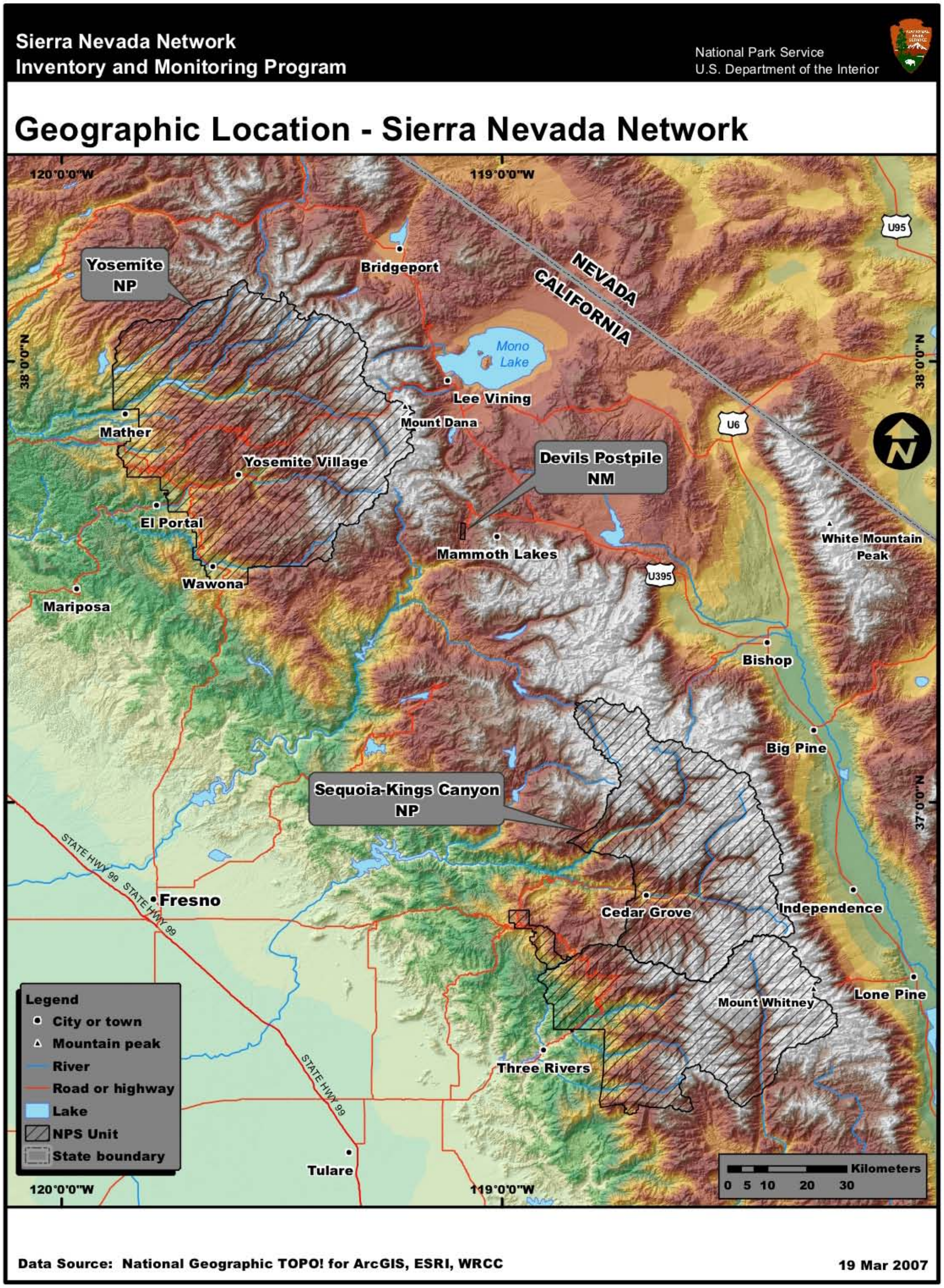


Figure 1. Map of park units within the Sierra Nevada I&M Network, all of which are located in the Sierra Nevada mountain range of California. The drainage divide (along which water drains either east or west) and crest (the highest portion) run along the eastern edge of the range.

1.2 Climate of the Sierra Nevada

Weather and climate describe the condition and variability of the atmosphere in a given place. Weather describes short-term atmospheric conditions—what one experiences at a moment or over a few days. Weather data are instantaneous or short-term observations of current conditions (e.g., the current temperature or cloudiness, today's maximum temperature, yesterday's total precipitation). Climate refers generally to long-term statistical characterizations of atmospheric conditions. The term climate is used in this document to indicate the long-term monitoring of both discrete weather events and climatic conditions.

The goal of this protocol is to report on the status and trends of specific climate parameters (e.g., temperature, precipitation, etc.) at scales relevant to park staff. Although status and trends will be reported individually by parameter, these parameters are inextricably linked because they are all driven by larger-scale atmospheric processes (Figure 2). SIEN developed models to conceptualize the interrelationships among major focal systems in its Monitoring Plan (Appendix F: Ecosystem Conceptual Models in Mutch et al. 2008) and the atmospheric component is summarized below.

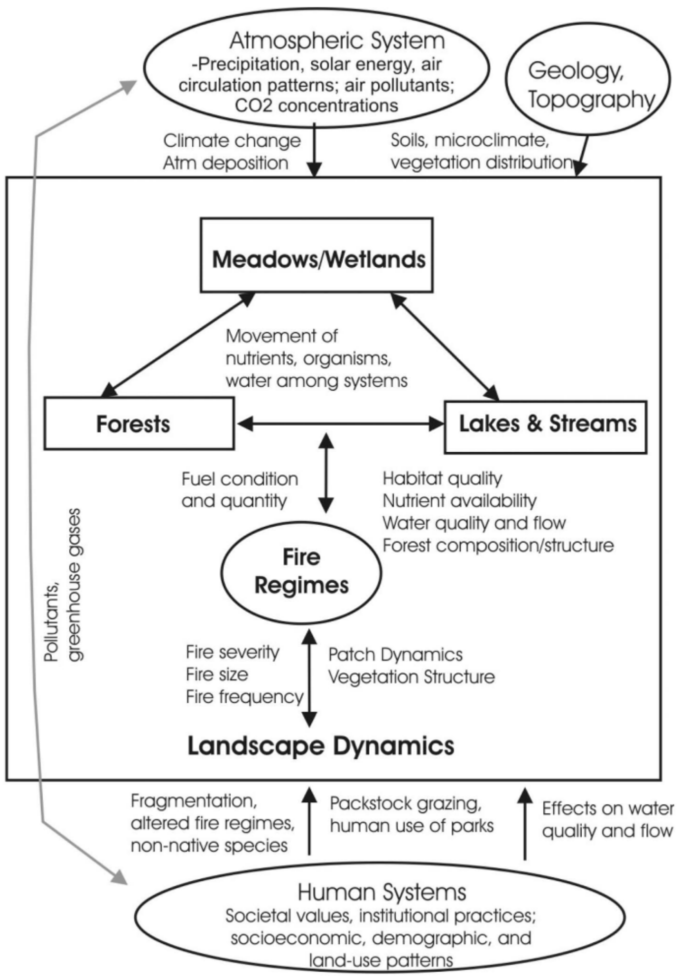


Figure 2. Conceptual model situating climate and atmospheric systems within the context of interacting focal systems in the Sierra Nevada. Main focal systems for SIEN vital signs monitoring are boxed; major drivers are circled (from Figure 2.3, Mutch et al. 2008).

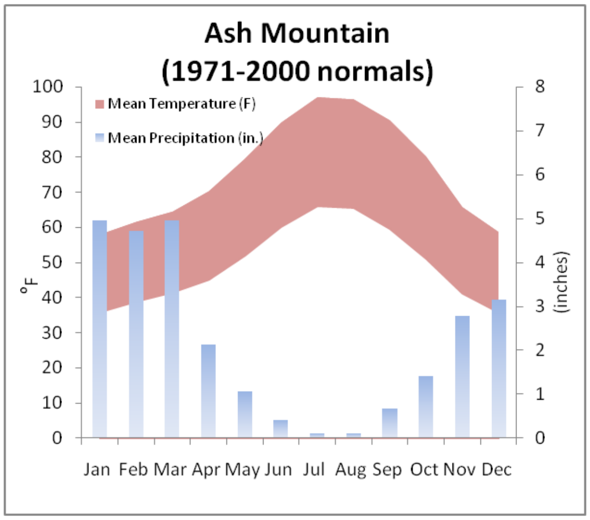
The atmospheric system drives multi-scalar weather and climate phenomena. Climate strongly influences landscape patterns and processes by determining the flux of both energy (solar radiation) and mass in the form of moisture (rain, snow, water vapor). Stine (1996) states that climate has a prominent influence on the following Sierra Nevada landscape components:

- Vegetation (type, biomass, distribution)
- Hydrology (size, distribution, fluctuations, and water quality of lakes and streams)
- Soils (thickness, stability, nutrient capacity)
- Landforms (rates of formation and loss)
- Fire (location, frequency, seasonal timing, intensity and/or severity)

Climate varies across both spatial and temporal dimensions. It affects forest type and other plant community distribution through influence on the soil water balance (Stephenson 1988, 1998). In the SIEN parks, precipitation increases and temperature (and thus evaporative demand) decreases with elevation. The mixed-conifer zone of the Sierra Nevada is sandwiched between low-elevation sites that are chronically droughty, and high-elevation sites that are too cold to be very productive (Urban et al. 2000). Thus, this zone is quite sensitive to climate variability (Graumlich 1993, Swetnam 1993). Predicted potential effects of anthropogenic climate change on the Sierra Nevada are likely to be highly synergistic, affecting a range of physical and biological systems in unpredictable ways (CIRMOUNT Committee 2006).

The climate of the southern Sierra Nevada can be characterized as Mediterranean at lower elevations (warm dry summers and wet cool winters) and boreal at high elevations (cold winters); more precipitation falls as snow than rain due to the large area of land at high elevations. Summers are dry and precipitation predominantly falls in the cool winter season, with snow possible at all elevations of all four park units (Figure 3). The rain/snow transition elevation zone is the zone above which precipitation falls predominantly as snow and typically ranges from 1500–1800 meters (4920–5900 ft). The elevations of this zone are expected to rise with climate change associated warming.

Mean annual temperature at a given elevation generally increases and mean annual precipitation decreases as one moves south from YOSE to SEQU. Consequently, most locations in YOSE are generally wetter than those at comparable elevations in SEKI (Figure 4). SEKI’s eastern high country is drier than both the western side of the high country and the Owens Valley farther to the east. DEPO, which is in a valley, receives less precipitation than its higher elevation surroundings.



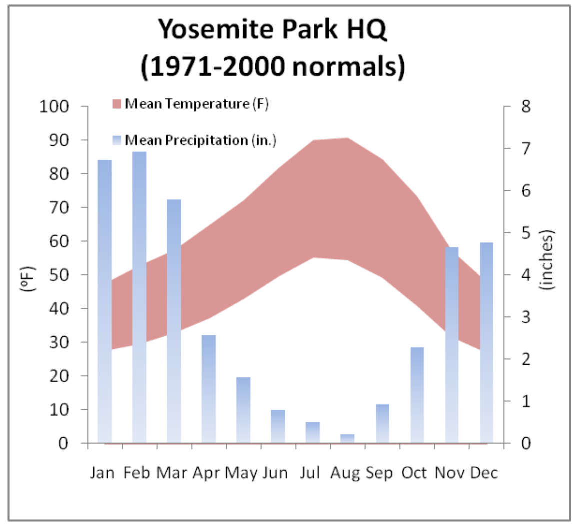


Figure 3. Monthly climate normals for two sites in the SIEN. Average minimum and maximum temperature and average total monthly precipitation for the period of 1971-2000 are displayed for the Ash Mountain (SEKI) and Yosemite Park Headquarters (YOSE). Note the unimodal pattern of seasonal precipitation cycle (wet winter, dry summer). Total monthly precipitation is displayed in vertical bars with highest values observed January through March. Temperature range is displayed as continuous block with highest values seen in July and August. SIEN will utilize English units rather than metric for reporting purposes.

For as long as instrumental weather records exist (since 1895), temperature and precipitation in the Sierra Nevada have exhibited high inter-annual variability. Nonetheless, it is documented that mean air temperature throughout the SIEN has risen since the mid-1970s, and the last 10 years of mean temperatures approach or exceed that of any other decade of temperatures on record. The mean minimum (or nighttime) temperatures have risen more dramatically than mean maximum (or daytime) temperatures. The mechanism for this is unknown but the Sierra Nevada is not alone in experiencing this uneven warming phenomenon (Edwards and Redmond 2011).

Although most climate models are unable to predict changes in the quantity of total precipitation (snow and rain), scientists expect that climate change will result in a widespread reduction in the ratio of snow to rain, which would have large effects on park ecosystems (Knowles et al. 2006). Although snowpack has declined throughout much of the U.S. from 1950-1999, no apparent trend has been observed in the total annual precipitation in the SIEN (Pierce et al. 2008, Edwards and Redmond 2011). Changes have been recorded, however, in the timing of snowmelt runoff as well as the percent of precipitation falling as rain rather than snow (Stewart et al. 2004, Hidalgo et al. 2009, Barnett et al. 2008). The snowpack of the Sierra Nevada acts as a natural reservoir, releasing water slowly through the spring and summer to park ecosystems and users outside the parks, primarily agriculture, in the Central Valley of California (Figure 5). An earlier snowmelt or a shift to less snow and more rain in the winter could result in less water availability to the valley throughout the growing season and California's water managers may need to re-organize their planning for releases from reservoirs. Therefore, California's water managers are particularly interested in snow, precipitation, and temperature trends in the Sierra Nevada.

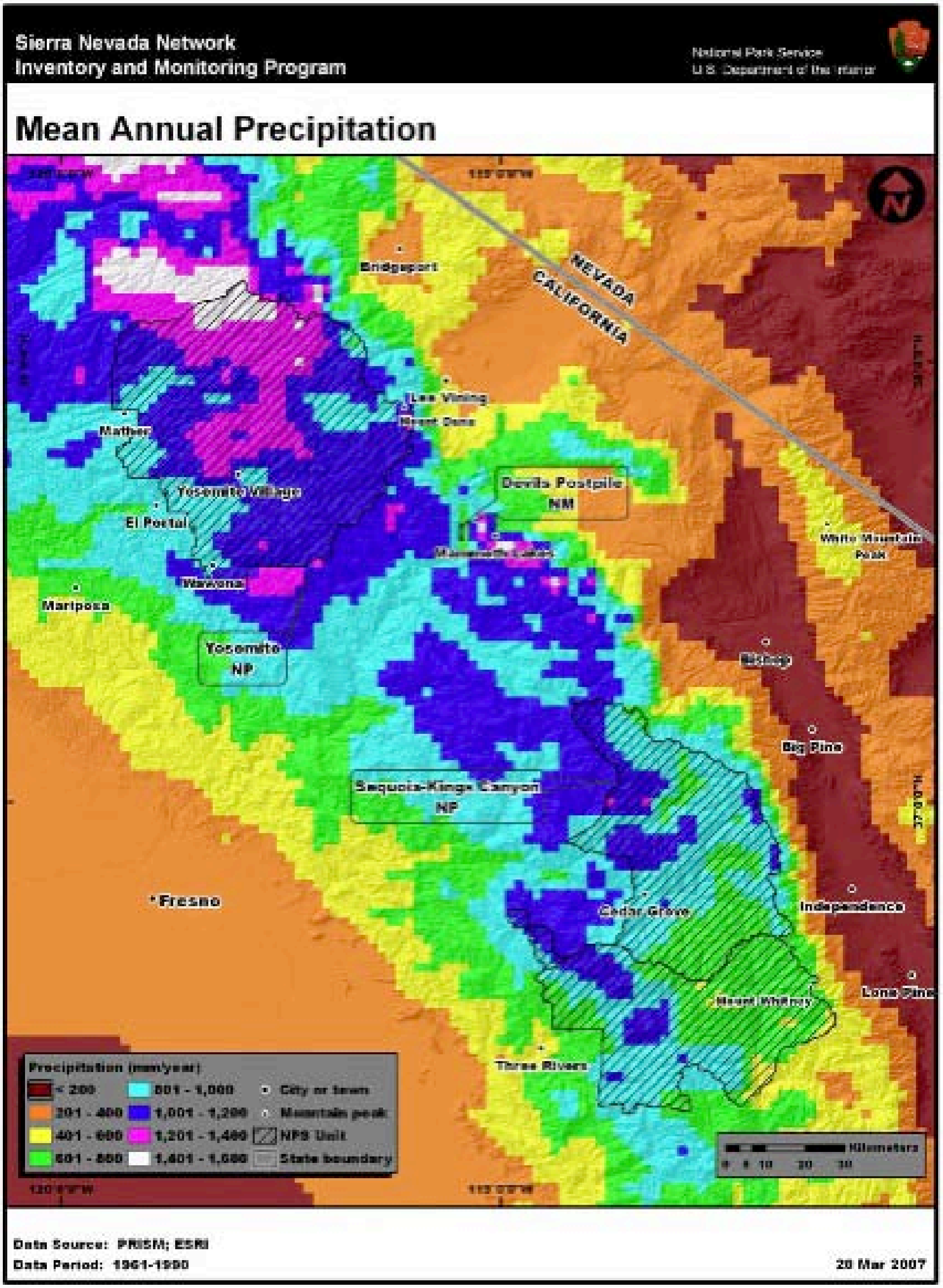


Figure 4. Sierra Nevada Network mean annual precipitation.

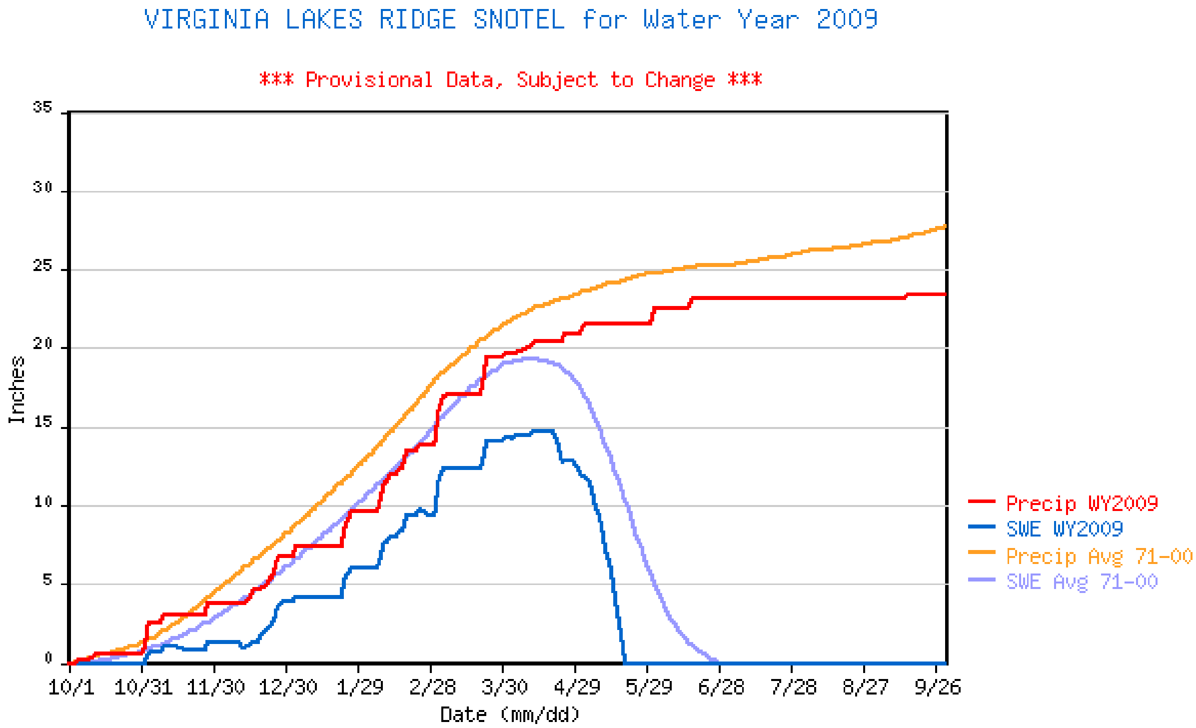


Figure 5. Historic and Water Year (WY) 2009 snow water equivalent (SWE) and precipitation at the Virginia Lakes Ridge Snotel station near the northeast boundary of YOSE. Note the peak snowpack in late March (blue lines) and the snowmelt period through June. Figure from: http://www.wcc.nrcs.usda.gov/snotel/California/california.html.

There are multiple broad-scale climate mechanisms that influence the variability of climate in the SIEN from hemispheric to the local scale. Below is a summary of commonly known causal associations between climate forcing mechanisms at multiple scales and the regional climate and weather that Sierra Nevada parks experience.

Hemispheric Teleconnections

- The El Nino-Southern Oscillation (ENSO) is a sub-decadal oscillation with a 2-7 year cycle and two phases. The El Nino phase is weakly related to wet winters in the Sierra Nevada: cooler, more frequent and longer precipitation events and increased snowfall. The La Nina phase is strongly correlated with dry winters in the Sierra Nevada: warmer, less frequent, shorter precipitation events, but higher flood likelihood. The connection weakens from south to north (stronger in SEKI than in YOSE).

- The Madden-Julian Oscillation (MJO) is a 40–70 day cycle with extended wet periods; warm storms with intense rain – “Pineapple Express” events; rain at high elevation; flooding, melting snow (“rain on snow” events).

- The Pacific Decadal Oscillation (PDO) is a 20 to 30-year cycle with ambiguous influence on the region (Mantua et al. 1997, Edwards and Redmond 2011).

Regional Influences

- The Southwest Monsoon (seasonal cycle) relating to summer heating in nearby desert regions can bring surges of moisture and instability that cause afternoon convective storms from late June to mid-September. During these events, the east slope, more so than the west, experiences thunderstorms, localized and intense precipitation, flooding, high winds, and wildfire ignitions

- Cyclonic Storms (non-cyclic) differentially affect the west slope. Thunderstorms are less likely but these events occur in all months, at any time of day. Cool season storms create widespread, uniform precipitation that increases with elevation.

Local Mechanisms

- Orographic Lifting causes the 5000-8000 ft elevational zone to receive more frequent and abundant precipitation than other elevations due to forced cooling of the rising air.

- Prevailing Westerly Winds combined with the effects of orographic lifting cause passing air masses to lose their moisture as they rise over the Sierra. This creates a rain shadow, leaving the eastern slopes much drier than the western slopes.

Other multi-scale mechanisms

- Drought is a lack of precipitation during the wet season. The most visible manifestation is reduced snowpack and earlier spring.

- Human effects on the Sierra Nevada atmosphere, terrestrial and aquatic ecosystems are thought to interact and compound one another. These effects are being intensely studied by the NPS and its partners. While some linkages are clear, others are complex, nonlinear and difficult to demonstrate.

1.3 Rationale for Monitoring Climate

SIEN selected “weather and climate” and “snowpack” as high-priority Vital Signs for monitoring because they are major drivers of Sierra Nevada ecosystems and they inform and affect park management and visitor experience (Mutch et al. 2008). In order to make informed management decisions within the SIEN parks, it is vital to have an understanding of how climate influences ecosystem components, interactions, patterns, and processes such as water dynamics, nutrient cycling, vegetation community composition, structure and productivity, and animal distribution and activity patterns. By characterizing and understanding climate patterns, it becomes possible to discern the effects of climate from other anthropogenic and non-anthropogenic drivers and stressors on ecosystem components and processes (and Vital Signs).

A local expression of climate change is already being felt in the Sierra Nevada, and scientific models predict the weather and climate will continue to change into the foreseeable future. These changes will provoke significant ecological change and necessitate tremendous human adaptation. Change in runoff amount and timing, movement of the rain/snow transition zone and increases in the rain/snow ratio will critically affect the availability of surface and ground water for biological and human use. A warming climate would change some of the snow to rain in basins where reservoir storage is limited and when a higher percentage of the total annual precipitation is delivered via winter rain rather than snow, runoff will occur months before evaporative demand. Bales and Rice (*in prep*) considered three climate change scenarios, 2, 4 and 6 ^{o}C; and each 2 ^{o}C rise in temperature corresponded to the current average temperature about 300 m lower in elevation. The analysis showed that temperature increases will shift the amount of snow covered area (SCA) as less precipitation falls as snow. The analyses also showed that timing of snowmelt shifts toward earlier in the spring at a rate of -6 to -7 days per degree Celsius. That is, with a temperature increase of 2^{o}C, 50% of the seasonal snowmelt will be depleted about 2 weeks earlier than present. Resource managers in the SIEN parks rely on climate data to document the rate at which these changes are happening and to prioritize management actions such as species relocations or habitat restoration. Further, if changes are detected in plant or animal communities, parks need to have adequate climate data to tease out causation for change, which in turn affects how they will manage for that change. The closer we track the local expression of global change, the more prepared we will be for the resulting changes in the natural landscape upon which we rely for its wealth of tangible and intangible values (Edwards and Redmond 2011).

Basic climate information from the reporting products of this protocol can be used by park visitors and managers for planning activities and by park managers for determining how and when to allocate resources within the parks. For example, the timing of first and last snows can determine road closures and level of access to different areas of the parks. Severe drought and the subsequent increase in fire danger may require parks to close areas or adjust strategies for managing wildland fires within the park. Further, significant changes in climate of the western United States since the last century have been documented, and these changes are predicted to continue, which will require land managers, water districts, and state and local governments to adjust and plan accordingly. These changes include increased temperatures and a shift toward earlier snowmelt timing in the western mountains and forests bioregion (Edwards and Redmond 2011).

Many organizations, such as the National Weather Service (NWS), the Western Regional Climate Center (WRCC), and the California Department of Water Resources (CDWR) collect weather data and operate websites that provide weather forecasts and climate summaries to NPS staff and visitors. The SIEN Climate Protocol will build upon these efforts, filling gaps and compiling figures and summaries from existing agencies into comprehensive reporting products that will be made available to staff and visitors. These include:

- Annual status reports for network parks. SIEN will compile data from these various organizations and sources into a single document and provide concise summaries of the results.

- Climate trend analyses and an interpretation of results that will be especially valuable for parks undertaking future scenario planning.

- Tabular data and additional QA/QC. The underlying tabular data for summaries and figures at some organizations' websites are not easily accessed. A significant value of this protocol is that it will acquire and store tabular data underlying the summary figures and tables. In addition, the quality of these data may be improved through automated QA/QC procedures before storage in the Climate Database and more rigorous QA/QC in preparation for trend analyses. All changes will be documented, and corrected data sets will be stored along with the original. These datasets are then available for covariate analyses of other monitoring or research data.

1.4 Description of Selected Climate Measures

SIEN will report on the following five climate measures which are tightly linked with ecological processes:

- Air temperature

- Precipitation

- Snow-water equivalent (SWE) and snow depth

- Drought

- Streamflow

Temperature

Temperature is a key driver of ecological processes across numerous scales from organismal biology to ecosystem structure in the SIEN and has a significant impact on all of the other climate measures - drought, precipitation, snowpack and streamflow. It can regulate the form in which precipitation will fall. Small temperature increases can alter the elevation of the rain/snow transition, the elevation above which precipitation falls as snow rather than rain. Understanding the relationships and trends in these variables is critical to recognizing potential threats to regional water supplies in California (Dettinger et al. 2004). Although this protocol does not claim to identify the signature of climate change, it can document the local expression of many types of climate variability and trends, which are of direct importance to concomitant variability in other ecological functioning.

Temperature determines the activation and efficiency of enzyme production, and therefore it regulates all aspects of life from microbial activity and plant growth rates to the body size of large predators. For instance, freezing temperatures can directly reduce the population size and number of breeding cycles in bark beetles and temperatures often determine foraging behaviors in birds. At an ecosystem scale, primary productivity and nutrient availability are driven by temperature via its effects on microbial activity (increasing temperatures increase microbial activity and the release of nutrients into the soil). Temperature is also one of the key determinants for tree line in the West because tree growth is limited by the low temperatures found at high elevations.

This protocol will utilize the daily minimum and maximum temperature recorded at the NWS Cooperative Observer Program (COOP) weather stations and at the automated weather station in DEPO (operated by Scripps Institution of Oceanography [Scripps]; herein referred to as DEPO-AWS) for reporting purposes.

Precipitation
Precipitation is measured as the volume and intensity of moisture deposition in liquid form. It may arrive as rain, sleet, snow or hail. Like temperature, precipitation can regulate processes on organismal and ecosystem scales. All organisms require water for growth and metabolism. At an ecosystem scale, precipitation determines nutrient availability via its effects on the weathering of rocks. Increased water availability also increases the rate of decomposition and the release of nutrients in the soil. As a result, wetter areas tend to have increased nutrient availability and productivity compared to drier areas. It is not only the amount of precipitation, but also the timing (e.g., winter vs. summer) and form of precipitation (e.g., snow, fog, and rain) that influences the type of vegetation present in an ecosystem. A major factor in the hydrologic cycle, the temporal and spatial variability of rainfall is recorded through a network of rain gauges.

This protocol will rely on the daily and monthly accumulated precipitation recorded at NWS-COOP weather stations and at the DEPO-AWS. Typically, NWS-COOP stations are visited each morning by an observer who records the volume of water accumulated in a rain-gage, melts any snow accumulated in tube similar to a rain-gage, and adds the volume of each to get the total accumulated precipitation for the previous 24 hours. It would be most common for a NWS-COOP station in SIEN to receive only snow or rain, but occasionally a station could receive precipitation in both forms in one day. The DEPO weather station is not a NWS-COOP station; it is automated. There are two tipping buckets at the station, one that is heated to melt snow and one that is unheated. The data are recorded every five minutes and sent via satellite to Scripps, the operator of the station. SIEN acquires the data from Scripps and uses the 5-minute data to calculate the total monthly and annual accumulated precipitation.

Snow-Water Equivalent (SWE) and Snow Depth
Winter snowfall is among the most dominant drivers of physical and ecological processes in the Sierra Nevada Region. Snow has innumerable direct effects on ecosystems, ranging from soil moisture recharge to strong impacts on winter survival of ungulate populations. Moreover, the land area above the rain/snow transition elevation makes up 91 percent of the land area in SEKI

and 86 percent of the land area in YOSE. A significant amount of this snowpack contributes to public use throughout California and is a primary driver of ecosystem dynamics throughout the parks. Therefore, the monitoring of snowpack evolution and ablation is of paramount importance.

The amount of water contained within a given volume of snow, SWE, is the product of snow depth and density. SWE is typically presented in units of inches and is the depth of liquid water that would result from melting the snow. This is determined manually using a snow tube or by melting the depth of the relevant snow pack (new or total, depending on which SWE is needed). Automated SWE observations are collected by snow pillows, which are envelopes of stainless steel or synthetic rubber, about 4 –6 feet square or rectangular, containing an antifreeze solution. Snow accumulating on the pillows exerts pressure on the solution. Automatic measuring devices in a nearby shelter house convert the weight of the snow into an electrical reading of the snow's water total water volume. By examining SWE and snow depth data together, one can gain an understanding of the amount of moisture contained in snow storms. A winter with many “wet” storms - storms that deliver snow with a high water content, would result in high SWE although the snow depth may be the same as a drier winter. Understanding the types of storms occurring annually in the parks is useful for examining climatic trends.

This protocol will rely on calculated SWE values from snow courses operated by the CDWR as reported on their website, the California Data Exchange Center (CDEC). The CDWR reports SWE for each snow course on the first of each month, January through May. The average SWE for each snow course is provided along with the “basin average” – the average of the courses in each major watershed, compared to the 50-year average. Commonly, the peak snow accumulation occurs near April 1, and the April 1 SWE is the metric most often compared between years.

SIEN will also report the snow depth at NWS-COOP stations. Snow depth can be measured by visually recording the accumulation at a NWS-COOP station with a measuring stick or at a snow course site during a snow course survey. Additionally, snow depth can be estimated through automated systems at snow courses or snow pillows.

Drought

Drought is an integrative climate parameter directly dependent on the measured primary variables precipitation, snow, and temperature. Documenting the circumstances that lead to drought conditions in the park will contribute to the understanding of other ecological observations or management decisions. For example, a given year with higher than average temperatures and lower than average precipitation or snow pack could result in drought. Multiple years with such conditions would lead to more extreme drought and may result in a change in policies regarding fire or changes may be seen in other vital signs such as surface water temperatures, streamflow, or plant productivity. In general, a drought is an extended period of time where an area has a deficiency of precipitation compared to an accepted baseline.

Drought is a normal recurrent feature of climate that occurs in virtually all climatic zones, although its effect differs by region. It is important to recognize that drought is a temporary aberration, different from aridity, which is a permanent feature of climate. Because there is no single, precise definition of drought, its onset and termination are difficult to determine and it is measured and described by a variety of metrics.

Many drought indices incorporate multiple data types into calculations that result in a single value ranging from negative (drought) to positive (wet). Drought indices can be reported at a variety of spatial scales, varying from climate regions to the globe, and at temporal scales ranging from weeks to decades. SIEN will report on conditions within the parks in relation to the drought status reported by the indices at the regional or state wide scale. The Palmer Drought Severity Index (PDSI) is a metric commonly used to determine and report drought throughout the U.S., however, through examination of the drought indices and consultation with the regional climatologist it was determined that the PDSI, which does not incorporate snow water content and is primarily suitable for homogeneous areas, was not a suitable way to look at drought in the SIEN where snow comprises the majority of precipitation. SIEN will report the Standard Precipitation Index (SPI) and Drought Monitor summaries.

The SPI is a metric that calculates the probability of recording a given amount of precipitation (McKee et al. 1993). The probabilities are standardized so that an index of zero indicates the median precipitation amount (half of the historical precipitation amounts are below the median, and half are above the median). The SPI was designed to quantify the precipitation deficit for multiple timescales. These timescales reflect the impact of drought on the availability of the different water resources. Soil moisture conditions respond to precipitation anomalies on a relatively short scale, whereas groundwater, streamflow, and reservoir storage reflect the longer-term precipitation anomalies. Two key distinguishing traits of the SPI are that it identifies emerging droughts months sooner than the PDSI and it is computed on various timescales, but it does not account for the role of heating on drought. SPI can be calculated for retrospective timescales from most immediate (proximate month and season) to sustained (multi-year). The National Climatic Data Center reports SPI for individual regions. The SIEN is located in California Climate Region 5, the San Joaquin Drainage.

The Drought Monitor (http://drought.unl.edu/DM/MONITOR.html) is a synthesis of multiple indices and impacts and represents a consensus of federal (U.S. Department of Agriculture [USDA] and National Oceanic and Atmospheric Administration [NOAA]) and academic scientists (National Drought Mitigation Center at University of Nebraska-Lincoln). The Drought Monitor produces a summary map of drought intensity for the nation and all states each week. Drought intensity is classified based on the PDSI, SPI, soil moisture, streamflow, and other indicators of drought such as vegetation health, groundwater levels, and SWE. It is on a scale ranging from abnormally dry (D0) to exceptional drought (D4). While the monitor provides excellent summary information on broad-scale conditions, local conditions (e.g., the park scale) may vary. We will include automated figures from the Drought Monitor and SPI for annual status reports. The drought indices mentioned here will not be used to make inference to specific locations in the SIEN but act as regional indicators of condition.

Streamflow

Streamflow, also referred to as discharge, is a measure of the volume of water that flows past a given point in a river or steam, recorded in cubic-feet per second. The mean daily streamflow displayed for the year in graphical format is a hydrograph. Streamflow in unregulated, unaltered watersheds is highly correlated with annual precipitation, spatially integrates precipitation over a catchment area, and can often reveal patterns that measurements from individual weather stations cannot. Peak streamflow refers to the highest instantaneous streamflow measurement for a given year. Stream stage and discharge are the primary indicators of water yield and demonstrate responses to precipitation, temperature, snow melt, and runoff throughout a watershed. For example, a wintertime spike in the annual hydrograph of a streamgage located below the rain/snow transition would indicate a storm with more rain than snow at higher elevations. Tracking the frequency and types of storms through examination of streamgage hydrographs can help explain ecosystem responses and help park staff or state water managers better plan for the future.

This protocol will rely on streamgage data from the U.S. Geological Survey (USGS) and figures from the USGS website the National Water Information System (NWIS). Although one of the streamgages used for reporting is operated by the electric company Southern California Edison, the records are reviewed and published by the USGS.

Spatial Climate Data

Of the climate measures listed above, air temperature, precipitation, snow depth, and streamflow are reported at the point-based level. This means that summaries can only be made for individual climate or streamflow stations. Because large portions of the SIEN are inaccessible and do not have weather stations, SIEN will also utilize spatial climate data to better understand conditions across the network.

The spatial climate data are gridded data sets that typically include temperature, precipitation, or SWE for a given period across a region. Most are created by statistically interpolating data values from irregularly spaced station locations to a regular grid (Daly et al. 2002). Because such data can be linked to Geographic Information Systems (GIS), incorporated in modeling exercises, and used to estimate climate for areas where weather stations are sparse or absent, they provide an essential resource to park managers and scientists. An up-to-date gridded surface climate dataset can spatially represent the status of climate zones, provide details for important management areas, and put park climates in perspective of the surrounding region’s climate. We will use data developed by PRISM (Parameter-elevation Regressions on Independent Slopes Model - http://www.prism.oregonstate.edu/). PRISM is based on local regression that accounts for spatially varying elevation relationships, effectiveness of terrain as barriers, terrain-induced climate transitions, and cold air drainage and inversions (Daly 2006). This approach has a long history of use in the western United States, and it has been shown to provide highly robust products in a wide variety of studies (Daly et al. 2008).

1.5 Monitoring Goals and Measurable Objectives

While there are numerous additional measures of climate collected in the SIEN parks, such as wind speed and humidity, this protocol will focus on the key parameters temperature, precipitation, snowpack, drought and streamflow. Below, we state our monitoring goals and objectives and further define the key parameters. Finally, we describe our monitoring approach and reporting products.

1.5.1 Monitoring Goals

Primary monitoring goals for this protocol are:

- To provide SIEN parks with summaries of the status and trends observed in key climate parameters at various temporal scales (daily, monthly, and annual) at selected weather stations.

- To describe variations and detect changes in key climate parameters relative to established baselines.

Secondary goals for the protocol are:

- To provide access to climate datasets for use as a covariate in analyses of other vital signs (e.g., high-elevation forests, rivers, lakes monitoring) or other research projects.

- To support cooperating agencies in maintaining continuous, high-quality weather data records at the stations they operate within SIEN parks, particularly at benchmark stations (e.g., COOP and Scripps), and maintain working relationships with other support personnel, including the various operators of monitoring stations. Although SIEN cannot assume primary responsibility for operation and maintenance of weather stations in and around SIEN parks, we will provide support by assuring that responsible agencies are performing regular maintenance of stations. In addition, we will perform checks of data completeness each year, and attempt to resolve issues to ensure that data are consistently collected.

- To position the SIEN Physical Scientist as a resource and central point of contact for park staff and cooperating agencies in need of climate information or assistance.

1.5.2 Monitoring Objectives

<u>Temperature</u>

1) Summarize monthly and annual temperature at the local scale (e.g., selected stations)

   Metrics:

- Minimum, maximum, and mean monthly and annual temperatures and departures from an established 30- year normal at selected weather stations. Required data include temperature records from NWS-COOP stations and the DEPO-AWS.

- Departure from a 30-year normal for minimum, maximum and mean monthly temperature for the SIEN climate region (PRISM data).

2) Analyze and report on inter-annual variability and trends in daily, monthly and annual temperature

   Metrics:

   - Variation and change in the probability of extreme temperature events and other features of daily frequency distribution.

   - Spatial and temporal variability and trends in minimum, maximum, and mean monthly and annual temperatures

   - Interpretation of results from the perspective of regional coherence and hemispheric teleconnections (ENSO, PDO, MJO).

<u>Precipitation</u>

1) Summarize monthly and annual accumulated precipitation at the local scale (e.g., selected stations)

   Metrics:

   - Monthly and annual accumulated precipitation and departures from an established 30-year normal at selected weather stations. Required data include precipitation records from NWS-COOP stations and the DEPO-AWS.

   - Departure from a 30-year normal for monthly and annual accumulated precipitation for the SIEN climate region (PRISM data).

2) Analyze and report on inter-annual variability and trends in daily, monthly and annual precipitation
   Metrics:

   - Variation and change in the probability of extreme precipitation events and other features of daily frequency distribution

   - Spatial and temporal variability and trends in minimum, maximum, and mean monthly and annual precipitation

   - Interpretation of results from the perspective of regional coherence and hemispheric teleconnections (ENSO, PDO, MJO)

Snow water equivalent and snow depth

1) Summarize monthly snow water equivalent and snow depth at the local and watershed scale (i.e., selected stations and watershed averages)

   Metrics:

   - Average SWE of all snow courses within each major watershed and departure from established 50-year normals. Required data include SWE statistics by watershed from CDWR

   - Monthly snow accumulation and departures from average at selected stations. Required data include snow depth at NWS-COOP stations.

2) Analyze and report on inter-annual variability and trends in daily, monthly and annual snow depth and snow water equivalent

   Metrics:

   - Variation and change in the probability of extreme snow events and other features of daily frequency distribution.

   - Spatial and temporal variability and trends in monthly snow water equivalent and snow depth.

   - Interpretation of results from the perspective of regional coherence and hemispheric teleconnections (ENSO, PDO, MJO).

   - Number of days with snow cover, areal extent, and timing of snowmelt

   - Frequency of extreme snow cover/SWE events beyond a defined threshold

   - Interpretation of results from the perspective of regional coherence and hemispheric teleconnections (ENSO, PDO, MJO)

Drought

1) Summarize monthly and annual drought status at the regional scale.

   Metrics:

   - Frequency and duration of drought beyond established thresholds as defined by major drought monitoring programs – Drought Monitor and NOAA Standardized Precipitation Index.

2) Analyze and report on inter-annual variability and trends in drought from the perspective of regional coherence and hemispheric teleconnections (ENSO, PDO, MJO).

Streamflow

1) Summarize mean daily, monthly and annual streamflow at the local scale (i.e., selected stations)

   Metrics:

   - Mean daily, monthly, and annual discharge and departures from the average. Required data include USGS hydrographs for selected streamgages.

   - Peak discharge timing and departure from the average

2) Analyze and report on inter-annual variability and trends in daily, monthly and annual streamflow

   Metrics:

   - Intra- and inter-annual variability and trend analyses and interpretation in light of SWE, drought, precipitation, seasonal temperatures, and hemispheric teleconnections (ENSO, PDO, MJO)

1.6 Monitoring Approach and Reporting Products

Constrained by limited funding and faced with difficult decisions on the allocation of funds among vital signs, SIEN and park staff decided that it was not logistically or financially feasible to add new weather stations throughout SIEN parks, despite significant gaps in coverage, due to the prohibitive costs of scoping, installing, and maintaining new weather stations. We determined that the best application of our limited budget is to harvest, analyze, and report on data from established weather monitoring programs that provide consistent, long-term, and high-quality climate records for our parks and region (Appendix A). This approach, as documented in the SIEN protocol narrative and SOPs, is patterned after the Rocky Mountain Climate Protocol (Rocky Mountain Climate Working Group 2010), which resulted from a multi-network collaboration among ecologists, data managers, and climate professionals. Through multiple meetings, these professionals determined the most efficient and effective methods to assess climate on an annual and periodic basis and report the findings in a useful format to park staff. The meetings resulted in the reports, A Framework for Climate Analysis and Reporting for Greater Yellowstone (GRYN) and Rocky Mountain (ROMN) Networks (Kittel et al. 2010) and The Development and Analysis of Climate Datasets for National Park Science and Management (Kittel 2009), upon which the Rocky Mountain Climate Protocol is based. We recognize these reports and the protocol as the most collaborative and comprehensive guidance for our climate reporting protocol at this time. Because our monitoring approach lacks a probabilistic sampling design, we recognize that we are unable to use station data to make statistical inference about status or trends beyond the station-level.

The protocol will focus on two types of data: (1) point data from weather stations, streamgages, and snow courses distributed across the SIEN and surrounding region that can provide daily measures of parameters such as temperature and precipitation; and (2) gridded data sets from PRISM that provide modeled estimates of climate parameters for the continental United States. The PRISM modeled precipitation estimates incorporate point observation data, a reliable digital elevation model (DEM), and expert knowledge of complex climatic variables that result in high resolution, continuous, digital grid estimates of total annual precipitation.

In the future SIEN may incorporate data from other stations such as Remote Automated Weather Stations (RAWS) and Air Quality Monitoring Stations. At this time SIEN will utilize data from those stations listed in Chapter 2.

To provide a product useful for all vital signs and parks it is important to acquire, process, manage, and deliver climate information at a variety of different spatial and temporal scales. When feasible, data will be provided at scales necessary for understanding changes in other vital signs (e.g., watershed scales).

Existing data sets from NWS-COOP stations and USGS streamgages will be acquired and archived in the I&M Enterprise Climate Database (hereafter referred to as the Climate Database), a SQL Server Database that has been developed by the National Inventory and Monitoring Division (IMD) for this specific purpose. We will summarize, analyze, and report on climate data at scales relevant to the parks via several reporting products. SIEN will publish and share this protocol, standard operating procedures, and scheduled reports in the NPS Integrated Resouce Management Applications (IRMA) data store. Published climate status reports and climate trend reports for network parks will list and describe data sources used, including data quality reviews and results, and how to access both original and quality-checked data sources.

1. *Annual Climate Status Report*

   This report will provide a descriptive summary and analysis of selected climate parameters of the past year. The purpose of each annual status report is to support park science, management planning, and interpretation/education on an annual basis. Reports will use English units (e.g., °F) rather than metric (e.g., °C) because English units are the reporting standard and are used by the Climate Database.

   After undergoing cursory QA/QC procedures by the responsible agencies, data for the calendar year and water year are available in late March or April of the following year. In April, data from the previous calendar year will be uploaded to the Climate Database and available for use in the status report or covariate analyses associated with other research or monitoring. Status reports will build on climate trend reports and put the year's climate in context of longer-term variability patterns and trends.

2. *Climate Inter-annual Variability and Trend Report*

   The climate trend analysis report will be produced on a 5 year cycle and will present rigorous analyses of inter-annual variability, long-term historical trends, and teleconnections with hemispheric climate patterns (e.g., the PDO).

   The protocol's small annual budget allocation will used towards an agreement or contract with a climatologist to produce a trend report on a 5-year schedule. Because many I&M Networks will be adopting this same approach towards climate monitoring (e.g., data harvesting and analysis) and using the same database, it may be efficient and economical to combine resources and collectively fund a climate expert to analyze each network's harvested data and produce trend reports for individual networks. Trend analyses require the development of specialized methods and models for data filling and QA/QC which can be time consuming. Efficiency may be achieved by contracting an individual to develop and employ these methods and models and apply them for multiple I&M networks at one time.

3. *SIEN Website*

   SIEN will maintain a climate monitoring webpage on the network website that will provide links to climate information, reports, and data that are relevant to the parks. The website also hosts a map of all weather stations operating in the SIEN parks and hyperlinks to the associated data.

1.7 Future Opportunities

SIEN is exploring the feasibility of conducting more detailed trend analyses on snow-related parameters. Through a cooperative agreement with researchers at UC Merced (Roger Bales and Bob Rice), SIEN is exploring methods that use ground-based and remotely-sensed data to evaluate status and trends in snow-covered area, snow water equivalent, and depletion rates and timing at the watershed scale (Bales and Rice *in prep*). The first step in this process is to establish repeatable methods, which is contingent on successful completion of the initial 2000-2009 trend report and documentation of methods. In the future, the second step would be to secure support and funding, potentially through various climate change-related initiatives.

2.0 Sample Design

Relevant SOPs:

- SIEN SOP 1: Station Selection

2.1 Climate Zones

A climate zone is an area that has spatially and physically coherent regimes – a set of features and processes that distinguish its climate, and is subject to the same broad-scale atmospheric and adjacent ocean circulation dynamics throughout the area (Kittel 2008). Typically, climate regions correspond closely to biotic regions. Many national and state climate monitoring networks, such as the National Ecological Observatory Network (NEON), have divided the U.S. into large climate regions. All SIEN parks fall into NEON climate domain #17 and within National Climatic Data Center (NCDC) California Climate Division #5: San Joaquin Drainage (Figure 6). The NCDC is a branch of the National Oceanic and Atmospheric Association (NOAA) that has identified divisions that are statistically-determined areas with homogeneous climate (NOAA 2011).

This protocol treats the entire network as a single homogeneous climate zone for analysis. Not only is this logical because all SIEN parks reside on the western slope of the southern Sierra Nevada mountain range, a relatively homogeneous physiological unit, but explicit evidence for this approach comes from recent correlation analyses, various classification schemes, other research, and data. On a more tangible level, weather stations within SIEN parks have demonstrated (a) similar patterns of intra-annual (seasonal) variability in precipitation and temperature (mean temperature and diurnal temperature range), and (b) similar patterns of inter-annual (year to year) dynamics in precipitation and temperature (Edwards and Redmond 2011). In SIEN reports, results will be displayed for all stations together. Anomalies between stations may indicate a need for more in depth examination of the data or processes that are at work in specific areas.

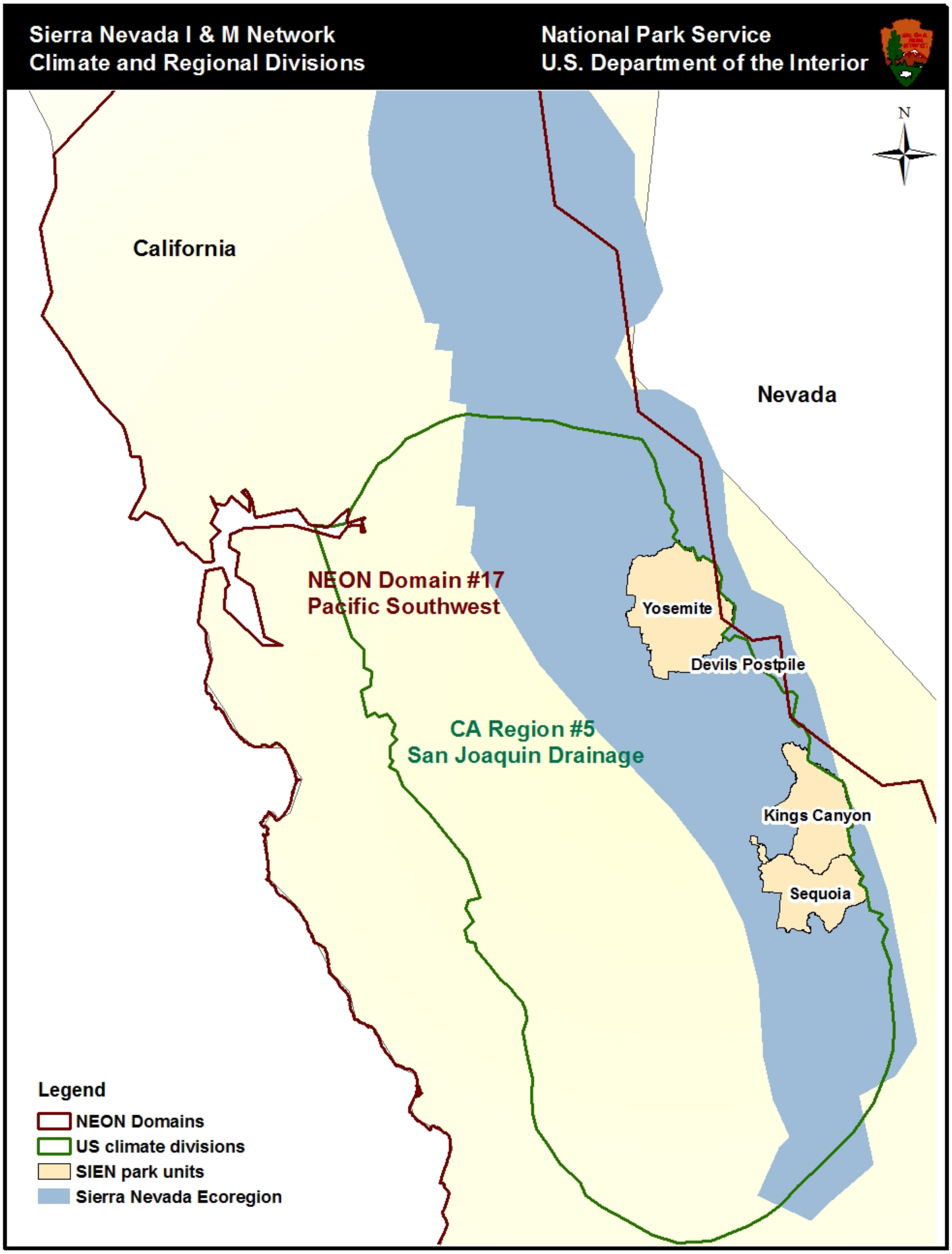


Figure 6. The location of SIEN parks within selected climatological, biological and/or physical landscape classification systems. All SIEN parks lie west of the Sierra Divide, and functionally, the crest is an appropriate place for a climate zone division. Note that the U.S. climate division (CA Region #5) properly tracks the Sierra Divide, whereas the interpolated boundary of the NEON Domain (#17) does not.

2.2 Selection of Sampling Locations (Weather Stations, Snow Courses, Streamgages)

A sampling design specifies how data are to be collected over time and space. Ideally, to describe climate within each network and have strong inference to the area of interest, probabilistic sampling would have been used to locate sampling locations (i.e., weather stations and snowcourses). However, such stations are expensive to install, operate, and maintain, and SIEN has only a small budget for this monitoring effort. In addition, other agencies have already invested time, money, and expertise to determine appropriate locations for stations, have installed, and currently operate these stations. Hence, this protocol relies on data from stations installed by other agencies. This approach is desirable because many existing stations have long periods of record which can be used to place results into a historical context.

Although we will not use a probabilistic sampling design, we apply specific criteria and expert opinion to select a judgment sample of existing weather stations to obtain as much high quality data as possible. Because we are not able to select spatially-distributed stations with a probabilistic design, we recognize that we will not be able to make network-wide inferences when trend analyses are conducted using station data. Grid-based PRISM data will be used in annual reports to display approximate climate conditions across the parks, although PRISM does have some limitations. For example, PRISM models are based on station data. In SIEN, stations are sparse and not evenly distributed, and park landscapes and associated climatic conditions are complex. These conditions reduce the ability of the PRISM models to accurately represent conditions in all areas of the parks. In SIEN reports, all figures created from PRISM data will include a statement about the limitations of these data.

2.2.1 Existing Station Types and Data Sources

Numerous climate, weather, and hydrologic monitoring programs operate in the SIEN, each with a unique type of monitoring station. SIEN will use and report on data from several, but not all of the programs. We refer to the stations that will be reported in the annual status reports as “benchmark stations”. A list of relevant monitoring programs is included here along with the types of data that will be used from each program (Table 2). SIEN will use a combination of station- and grid-based data for reporting purposes. Although only data from benchmark stations will be used for annual reports, SIEN will track the operation and location of other stations because the data may be useful for filling data gaps at benchmark stations and may be a valuable data source for modeling efforts undertaken by other agencies or researchers that may contribute to the understanding of climate processes and trends in the SIEN.

Most stations are located where they are accessible, thought to be representative of an area, and clear of factors that could affect measurements (i.e., tree canopy or waterfalls). A SIEN-sponsored inventory (Davey et al. 2007) identified all weather stations in or near SIEN network parks and this list is continually updated by SIEN staff to reflect stations that have been added, moved, or shut down (Appendix B).

The criteria for stations included in our annual status reports are that 1) data handling procedures are known and rigorously followed, 2) the subsequent data quality is high, and 3) the stations have a complete and long period of record. To summarize, for annual status reports, we will only include precipitation and temperature data from NWS-COOP stations within YOSE and SEKI and the DEPO-AWS (Figure 7). Although the DEPO-AWS does not meet all of our selection criteria because it does not have a long record, it is the best representation of climate at DEPO. The nearest stations with extensive records are nearby at Mammoth Mountain. However, due to the geographic differences between DEPO and the location of the existing stations on Mammoth Mountain, they do not accurately represent DEPO climate. Because the DEPO-AWS does not have a long period of record, there will be no comparisons to a historical "normal" until at least 10 years of data have been collected.

Data from more than 90 snow courses within the SIEN watersheds will be used in the annual reports (Figure 8). We will rely on data summaries for the stations that are calculated and posted by the CDWR Snow Survey Program on the California Data Exchange Center (CDEC) website. The number of courses surveyed and reported varies slightly each year. SIEN will acquire, format and archive the tabular data from the CDWR snow courses locally in a Microsoft Access database because these data are not appropriate for the Climate Database. Although some snow courses also collect precipitation and temperature data, SIEN will not use that data for reporting purposes because the stations do not meet our quality control requirements

Many stations associated with programs in Table 2 meet portions of criteria listed above, but do not meet all conditions. For example, many RAWS stations have a long period of record, but commonly have data gaps of years or months which make it impossible to calculate annual or 30-year averages. Further, the maintenance of the RAWS stations and QA/QC of the data varies from station to station. In comparison, these procedures are consistent and documented for the NWS-COOP stations that we have chosen for use in our reporting.

2.3 Selection Criteria

The following are minimum requirements for acquiring climate data. Additional criteria are detailed in "SOP 1: Station Selection," which is based on guidance provided in Gray (2008) and Kittel et al. (2010) to select data sources to support climate monitoring analyses:

1. Data are from an established program
2. Data are from stations within the area of interest, the park boundary, or the climate zone
3. Data are certified for release as final
4. Data are in digital format
5. Literature supports that the measurement/parameter (e.g., temperature, precipitation) has an effect on resources within and around the park
6. Data are freely available to the networks and park units

In addition, the following documentation must accompany and be stored with the data or be available and accessible at all times via an apparent link from the locally-stored data:

1. Collection/processing procedures
2. Quality assurance and control procedures, including all changes or additions to raw data values, such as data set normalization or extrapolation to populate missing values
3. Station location, setting, and history including instrumentation specifications and changes to instrumentation

Table 2. List of organizations that collect climate data in or near the SIEN parks. Bold font indicates data that will be used by SIEN for reporting purposes. Refer to Appendix A for detailed descriptions of these programs.

Organization & Program	Program Acronym	Number of stations in the SIEN parks	Data to be used by SIEN	Spatial scale	Temporal scale
Oregon State University (OSU) - PRISM Climate Group	**PRISM**	**NA**	**Temperature, precipitation**	**4 km grid**	**Monthly**
National Weather Service (NWS) - Cooperative Observer Program	**COOP**	**7**	**Temperature, precipitation, snow depth**	**Local stations**	**Daily, Monthly, Annual**
Cal Dept of Water Resources (CDWR) Snow Survey Program	**—**	**37 (+56 outside the park used for reports)**	**SWE**	**Local snow courses, snow pillows and watershed averages**	**Monthly and April 1**
U.S. Geological Survey (USGS) - Streamgage	**----**	**3 (+1 outside SEKI used for reports)**	**Streamflow**	**Local stations**	**Daily**
Scripps Institution of Oceanography (SIO)	**----**	**1**	**Temperature, precipitation**	**Local station at DEPO**	**Daily, Monthly, Annual**
The Drought Monitor (multi-agency) and the National Climatic Data Center's Standardized Precipitation Index (SPI)	**----**	**NA**	**Standardized Precip Index, Drought Intensity**	**Regional**	**Monthly**
NOAA – Snow Data Assimilation System	SNODAS	NA	Snow-water equivalent (SWE)	1 km grid	Monthly
Remote Sensing Data. Multiple sources from: http://rsd.gsfc.nasa.gov/rsd/RemoteSensing.html	----	NA	NA	Multiple scales	----
Natural Resource Conservation Service (NRCS) - Snowpack Telemetry Program	SNOTEL	Multiple stations north of YOSE boundary – none in parks	NA	Local stations	----
Citizen Weather Observer Program	CWOP	1	NA	Local stations	----
Climate Reference Network	CRN	1	NA	Local stations	----

Table 2. List of organizations that collect climate data in or near the SIEN parks (continued).

Organization & Program	Program Acronym	Number of stations in the SIEN parks	Data to be used by SIEN	Spatial scale	Temporal scale
Desert Research Institute	DRI	1	NA	Local stations	----
Merced Irrigation District	MID	1	NA	Local stations	----
National Park Service	NPS	1	NA	Local stations	----
Remote Automated Weather System	RAWS	11	NA	Local stations	----
University of California – Santa Barbara (UCSB)	----	3	NA	Local stations	----
US Army Corps of Engineers (USACE)	----	5	NA	Local stations	----
Air Quality Stations					
California Air Resources Board (CARB)	----	4	NA	Local stations	----
Environmental Protection Agency - Clean Air Status and Trends Network	CASTNet	1	NA	Local stations	----
NPS Air Resources Division Gaseous Pollutant Monitoring Program	GPMP	2	NA	Local stations	----
National Atmospheric Deposition Program (NADP)	----	2	NA	Local stations	----
Interagency Monitoring of Protected Visual Environments (IMPROVE)	----	1	NA	Local stations	----
Mercury Deposition Network (MDN)	----	1	NA	Local stations	----

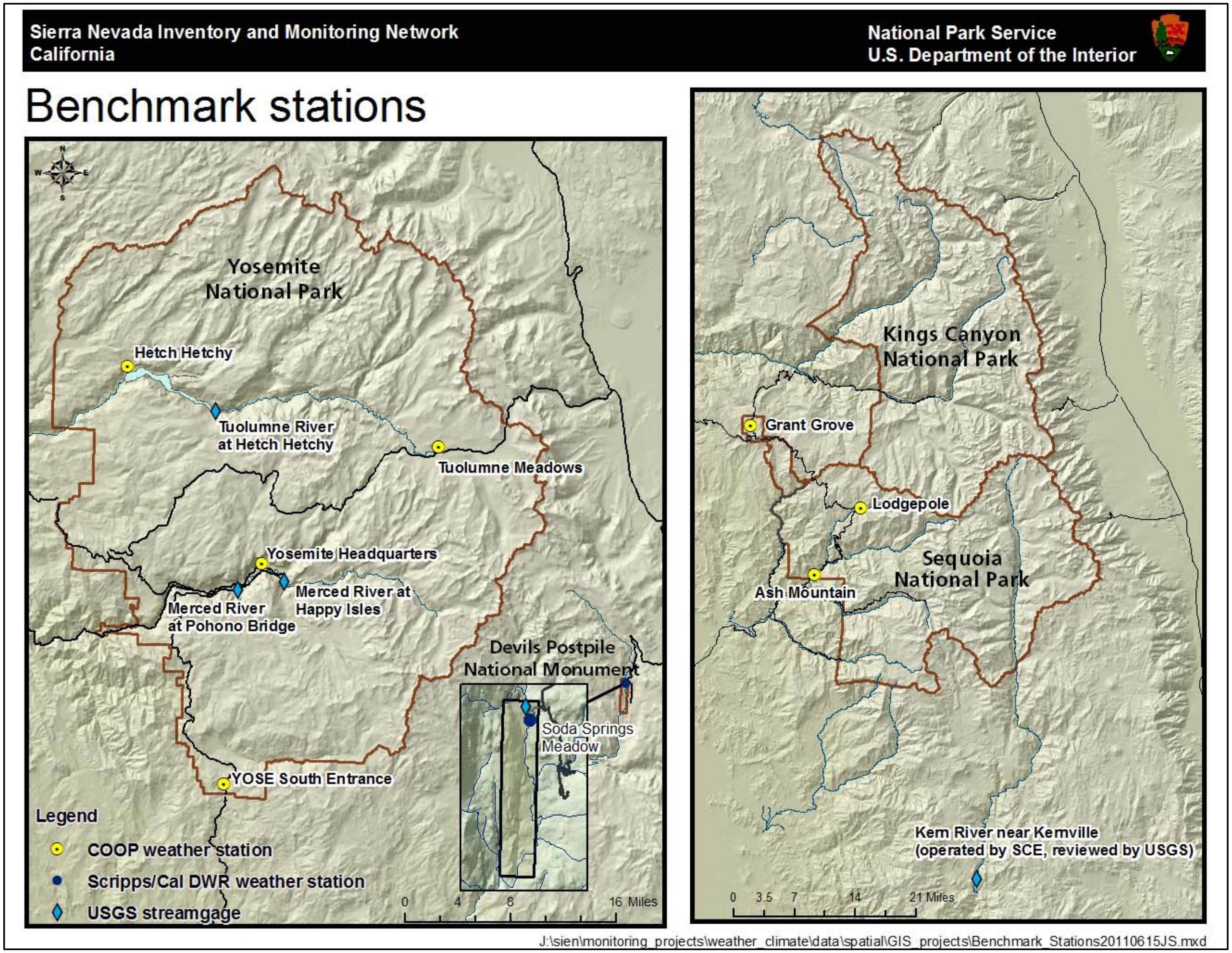


Figure 7. Map of the NWS-COOP weather stations, the automated weather station (AWS) operated by Scripps in DEPO, and USGS streamgages included in annual reports**.**

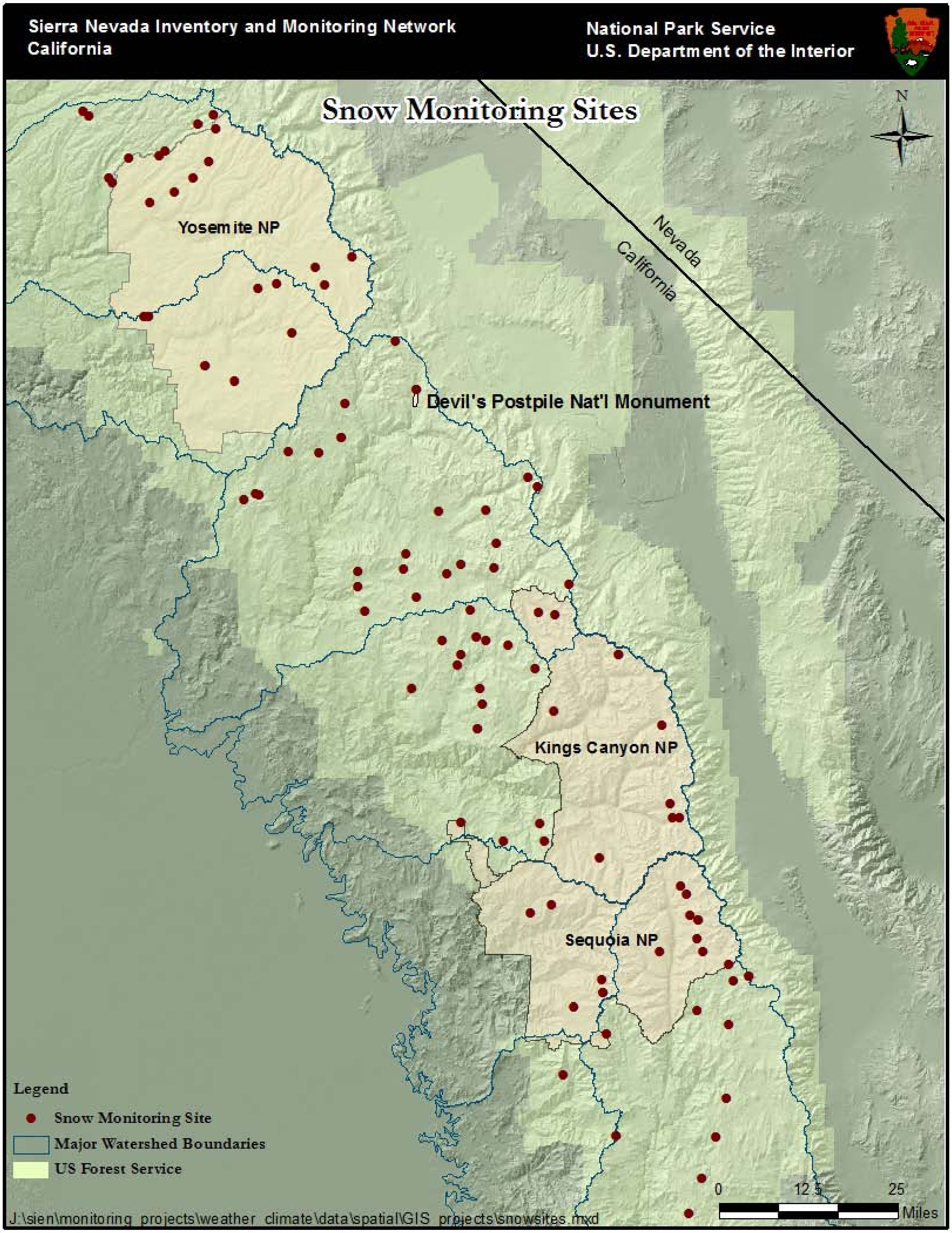


Figure 8. Map of CDWR snow monitoring sites (snow courses and snow pillows) in the Sierra Nevada Network. Although some CDWR snow monitoring sites collect precipitation and temperature, SIEN will not use those data for reporting.

3.0 Data Management, Acquisition, Quality Control and Processing

Relevant SOPs:

- SIEN SOP 2: Quality Control
- SIEN SOP 3: Tabular Data Acquisition and Processing
- SIEN SOP 4: Processing Geospatial Data for Climate Status Reports
- SIEN SOP 5: Annual Status Reports
- SIEN SOP 6: Trend Analysis and Reporting
- IMD SOP: Requesting Essential Parameter Values
- IMD SOP: Submitting Data for Upload
- IMD SOP: Connecting to the I&M Climate Database
- IMD SOP: Climate Grid Analysis Tools (CGAT)
- IMD SOP: Climate Database Schema

3.1 Data Management

The data management system that supports the climate protocol is designed to ensure that analyses and reporting are based on standard and documented data sources from credible observation programs. Data management helps ensure that all data used in analysis and reporting are adequately referenced, easily accessible and of the highest quality.

3.2 Roles and Responsibilities

Successful data management involves overlapping and shared responsibilities among all project staff over the entire life cycle of the data and products collected and generated as part of the project. Project personnel collectively develop, follow, and document standards for information needs, data quality, analytical inputs and outputs, data processing procedures, and reporting requirements to meet the stated objectives of the project.

3.3 Data Standards

Where we have the ability to do so, we intend to meet the following standards for climate data as a basis for scientific credibility: (1) qualitative accuracy, (2) completeness, (3) consistency, (4) precision, (5) timeliness, (6) uniqueness, and (7) validity. Since this protocol relies on data harvested from sources other than our own direct collection, we will use established data providers whose data sources are documented relative to these listed standards. Data qualifiers are normally reported with results to inform readers of important restrictions or limitations to the data. Using reliable sources and carrying forward such qualifiers to the reporting audience allows us to avoid situations where the reporting audience could potentially misconstrue the levels of precision and inference conveyed by reported data. If project staff or report readers identify questionable results or become aware of changes to source data after a report is made, an assessment by project staff will determine whether it is necessary to re-do any or all of the analysis and reporting.

All necessary non-spatial data are acquired in or converted to a common and consistent format for analysis, reporting, storage and distribution. For example, some data are manipulated and formatted for presentation in Microsoft Excel. Geospatial data are formatted for use in Department of the Interior and NPS standard Environmental Systems Research Institute (ESRI) ArcGIS applications. Reports and other distributed products are formatted as Adobe PDF format. Derived data are treated identically. Each derived and developed data set or product meets the standards for format, quality, organization, and maintenance as set forth in the applicable SIEN or NPS data management plans (Cook and Lineback 2008, NPS 2008a).

3.4 Data Sources

The following types of data will used by SIEN to produce annual status reports:

1) Raster data obtained from PRISM

2) Tabular data that is acquired by IMD and archived in the Climate Database

3) Tabular data for DEPO-AWS that is formatted by SIEN for upload and storage in the Climate Database

4) Derived data (data that have been processed into figures or summary statistics) that are available through other websites (i.e., USGS streamgage hydrographs and CDWR snowcourse monthly summary statistics)

In Chapter 2 – Sample Design, we identify the numerous sources of climate data available. As previously mentioned, not all data sources meet our acquisition criteria; therefore, we have selected a subset of these sources for our reporting purposes. These sources are listed in Table 3 along with the data used from each source, and how we will acquire the data. More detailed instructions for acquiring data are included in "SIEN SOP 3: Tabular Data Acquisition and Processing". As evidenced in the table, only NWS-COOP data will be acquired by IMD and stored in the Climate Database. The SIEN data manager will also format and upload DEPO-AWS data to the Climate Database according to "IMD SOP: Submitting Data For Upload" (Frakes 2012b). Although it would be ideal for all data to be stored in a common database, it is necessary to utilize climate data from many organizations to produce comprehensive reports and not all organizations are structured such that their data can be easily acquired for storage in a SQL Server database.

Some datasets reflect point-based conditions or observations (e.g., COOP), while others represent a defined spatial extent at a specified cell resolution (commonly known as "grid" or raster data (e.g., PRISM)). Normally, project staff will acquire high-quality, well-documented data from each system once per year. During the intervals between annual data downloads, park and network staff can obtain climate data using the links on the SIEN website. Data providers, formats, and content are expected to change through time, and project staff will watch for and adapt to changes in data availability and data acquisition and processing methods that may require more or less time and expertise in a given year or reporting cycle. The SIEN strategy is to focus on data acquisition, processing, analysis, and reporting for those climate metrics and indices that directly support stated monitoring objectives, especially for status reports. Further, SIEN is committed to ensuring that all data collected meet the data requirements listed in Chapter 2. Optional data sources, climate measures, and analyses can be incorporated as resources allow.

It should be noted that the IMD can acquire snow course data, but only for those courses that are managed by the Natural Resources Conservation Service (NRCS). Throughout the Western U.S., most snow courses are operated and the data is managed by the NRCS National Water and Climate Center. Several snow courses in the Sierra Nevada are managed by NRCS, but all courses within SIEN parks are managed by CDWR rather than NRCS. The data are stored and served through the California Data Exchange Center (CDEC). While SIEN will acquire the tabular data from CDEC and store it locally in an access database, we will primarily rely on the summary statistics calculated by CDWR that are provided on their website.

Devils Postpile Soda Springs Meadow Station Data

There are sensors belonging to both CDWR and Scripps on the DEPO Soda Springs meteorological station. CDWR data is available through the CDWR CDEC website (http://cdec.water.ca.gov/cgi-progs/staMeta?station_id=DPO), however, these data should be used with caution because CDWR equipment at the DEPO Soda Springs station does not receive regular and rigorous calibration, and data do not receive significant QA/QC before they are published online. The Scripps data is available upon request. The abbreviation 'DEPO-AWS' refers to the Scripps component of this station. In the future, Scripps plans to transfer their DEPO data to the Western Regional Climate Center (WRCC). Scripps does not perform QA/QC on the data on a routine schedule, although they will perform QA/QC on the data periodically. The time frame for any QA/QC may be significantly longer than SIEN is able to wait, possibly well after we hope to complete each annual report. We will address this problem by requesting data from Scripps on an annual basis. The SIEN Data Manager will import the data into an MS Access database, perform basic QA/QC to flag missing values and logical inconsistencies, and convert data collected at 1-minute intervals to daily and monthly averages. Processed data will be formatted so it can be archived in the Climate Database using "IMD SOP: Submitting Data for Upload" (Frakes 2012b).

3.5 Data Storage and Archiving

As described above, SIEN will utilize tabular, geospatial (PRISM), and derived data for reporting purposes. The sources and storage location of the data are included in Table 3. Geospatial data, including the grid-files, the image or database files that result from processing with Python, and any ArcMap or Excel projects, will be downloaded and archived locally on the SIEN shared network drive. A SQL Server database operated by the National I&M office will be the repository for tabular NWS-COOP, snow course, and DEPO-AWS data. The "IMD SOP: Climate Database Schema" describes the structure and table relationships within the IMD Climate Database (Frakes and Kingston 2011b, see Appendix D: Data Dictionary). Summary figures and tabular statistics that are directly reported through various agency websites will also be relied on for reports, although in some cases such as USGS hydrographs and drought summary maps, the underlying tabular data will not be stored by SIEN. Any derived figures or tables of statistics will be stored on the shared network drive.

Table 3. Data sources for annual status and periodic trend reports and where data will be stored.

Organization and/or Program	Data Used	Source	Data Storage – File types and storage location
OSU PRISM grid-based data	4-km grids of precipitation and min/max temperature	ftp://prism.oregonstate.edu/pub/prism/us/grids/	4-km grid files from OSU, image files output from Python scripts, and ArcMap projects all stored on the shared network drive
NWS-COOP	Min/Max temp, precipitation, snow depth	SIEN acquires summary statistics from the IMD Climate Database (IMD will acquire data from National Climatic Data Center)	Daily tabular data stored in the IMD climate database
California Dept of Water Resources Snowcourse Surveys	Monthly snow water equivalent at stations and watershed average	http://cdec.water.ca.gov/snow/current/snow/	Tabular data and monthly summaries stored by and served through the CDWR CDEC website and tabular data stored locally in an access database
USGS	Annual hydrograph and streamflow statistics	http://nwis.waterdata.usgs.gov/nwis/sw	Tabular data stored in the USGS NWIS database. Tabular data, hydrographs, and summaries served through the NWIS website. Tabular data not stored by SIEN
Scripps – DEPO weather station	Min/Max Temp and Precipitation	Request data from Scripps staff annually	Tabular data in 1-minute intervals converted to daily and monthly averages in MS Access database. Processed data uploaded and archived in IMD Climate Database.
The Drought Monitor (multi-agency) and the Standardized Precipitation Index (SPI)	Drought intensity	http://www.drought.unl.edu/dm/monitor.html http://lwf.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html	Summary figures from agency websites

3.6 Data Acquisition and Processing

There are multiple processes associated with formatting the data acquired from the sources in Table 3. Figure 9 outlines the actions required of SIEN and IMD staff to acquire and process data and produce reports. Data processing encompasses all activities required to make the data useful for immediate and long-term analysis, reporting, and distribution, as well as storage. Processing includes documenting data sources, acquisition methods, any structural, format, and content changes; quality assurance processes, and transformations, manipulations, or summarizations that prepare the data for use in the appropriate analysis applications, reporting systems, and data storage systems.

Grid-based data (PRISM) are acquired through their program websites. They are then processed with Python scripts using the toolsets developed by the IMD and detailed in “IMD SOP: Climate Grid Analysis Tools (CGAT)” (Sherrill and Frakes 2011). The CGAT scripts result in image files that are then incorporated into maps in ArcMap using the instructions in “SIEN SOP 4: Processing Geospatial Data for Climate Status Reports”.

In general, tabular data from NWS-COOP stations is acquired by the IMD from the National Climatic Data Center (NCDC) each year, typically in April, upon request by SIEN following the steps outlined in “IMD SOP: Requesting Essential Parameter Values” (Frakes 2012a). The “IMD SOP: Connecting to the I&M Climate Database” provides a detailed description of how to connect to the IMD Climate Database, the tools/queries that are available to generate climate summaries, and how to use them (Frakes and Kingston 2011a). As detailed in the SOP, the easiest method for SIEN to access the database is through the use of SQL Server Management Studio Express (SSME). Through SSME, users are able to connect to the database remotely and retrieve summary climate statistics or tabular daily data for individual stations using functions written by the IMD. Alternate options for accessing the database, also detailed in the IMD SOP, include python scripts, MS Access, R Code, and MS Excel. The summary values can then be manipulated in SigmaPlot or Excel to create figures for the report.

The California Department of Water Resources (CDWR) Snow Survey program posts summary statistics for the snow water equivalent for each course in the SIEN as well as calculated averages for each major watershed. SIEN will acquire, process, and archive the snow course SWE data in a locally stored Access database, but will primarily pull the summary statistics directly from the CDWR website for reporting purposes. Furthermore, drought and streamflow summary figures can be created from each program’s website. SIEN will use data and tools on the user-friendly USGS NWIS website to create the annual hydrographs and acquire summary statistics for annual status reports. The appropriate websites are listed in Table 3 as well as “SIEN SOP 5: Annual Status Reports”.

The Climate Database and associated SOPs are expected to change through time and SIEN project staff will watch for and adapt to changes that may require more or less time and expertise in a given year or reporting cycle. Tools, such as R code, that can be used to produce graphics for annual status reports may become available over time as networks, parks, or IMD staff have time to develop them. The recommended strategy is to focus on data acquisition, processing, analysis, and reporting for those climate metrics and indices that directly support stated monitoring objectives. Optional data sources, climate measures, and analyses can be incorporated as resources allow.

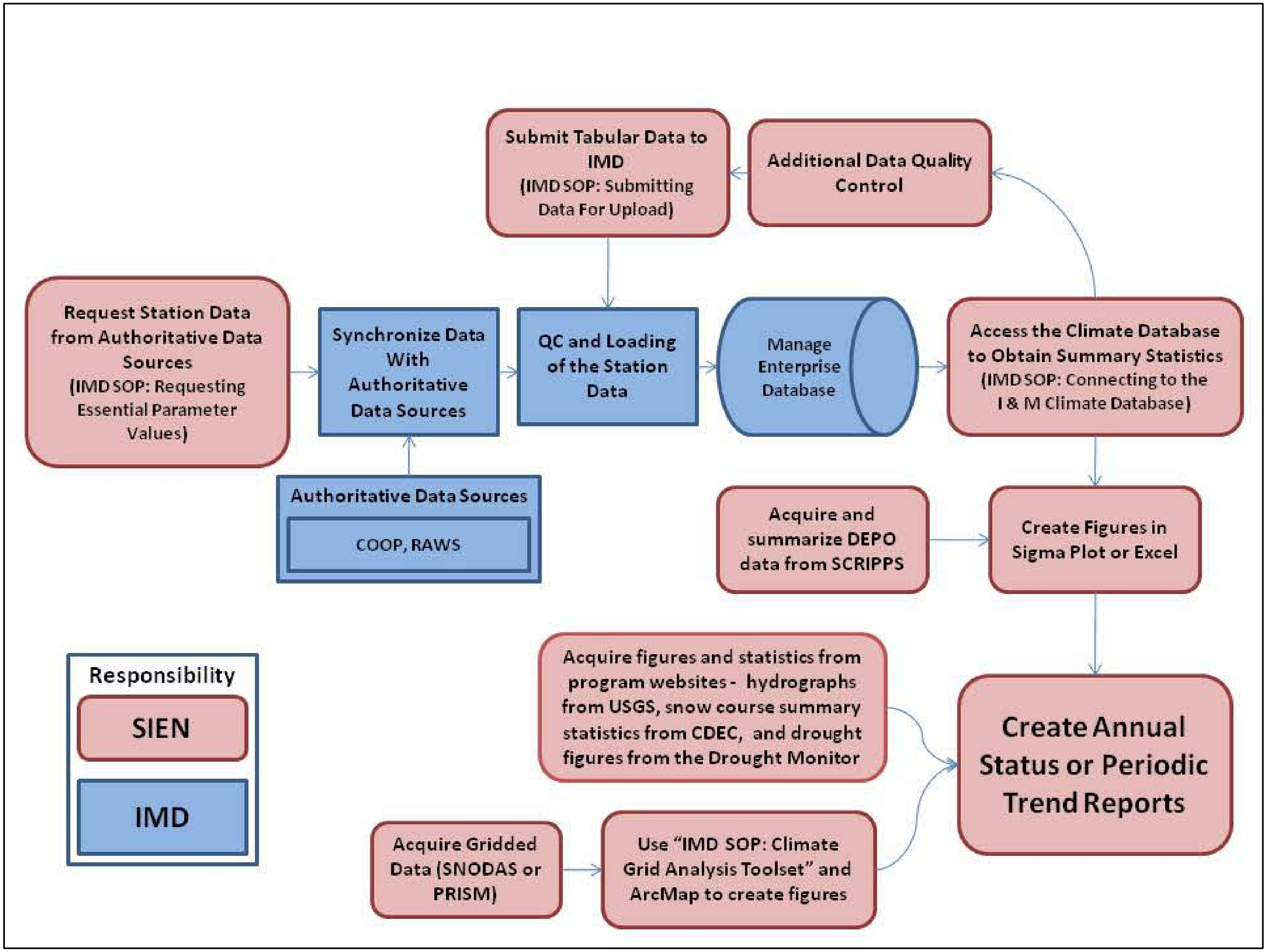


Figure 9. Workflow associated with data acquisition, processing, and reporting (modified from Frakes and Budde 2012).

3.7 Quality Control of Data

Quality control of data used for SIEN annual status and periodic trend reports is a multi-step and multi-agency process. The process varies for each type of data used by SIEN. The procedures undertaken to quality check data for annual status and long-term trend reports are identified in "SIEN SOP 2: Quality Control" and are summarized below. The SOP is based on the strategies for processing data to handle errors, inhomogeneities, and missing values outlined in Kittel et al. (2010) and Kittel (2009). A data record is said to be inhomogeneous when inconsistencies result from changes in station location, observation time, instrumentation, surrounding environment, observation practices, sensor drift, or others (National Climatic Data Center 2002). Statistical methods have been developed to adjust for such changes so that historic data conform to the most recent period.

For data from COOP stations, QA/QC begins when data is sent to the NCDC. The NCDC processes these surface climatological data through one system that utilizes standardized QA/QC algorithms and procedures (Del Greco et al. 2007). Next, the data is acquired by the IMD and loaded into the Climate Database. The IMD runs screening procedures for missing values, duplicate values, and inconsistencies on the data and reports errors back to the NCDC. The NCDC may choose to correct the errors, if possible, and the following year, the IMD will acquire the corrected data for each station for the period of record again. Each year, the IMD acquires the entire period of record and the previous dataset is overwritten. When analyzing data for annual status reports, the final QA/QC steps occur after SIEN acquires the data from the Climate Database. First, SIEN utilizes queries within the Climate Database to identify missing values and check for instances where the daily minimum temperature is above the daily maximum. Next, the daily values are graphed as a time series to identify and suspicious values. If errors are identified, SIEN acquires the original reporting form from NCDC and checks whether the values were entered incorrectly. If changes are made to the original data, the edited data set is uploaded to the Climate Database with a new version number.

When data are missing, SIEN will note the number of days missing from each month in the graphs that go in the annual reports. SIEN follow procedures used by WRCC which specify that months with greater than five days of missing temperature data are excluded from figures and from calculations of annual averages.

The data from the USGS Streamgages, PRISM, and CDWR snow surveys are typically released in a preliminary form with flags stating that the data has not gone through a thorough review. SIEN will not obtain data until they are posted in their reviewed and "final" versions. Each program has its own unique data review procedures that SIEN accepts as sufficient.

Quality Control procedures are more extensive when preparing data for long-term trend reports, are summarized below, and detailed in Kittel (2009).

3.7.1 Annual Status Report Quality Control

Data quality control requirements and activities for annual reports are limited in scope relative to the more comprehensive and time-consuming expert review required for data used in Climate Variability and Trend reports. These reports provide an accurate synopsis of the previous year's climate and, following the first trend report, will place the year's results in the context of historical averages and trends. The guiding strategy for status report quality control is:

- For certain datasets, accept the standard quality control and corrections made by the originating data provider.
- Ignore complex issues, such as outlier values or drifting values due to calibration issues, based on documented rationale. In reports, provide qualifiers and caveats to explain how these issues may affect the temporal and spatial analysis (Kittel 2009).
- Fill missing data or correct values by examining the original data wherever possible.

These tasks are outlined in “SIEN SOP 2: Quality Control” by variable, objective, and analysis. This strategy generally encompasses levels of quality control implemented in other network protocols and analyses, such as for Northern Colorado Plateau Network (Garman et al. 2004, Garman 2009) and Central Alaska Network (Sousanes 2004, Keen 2008).

Because of the limited amount of quality control that will be undertaken, data and analyses presented in status reports will include caveats that data and results are provisional and subject to being updated in the trends report.

3.7.2 Climate Variability and Trends Report Quality Control
Quality control for Variability and Trends analyses is far more rigorous in addressing complex data issues (see Kittel et al. 2010). Data checking and cleaning steps are outlined in “SOP 2: Quality Control” by analysis objective and variable and also in “SIEN SOP 6: Trend Analysis and Reporting.” Detailed steps to create useful climate datasets can be found in Kittel et al. (2010) and Kittel (2009).

Sophisticated quality control processes involve tailored treatment of station records that requires climatological expertise beyond that typically found within I&M networks. These processes include the use of regression analyses and models to account for missing data or weather station changes. Quality control procedures for trends reports will be conducted by or in close collaboration with a climatologist. For the first variability and trend report, the contracted climatologist will document their procedures so that the SIEN Physical Scientist may be able to replicate their methods for future reports. However, due to the complexity of the procedures, it is unlikely that the SIEN Physical Scientist would be to complete the report independently, but the amount of time and funding required for the contracted climatologist may be reduced.

3.8 Records Management, Data Maintenance, and Archiving
The objective of records management, data maintenance, and archiving procedures is to ensure that protocol data, analyses, products, and reports are well-documented, easily accessible, shared, and properly interpreted by a broad range of users in perpetuity (NPS 2008a). The SIEN Data Manager, in consultation with the Physical Scientist, will modify the established SIEN monitoring project file structure to accommodate the specific needs for storing and managing climate data. In most cases, “raw data”, data that has been subjected to preliminary checks for completeness and validation procedures but not rigorous error checking or QA, will initially be stored in the Climate Database. Versions of these datasets which have been error checked, corrected, and certified as complete by the SIEN Physical Scientist and Data Manager will subsequently be submitted for upload to the Climate Database with metadata documenting these changes.

The data directory for the climate monitoring protocol on the local SIEN network file server will store spatial and tabular data products that have been derived from the raw data. These include data summaries and analyses and data that have been manipulated, converted, or transformed from their original form. This repository may also include copies of certified data sets that are frequently accessed or requested.

3.9 Metadata Compilation

Metadata include the type of station, equipment specifications, the types of data reported, repairs reported, site characteristics, station moves, and data completeness statistics. SIEN maintains a geodatabase which contains two shapefiles comprising all known weather stations in and near the SIEN. The geodatabase is located on the shared network drive: \\Inpsekihqgis1\sekigis\sien\monitoring_projects\climate\data\spatial\MASTER_DATA (see Appendix B to view the current list of stations in the parks)

When new weather stations are added or changes are made to existing weather stations, the changes are made in the attribute table of the layer in ArcMap. Follow these steps to make changes:

1) Open an existing project in ArcMap that has the shapefile weatherstations_SIEN or weatherstations_near_SIEN.
2) Open the attribute table for the layer and start editing (Figure 10).

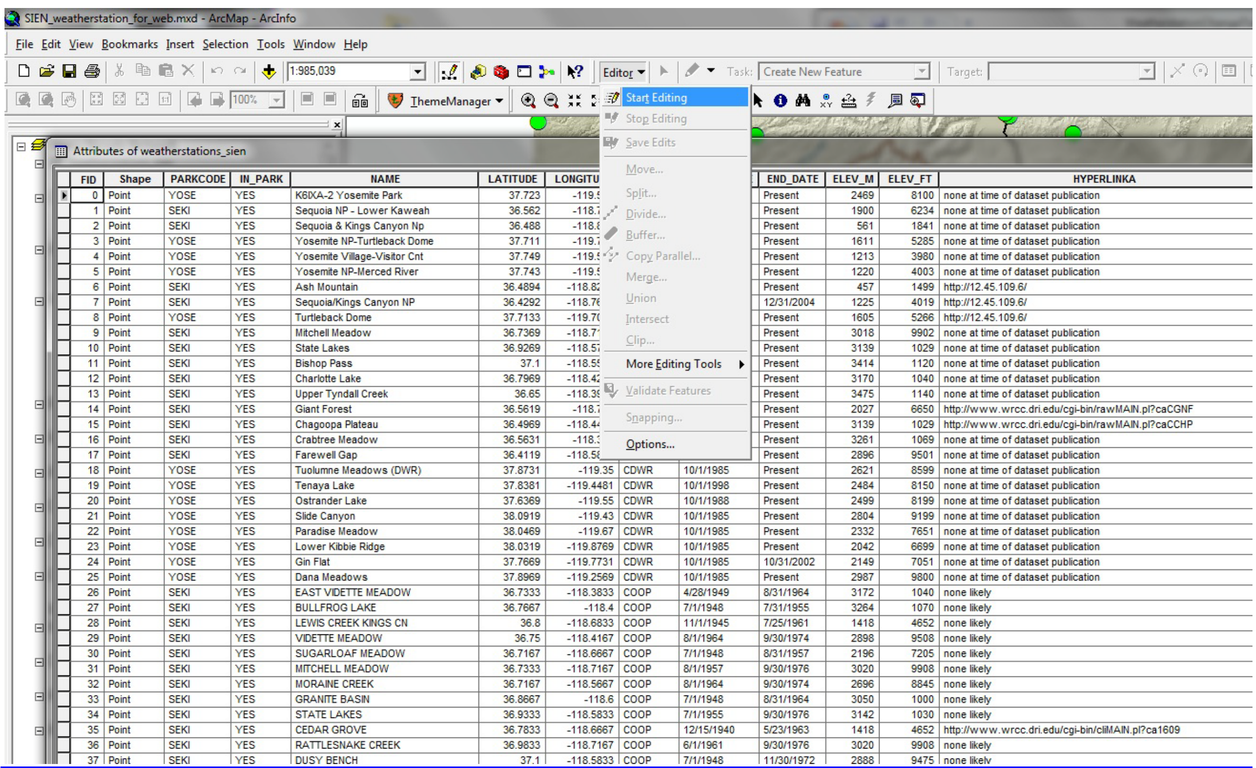


Figure 10. Weather station metadata is stored in the Weatherstations_Sien or Weatherstations_near_SIEN shapefiles and edited in ArcMap.

Additional tasks relating to metadata compilation include the following:

- Maintain detailed metadata information about each weather station: contacts, data quality issues, station reconfiguration and relocation, action items, data accessibility, and photographs. This information does not warrant creation of a specific relational database for storage.

- Review data completeness as part of the annual status reporting process. Swiftly deal with data collection and management needs.

- As new stations are installed in SIEN parks, work with the operating partner (especially if it is a small network or university) to ensure adequate attention is paid to (1) station maintenance, (2) data quality control, (3) data delivery, and (4) adequate metadata documentation. For new partners, help set up a relationship with the WRCC for data archiving.

- Most importantly, maintain a working relationship with station operators and support personnel. This is critical to achieveing all of the other support functions.

4.0 Data Analysis and Reporting

Relevant SOPs:

- SIEN SOP 5: Annual Status Reports
- SIEN SOP 6: Trend Analysis and Reporting
- SIEN SOP 7: Sharing Climate Data, Information and Reports
- SIEN SOP 8: Protocol Revisions

The goal of the protocol is to produce consistent and comparable monitoring information that can be used to support park management and decision making. The NPS aims to "improve park management through greater reliance on scientific knowledge" (NPS 2006), and the production of reports and other communication of scientific results transforms climate data into useful climate information. Our specific internal audiences include (1) park managers, (2) park resource professionals and other park staff, including interpretive staff, and (3) SIEN staff. Our external audiences include: (1) the academic community, (2) other government agencies, (3) nonprofit or non-governmental organizations, and (4) the general public.

The routine preparation of reports on a predictable and recurring basis of data summaries and basic interpretation can: (1) foster program support by establishing a client base, (2) motivate continued progress in program components, and (3) serve as the foundation for more comprehensive interpretive reports.

4.1 Reporting

Climate data will be analyzed and reported every year in status reports, as well as every five years in trend reports. Short two or three-page resource briefs, summaries of the longer reports, are published following the completion of each report. All reports will be published through the National Park Service's Natural Resources Technical Report (NRTR) series or Natural Resources Data Series (NRDS). The format for these reports and more detailed descriptions of the suggested content are included in SOP 9.

4.1.2 Reporting Products

Our monitoring goals and objectives and how we intend to report on these goals drive the data analysis portion of the protocol. For this reason, we begin the discussion of data analysis with an overview of our reporting products. This climate monitoring program will provide four reporting products to meet its goals:

1. *Climate Status Report*. This report provides a largely descriptive summary analysis of the climate of the past year or few years. Prepared and released on an annual cycle, the climate status report covers the previous calendar year (previous water year for snow and streamflow – Oct 1 to Sept 30). The purpose of the report is to support park science and management planning on an annual basis. Additionally, the status report puts the year's climate in context of longer-term variability patterns and trends identified in preceding trends and inter-annual variability report.

2. *Climate Variability and Trends Report*. This report will be produced on a 5-year cycle and will present rigorous analyses of inter-annual variability, long-term historical trends, and teleconnections with hemispheric climate patterns (e.g., the PDO). To describe the variability and trends of climate in a scientifically-defensible manner requires a substantial investment in quality control that is beyond the scope of what can be accomplished without climate science expertise (Kittel 2009). Therefore, these reports will require contracting or collaboration with a climatologist.

   The 5-year cycle for this report permits a high level of station data quality checks and correction since the previous report, and detailed analyses of long-term patterns in the annual, monthly, and daily climate record. These analyses will include comparison to an established baseline; Kittel et al. (2010) recommend that this be most recent 30-year "climate normal" period: (e.g., 1971–2000 for reporting years through 2010 or 1981-2010 for reporting years 2011 through 2020). The purpose of these trend reports is to provide park management, research, and public outreach with reliable, pertinent assessments of changes in park climates. Ideally, these reports will be turned into manuscripts and submitted to peer-reviewed journals. These assessments will be multi-faceted and will include:

   a. Evaluating a suite of ecologically-significant climate variables including temperature, precipitation, snowpack, drought, and surface hydrology
   b. Assessing a spectrum of climate dynamics in terms of daily (e.g., occurrence of extremes), inter-annual, and long-term behavior
   c. Testing connections to regional and hemispheric climatic processes

3. *SIEN Climate Webpage.* Each I&M network is responsible for developing and deploying website content specifically designed for that network's parks. The aim is to provide informative products: reports, maps, graphs, climate measures, and hyperlinks to various climate resources of interest. The SIEN Data Manager is responsible for the functionality of the website, and the SIEN Physical Scientist and cooperating climatologists are responsible for developing and maintaining current and useful content. Procedures and links to climate resources are provided in "SIEN SOP 7: Sharing Climate Data, Information, and Reports". SIEN is in the process of better understanding the amount and types of website content that are appropriate and viable given limited resources.

4. *Periodic Resource Briefs.* These publications will be brief summaries (less than 4 pages) of the climate of the previous year. The report will place the year's climate in an historical context by comparing the results to the 30-year average or other particularly dry, wet, hot or cold years. The aim is to provide an informative summary for other park staff including interpretive rangers, resource managers, or researchers which can guide readers to the annual status report for further information on any findings of interest.

4.1.3 Overview of analyses for reporting
Analyses associated with the protocol can be divided into two categories: those for the status reports are primarily descriptive in nature, and those for the trends reports are statistical tests and require advanced quality control checks prior to being executed. The former can be completed by

I&M staff, while the latter are expected to require input from a climatologist. Consequently, this protocol provides more detailed descriptions of analyses and reporting requirements for status reports (SIEN SOP 5) and a more general description of the content and quality control process for trend reports (SIEN SOP 6). Both SOPs will be refined over time as the reporting process is repeated.

Analyses and reporting for both the status and trends reports will occur at the scale of points (stations), parks, and the climate zone - an area of interest immediately surrounding the parks. SIEN intends to provide one annual report that covers all four park units. Report presentation and content are expected to change over time to improve efficiency and better meet the needs of park staff. It is expected to take multiple years to get the status reports to their desired level of detail and clarity of presentation. The data acquisition, processing, analysis, and reporting will become streamlined as the report-generating process is repeated, thereby allowing more time for network staff to focus on additional analyses and improving the format of the presentation.

4.1.4 Climate Variables Used in Reporting

The status and trends reports utilize variables that fall into four general categories:

- *Primary variables.* Key climate parameters directly measured (minimum and maximum temperature, precipitation, streamflow and SWE).
- *Integrative variables.* Variables expressing combined effects of primary variables (e.g., drought). These are either directly measured or derived from primary variables.
- Variables indicating the timing or length of a seasonal process (e.g., accumulated growing degree days, frosts). These are calculated from primary or integrative variables. Presently, SOPs for the acquisition, quality control, analysis, and reporting on these variables are not incorporated into the protocol. Such variables are more likely to appear in trend reports than status reports.
- *Secondary.* Other climate-related variables of interest, but currently not covered by the protocol (e.g., surface wind, solar radiation). Presently, SOPs for the acquisition, quality control, analysis, and reporting on these variables are not incorporated into the protocol. They may be added in the future as funding and staffing allow. Such variables are not likely to appear in status reports.

The metrics and analyses associated with primary and integrative variables have been described in detail in the monitoring objectives section of Chapter 1.

4.1.5 Status Reports

The suggested content for annual status reports is detailed in “SIEN SOP 5: Annual Status Reports.” In brief, the status reports provide a general summary of climate parameters that are relevant to ecological processes (temperature, precipitation, snowpack, streamflow, and drought) for a specific calendar year or water year (1 October–30 September) and place those results in a historical context with comparison to long-term averages. Results will be reported at multiple scales, from the station level to the climate region. Figure 11 is an example of a graph that would be included in an annual status report to highlight departures from historic averages at individual NWS-COOP stations.

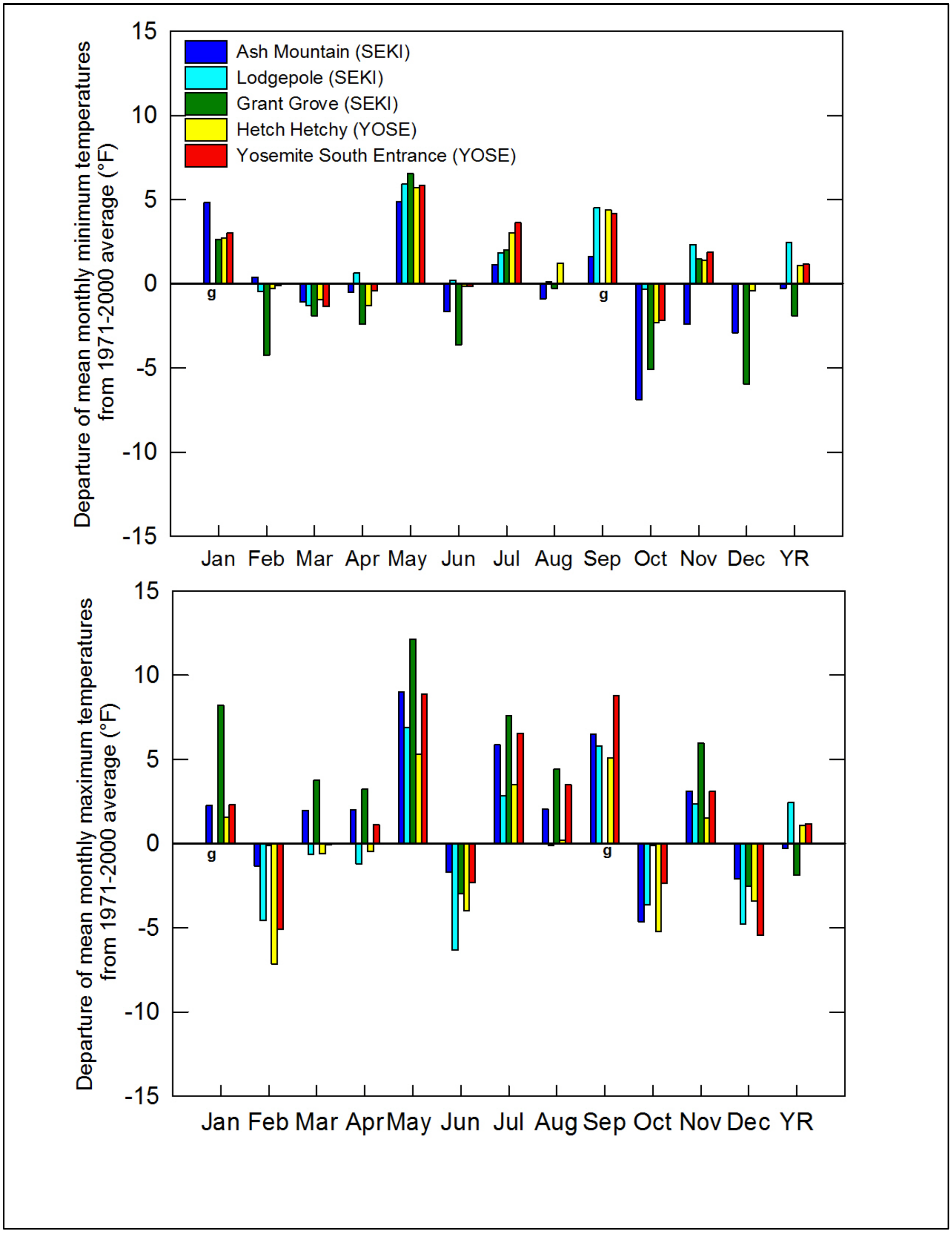


Figure 11. An example of the type of graph that will be included in annual status reports. This figure displays the departure of the monthly minimum and maximum temperatures from the historical average at selected NWS-COOP stations.

Data acquisition and formatting procedures will be performed according to the methods outlined in the preceding chapters. Data analysis for the status reports is largely descriptive and involves the creation of multiple graphics to aid with data comparisons and interpretation. Reports are expected to contain only essential data in the early years of reporting and to be expanded over time to include additional variables, a variety of presentation formats, and/or interpretation of annual results in the context of findings from periodic trend analyses.

For the most part, status reports need only include graphs that pertain to the reporting year. This approach saves effort in physical report preparation and keeps the narrative to the most important annual features and long-term dynamics, while still making adequate information available to users.

The narrative of the status reports integrates information across variables. The intent is not to be all inclusive of the data available, but to provide a succinct interpretation of the year's climate.

An important purpose of the report narrative is to aid with understanding the dynamics of other park resources and reveal areas for further analyses. By highlighting the departures from normal conditions, we may elucidate links between climate and other vital signs. For instance, our status reports may provide an impetus for park or I&M staff to investigate the effects of climate variables on lake or stream water quality. The annual report is not intended to provide a comprehensive analysis of climate, a physical understanding of why changes occur, or to determine long-term trends in climate.

4.1.6 Trend Reports

An outline of suggested procedures for creating inter-annual variability and trends reports (trends reports) are compiled in "SIEN SOP 6: Trend Analysis and Reporting". SOP 6 references Kittel (2009) for specific analytic approaches to trend analysis, but the actual methods used will ultimately be determined by the climate expert completing the report. The climatologist will take advantage of his/her knowledge of the latest techniques in climate analysis and use his/her expertise to determine the most appropriate approach for our data sets and to meet our objectives.

The goal of trend reports will be to determine the status and trends in key climate variables at the park or regional scale. Trends reports contain a comprehensive and rigorous analysis of only the highest-quality, long-term records from stations relevant to a park or climate zone. An important difference between the status and trend reports is the level of quality control procedures performed on the data. To avoid potentially reporting erroneous trends in climate variables, time and expertise are invested in sophisticated quality control measures and rigorous trend analyses. Historic and current data for selected stations will be assessed for quality and completeness and corrected where necessary by or in consultation with a climatologist. Kittel (2009) outlines the acquisition, quality control, and analyses that must be performed for trend reports.

Trend reports require a substantial investment in quality control and data management because the datasets used for these analyses require quality control measures and error processing specific to particular analyses. This necessitates the management of multiple copies of very similar data, with each copy having different quality control depending on the intended analysis. Additionally, prior to performing the trend analyses, multiple models are often used to fill data and/or perform serial correlation. As a result of the specialty knowledge needed to perform these techniques and the high level of quality control required prior to trend analyses, these reports will require expertise beyond that currently held within the SIEN I&M Program. The Climate Database will allow for the storage of the multiple versions of datasets along with details about any modifications made to the original data set.

Because of the significant time and effort involved in developing a robust historic record of climate for each region or park, it is expected that the first trends report will be the most labor intensive and require significant time and expertise from a climatologist. All subsequent status reports and trends reports may use data sets from and make reference to the climate record from the first trend report as the baseline and best record for the region.

4.1.7 SIEN Climate Website
SIEN will develop and maintain online resources for the protocol in conjunction with the network's website. These online resources will include all status and trends reports and information on how to access the data used in the creation of the reports, as well as current and archived protocols, standard operating procedures, resource briefs that summarize reports, and links to additional climate information. See http://science.nature.nps.gov/im/units/sien/monitoring/WeatherAndClimate/WeatherAndClimate.cfm

4.2 Protocol Updates
The methodology and procedures of the protocol will be initially evaluated after 1–2 years and then at every completion of a trend monitoring cycle (every 5 years). These reviews may include suggested modifications to data acquisition, analysis, and reporting based on either scientific considerations or budgetary constraints. An official programmatic review of the protocol approach to vital signs, including the objectives, designs, methods, and results as well as the ability of the network to sustain these protocols, will be conducted after approximately five years. The Physical Scientist and climate work group will review the narrative, SOPs, associated database, and other products. Changes are logged according to procedures outlined "SIEN SOP 8: Protocol Revisions".

The protocol narrative and each SOP contain a revision history log that is completed for each change, reasons for changes, and to assign a new version number to the revised SOP or narrative. Careful documentation of changes to the protocol and a library of previous protocol versions are essential for maintaining consistency in data acquisition and for appropriate treatment of the data during data summary and analysis.

5.0 Personnel and Operational Requirements

Relevant SOPs:
- SIEN SOP 4: Processing Geospatial Data
- SIEN SOP 5: Annual Status Reports
- SIEN SOP 8: Protocol Revisions

This chapter describes required personnel and funding resources, and roles and responsibilities for the protocol.

5.1 Project Management

The assigned project leader, the SIEN Physical Scientist, develops an annual climate monitoring work plan in coordination with the SIEN Program Manager and Data Manager as part of annual work planning. Annual planning will address the acquisition, management, analysis and reporting of climate data, and identify needs and projected costs not explicitly defined in this chapter.

5.2 Annual Workload

Accomplishing protocol objectives is a collaborative effort among SIEN staff, National Inventory & Monitoring Division (IMD) staff, and outside organizations and personnel. All SIEN-related work identified in this protocol will be completed by the SIEN Physical Scientist, Data Management Team, Program Manager, and Science Communication Specialist. Annual tasks include submitting a request to IMD for data acquisition, retrieval of summary statistics from the database, creation of figures for the annual status report, and finally, compilation and publication of annual status report (Table 4).

Table 4. Annual schedule of activities for the SIEN climate protocol.

Timing	Activity	Responsibility
October–January	• Determine annual budget • If applicable, prepare task agreements with cooperators	Program manager (assistance from the Physical Scientist)
March	• Submit data request to IMD (or email with confirmation that the list of stations for request is identical to the previous year)	Physical Scientist
Late March/ Early April	• Download PRISM geospatial data and process using IMD SOP: CGAT	Data manager or technician
Late April	• Pull summary statistics from the Enterprise Climate Database. If possible, utilize tools to create figures from the database.	Data Manager or technician
May	• Update project metadata records for previous year	Physical Scientist
	• Utilize summary statistics to create figures	Physical Scientist
	• Prepare annual status report and resource brief	Physical Scientist Science Communication Specialist
June	• Publish annual reports	Physical Scientist
	• Upload reports to SIEN website	Data Manager or Science Communication Specialist
September	• Close out year-end budget and prepare administrative reports	Program manager

5.3 Partnerships

SIEN is partnering with IMD in Fort Collins, Colorado. IMD is a branch of the Natural Resource Stewardship and Science Directorate (NRSS), which provides scientific and technical support to national parks for the management of natural resources. IMD has agreed to provide basic services to the I&M Networks regarding climate data. These services include the operation of a SQL Server database to house climate data, the acquisition of the data from nationwide programs, and the development of several tools for querying data and calculating summary statistics. Additionally, IMD and several I&M networks are working with Mike Tercek, an outside contractor, to develop tools which would easily analyze data from the database and output figures for annual reports. IMD and networks will continue to develop tools using Python or R scripts to easily output summary graphs and tables. As these tools are developed, all networks will benefit; the creation of annual status reports will become more automated and less time consuming. These procedures will be incorporated into “SIEN SOP 5: Annual Status Reports” as they are developed.

The success of this protocol is dependent on many other agencies and NPS staff. SIEN staff will work closely with NWS and NPS staff to ensure that data at NWS-COOP stations are consistently being collected and reported.

In addition, SIEN currently has a task agreement with researchers Roger Bales and Bob Rice at the Sierra Nevada Research Institute at the UC Merced to develop methods to better understand changes in snowcover across park landscapes using both ground-based and remotely-sensed data. Bales and Rice have collaborated with other universities, the California Department of Water Resources and NASA scientists to develop methods that combine these two different types of datasets to estimate snow-covered area, snow water equivalent, and snow depletion rates at watershed scales. The purpose of this task agreement is to apply these methods to estimating snow parameters for the six major SIEN watersheds for available years of record, 2000-2009, analyze resulting estimates temporally and spatially, summarize methods and results in a final NPS technical report, and recommend tools and approaches for a long-term snowcover monitoring plan. The products of the task agreement are currently being finalized. Pending the outcome of this project, SIEN may explore the potential for continuing this partnership, the feasibility of conducting these analyses on a periodic basis, and potential funding opportunities.

5.4 Qualifications

Each position identified in Table 5 requires minimum background knowledge, skills, and abilities. The Physical Scientist must be familiar with methods to provide climate summaries that are useful to national park managers and other park staff. This includes sample design, data requirements, data processing, analysis tools and approaches, and professional written and verbal communication of scientific results and ideas. The project leader and/or cooperators with subject-matter expertise will apply data preparation and computer analysis techniques that meet current standards of the climate science community. The administrative lead must be familiar with the standards and requirements for the NPS I&M Program and use of cooperative and interagency agreements. The Data Manager must work with the project lead to understand and provide for climate data input and processing requirements, related database applications, tools and procedures, basic website design and/or maintenance, and metadata production tools. The production of trend and inter-annual variability reports requires climate science expertise (Kittel et al. 2010). The network will need to obtain outside expertise to create the trend reports if there is not a climate expert either on staff or available through a collaborator.

Table 5. Roles and responsibilities associated with implementation of the SIEN Climate Protocol.

Protocol role	Responsibilities	Time required	Network position
Administrative Lead	• Provides program oversight and administration • Tracks and reports project budget, requirements, objectives, and progress toward meeting objectives • Serves as NPS key official, agreement technical representative, contracting officers technical representative on agreements or contracts for climate monitoring work • Reviews reports and other products for completeness and compliance with I&M Program guidance and specifications • Liaison to WASO programs, offices, and other I&M networks	1 pp	Program manager
Physical Scientist / Project Lead	• Plans and coordinates project operations, namely the acquisition, analysis and reporting of climate data • Performs or oversees maintenance and archiving of project records • Interprets data and prepares reports and other products (posters, website content, etc.) • Identifies need for advanced technical assistance in analysis and interpretation of data and prepares scope of work for agreements and/or contracts • Coordinates and ratifies changes to the protocol	2 pp	Physical Scientist
Data Management Team	• Advises on data and information management activities • Acquires, organizes, manages, and process spatial data in preparation for analysis, distributions, and archiving • Post metadata, reports, and other products to NPS delivery systems and network website • Provides data management training as needed • Prepares of oversees development of metadata for spatial and tabular data sources and products	3 pp	Data manager and data management assistant
Science Communication	• Assist with resource briefs • Update climate website text	0.5 pp	Science Communication Specialist
Climatologist	• Prepare periodic trends reports	6 pp for initial; 3 pp thereafter	Non-network position

5.5 Training

If cooperative agreements are used, personnel in this position must have completed the NPS Agreement Technical Representative training. Cooperative agreements would likely be managed by the Program Manager.

Additionally, management and analysis of the climate data associated with this protocol relies on numerous software programs including SQL Server Management Studio Express (SSME), ArcMap, MS Excel, SigmaPlot, and PythonWin. A member of the SIEN data management team will assist with the training of the Physical Scientist in the use of SSME and ArcMap and IMD staff will assist with the training of SIEN staff in the use of PythonWin. Further, instructions for the operation of PythonWin are included in "SIEN SOP 4: Processing Geospatial Data for Status Reports". We assume that all SIEN staff will have a basic understanding of MS Excel and SigmaPlot; the graphing functions in these programs are relatively simple and should be manageable with the assistance of the "help" functions in the programs. Further, in many cases, previously developed templates will be utilized so that only the current year's summary statistics need to be input for the graphs to be updated. If a SIEN staff is not familiar with SigmaPlot or the graphing functions in MS Excel, they will be encouraged to undertake the online trainings available through the SigmaPlot website or the Skillsoft e-learning courseware.

5.6 Facility, Software and Equipment Needs

This protocol requires minimal specialized equipment or facilities. The Data Manager will ensure that the necessary software are installed and updated. The software needed for implementation of this protocol includes SQL Server Management Studio Express, ESRI ArcMap, and PythonWin. Additional software that may be used include SigmaPlot, Excel, and R.

5.7 Budget

This protocol harvests, analyzes, and reports on data from established weather monitoring stations and programs that are operated and maintained by other organizations. As a result, the primary annual expenses for implementing this protocol are related to SIEN staff salaries and general support (Table 6). SIEN's annual operating budget provides computer hardware, standard software, and other office needs and thus they are not listed here. Estimates of the amount of staff time dedicated towards the overall implementation and specifically toward data management are listed in Table 6. The protocol focuses on data harvesting, management, QA/QC, analysis, reporting, and archiving, and the data manager is expected to participate in Climate Database and climate data management discussions at the annual, National I&M Program-sponsored Data Manager's meeting. Thus, almost three-fourths of the annual budget is used towards data management.

SIEN anticipates producing a climate trend report on a 5-yr schedule. The initial trend report requires QA/QC of the entire historical data set for multiple stations, while subsequent trend reports require QA/QC of only the previous 5 years of data; hence the first trend report will be more costly to produce than subsequent reports. To produce an individual network trend report for the first time, we anticipate a cost of up to $30,000; however, this cost may be reduced by pursuing a multi-network contract or agreement. As the need for a historic climate analysis was identified as a high priority during preparation of the SEKI Natural Resources Condition Assessment, SEKI and SIEN may also combine funds to support development of this initial trend report.

Table 6. Annual budget for climate monitoring protocol. Costs are primarily related to salary, and are largely devoted to data management activities.

Expense	Time allotted for duties	Cost ($)	Data Management	
			% of time	Cost ($)
Personnel				
Program Manager (GS-13)	1.0 pay period	4,666	30	1,400
Data Manager (GS-11)	1.5 pay periods	5,325	100	5,325
Physical Scientist (GS-11)	2.0 pay periods	6,400	50	3,200
Data Management Technician (GS-7)	1.5 pay periods	3,300	100	3,300
Science Communication Specialist (GS-11)	0.5 pay period	1,600	30	481
Travel				
Climate Database Meetings at Annual National Data Manager's meeting		1,750		1,750
Periodic cost (every 5 years)				
Climatologist (GS-12 level)- estimates reflect half the time commitment for subsequent trend reports	6 pay periods; then 3 pay periods	30,000 15,000	50% 50%	15,000 7,500
Total Annual Cost for years with no trend report		21,291		15,456
Total Annual Cost during 1st Trend Report Year		51,191		30,456
Total Annual Cost during subsequent trend report years		36,291		22,956

Literature Cited

Bales, R. C. and R. Rice. In Prep. A snowcover assessment for the Sierra Nevada Network of National Parks. Natural Resource Technical Report. National Park Service, Fort Collins, Colorado.

Barnett, T. P., D. W. Pierce, H. G. Hidalgo, C. Bonfils, B. D. Santer, T. Das, G. Bala, A. W. Wood, T. Nozawa, A. A. Mirin, D. R. Cayan, and M. D. Dettinger. 2008. Human-induced changes in the hydrology of the western United States. Science 319:1080–1083.

CIRMOUNT Committee. 2006. Mapping new terrain: climate change and America's west. Report of the consortium for integrated climate research in western mountains. (CIRMOUNT), Misc. Pub., PSW-MISC-77. Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture. Albany, California.

Cook, R. R. and P. Lineback. 2008. Sierra Nevada Network data management plan. Natural Resource Report NPS/NRPC/NRR—2008/070. National Park Service, Fort Collins, Colorado.

Daly, C. 2006. Guidelines for assessing the suitability of spatial climate data sets. International Journal of Climatology 26: 707-721.

Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris. 2002. A knowledge-based approach to the statistical mapping of climate. Climate Research 22: 99–113.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. A. Pasteris. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. International Journal of Climatology 28: 2031-2064.

Davey, C. A., K. T. Redmond, and D. B. Simeral. 2007. Weather and climate inventory, National Park Service, Sierra Nevada Network. Natural Resource Technical Report NPS/SIEN/NRTR—2007/042. National Park Service, Fort Collins, Colorado. Available from: https://nrinfo.nps.gov/Reference.mvc/Profile?Code=649248

Del Greco, S. A., N. Lott, R. Ray, D. Dellinger, P. Jones, and F. Smith. 2007. Surface data processing and integration at NOAA's National Climatic Data Center, 87th AMS Annual Meeting, 13-18 January 2007, San Antonio, Texas, combined preprints [CD-ROM], American Meteorological Society, Boston, MA.

Dettinger, M. D., D. R. Cayan, M. K. Meyer, and A. E. Jeton. 2004. Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900-2099. Climatic Change 62:283-317.

Edwards, L. M., and K. T. Redmond. 2011. Climate assessment for Sierra Nevada Network Parks. NPS/SIEN/NR-2011/482. National Park Service, Fort Collins, Colorado. Available from: http://science.nature.nps.gov/im/units/sien/monitoring/Reports/SIENClimateAssessment_EdwardsRedmond_Final_20111230.pdf (accessed 26 March 2012).

Fancy, S. G., J. E. Gross, and S. L. Carter. 2008. Monitoring the condition of natural resources in U.S. national parks. Environmental Monitoring and Assessment 151:161–174.

Frakes, B. 2012a. Requesting essential parameter values from COOP, SNOTEL, streamgage, and snow course data for acquisition, compilation, basic quality control, and distribution by the NPS Natural Resource Program Center. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166927 (accessed 30 March 2012).

Frakes, B. 2012b. Submitting data for upload to the I&M Enterprise climate database. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166928 (accessed 30 March 2012).

Frakes, B., and P. Budde. 2012. The I & M Enterprise climate database project description. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/DownloadDigitalFile?code=443748&file=ClimateDatabaseProjectDescription.pdf (accessed 30 March 2012)

Frakes, B., and S. Kingston. 2011a. Connecting to the I&M Enterprise climate database for the purpose of data retrieval, summary and/or analysis. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2167699 (accessed 30 March 2012).

Frakes, B., and S. Kingston. 2011b. Climate database schema, table and field definitions. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2167699 (accessed 30 March 2012)

Garman, S. L. 2009. Climate monitoring in the Northern Colorado Plateau Network: Annual report, 2007. Natural Resource Technical Report NPS/NCPN/NRTR—2009/216. National Park Service, Fort Collins, Colorado.

Garman, S., M. Beer, M. A. Powell, and R. DenBleyker. 2004. Climate monitoring protocol for the park units in the Northern Colorado Plateau Network. Version 1.00, 15 December 2004. U.S. Department of the Interior, National Park Service, Inventory and Monitoring Program, Northern Colorado Plateau Network, Moab, Utah.

Graumlich, L. J., 1993. A 1000-year record of temperature and precipitation in the Sierra Nevada. Quaternary Research 39:249-255.

Gray, S. T. 2008. Framework for linking climate, resource inventories, and ecosystem monitoring. Natural Resource Technical Report NPS/SIEN/NRTR—2008/110. National Park Service, Fort Collins, Colorado.

Hidalgo, H. G., T. Das, M. D. Dettinger, D. R. Cayan, D. W. Pierce, T. P. Barnett, G. Bala, A. Mirin, A. W. Wood, C. Bonfils, B. D. Santer, and T. Nozawa. 2009. Detection and attribution of streamflow timing changes to climate change in the western United States. Journal of Climate 22:3838–3855.

Keen, R. A. 2008. Climate data analysis of existing weather stations in around the Central Alaska Network (CAKN), including Denali National Park and Preserve, Wrangell-St. Elias National Park and Preserve, and Yukon-Charley Rivers National Preserve. Unpublished report. National Park Service, Central Alaska Network, Fairbanks, Alaska. Available from: http://science.nature.nps.gov/im/units/cakn/Documents/2008reports/CAKN_Climate_Data_%20Analysis_%20Keen_2008.pdf (accessed 9 April 2012).

Kittel, T. 2008. A framework for understanding climate regions. Unpublished report. National Park Service, Fort Collins, Colorado.

Kittel, T. 2009. The development and analysis of climate datasets for National Park science and management: A guide to methods for making climate records useful and tools to explore critical questions. Final draft report (Dec. 10, 2009) prepared for the National Park Service Inventory and Monitoring Program. University of Colorado, Institute of Arctic and Alpine Research, Boulder, Colorado.

Kittel, T., S. Ostermann-Kelm, B. Frakes, M. Tercek, S. Gray, and 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, Montana, 7–8 April 2009. Unpublished final report (18 February 2010). National Park Service, Greater Yellowstone Network, Bozeman, Montana.

Knowles, N., M. D. Dettinger, and D. R. Cayan. 2006. Trends in snowfall versus rainfall in the western United States. Journal of Climate 19:4545–4559.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis. 1997. A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. Bull. Amer. Meteor. Soc.78:1069-1079.

McKee, T. B., N. J. Doesken, and J. Kleist. 1993. The relationship of drought frequency and duration to time scales. Preprints. Eighth Conference on Applied Climatology. Anaheim, California. 1993:179–184.

Mutch, L. S., M. Goldin Rose, A. M. Heard, R. R. Cook, and G. L. Entsminger. 2008. Sierra Nevada Network vital signs monitoring plan. Natural Resource Report NPS/SIEN/NRR—2008/072. National Park Service, Fort Collins, Colorado.

National Climatic Data Center. 2002. United States Climate Normals, 1971-2002: Inhomogeneity adjustment methodology. Online. Available from: http://www.sercc.com/normnws0320.pdf (accessed 13 April 2012).

NOAA (National Oceanic and Atmospheric Administration). 2011. Location of U.S. Climate Divisions. Online. Available from: http://www.esrl.noaa.gov/psd/data/usclimdivs/data/map.html (accessed 30 March 2012).

National Park Service (NPS). 2006. Management policies 2006. U.S. Department of the Interior, National Park Service, Washington D.C.

National Park Service (NPS). 2008a. Data management guidelines for inventory and monitoring networks. Natural Resource Report NPS/NRPC/NRR—2008/035. National Park Service, Fort Collins, Colorado.

National Park Service (NPS). 2008b. Inventory and Monitoring program description. Online. Available from: http://science.nature.nps.gov/im/ (accessed 30 March 2012).

Pierce, D. W., T. P. Barnett, H. G. Hidalgo, T. Das, C. Bonfils, B. D. Santer, G. Bala, M. D. Dettinger, D. R. Cayan, A. Mirin, A. W. Wood, and T. Nozawa. 2008. Attribution of declining western U.S. snowpack to human effects. Journal of Climate 21:6425–6444.

Rocky Mountain Climate Working Group. 2010. Rocky Mountain climate protocol: Climate monitoring in the Greater Yellowstone and Rocky Mountain Inventory and Monitoring Networks, Version 1.0. Natural Resource Report NPS/IMRO/NRR—2010/222. National Park Service, Fort Collins, Colorado.

Sherrill, K. R. and B. Frakes. 2011. Climate grid analysis toolset – Tools for assessing regional climatological trends: Standard operating procedure (Version 1.2). Natural Resource Report. NPS/NRSS/IMD/NRR—2011/411. National Park Service, Fort Collins, Colorado.

Sousanes, P. J. 2004. 2003 Long term ecological monitoring climate data summary Denali National Park and Preserve. Unpublished report. Denali National Park and Preserve, Denali Park, Alaska.

Stephenson, N. L. 1988. Climatic control of vegetation distribution: The role of water balance with examples from North America and Sequoia National Park, California. Dissertation. Cornell University, Ithaca, NY.

Stephenson, N. L. 1998. Actual evapotranspiration and deficit: Biologically meaningful correlates of vegetation distribution across spatial scales. Journal of Biogeography 25:855-870.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger. 2004. Changes in snowmelt runoff timing in western North America under a "business as usual" climate change scenario. Climatic Change 62: 217–232.

Stine, S. 1996. Climate, 1650-1850. Pages 25-30 *in* Sierra Nevada Ecosystem Project, Final report to Congress, Volume II: Assessments and scientific basis for management options. Wildland Resources Center Report No. 37. Centers for Water and Wildland Resources, University of California, Davis, California.

Swetnam, T. W. 1993. Fire history and climate change in Giant Sequoia groves. Science 262:885-889.

Urban, D.L., C. Miller, P. N. Halpin, and N. L. Stephenson. 2000. Forest gradient response in Sierran landscapes: the physical template. Landscape Ecology 15:603–620.

Appendix A: Data Sources

This appendix has been modified from Appendix A in the Rocky Mountain Climate Protocol (Rocky Mountain Climate Working Group 2010)

Acronyms

COOP	Cooperative Observer Network
CDWR	California Department of Water Resources
CDEC	California Data Exchange Center
ENSO	El Niño-Southern Oscillation
MJO	Madden-Julian Oscillation
NAO	North Atlantic Oscillation
NCDC	National Climatic Data Center
NWS	National Weather Service
NRCS	Natural Resources Conservation Service
NSIP	National Streamflow Information Program
PDO	Pacific Decadal Oscillation
PDSI	Palmer Drought Severity Index
PNA	Pacific/North American pattern
PRISM	Parameter-elevation Regressions on Independent Slopes Model
Scripps	Scripps Institution of Oceanography
SNODAS	Snow Data Assimilation System
SNOTEL	Snowpack Telemetry
SOI	Southern Oscillation Index
SWE	snow water equivalent
TNH	Tropical-Northern Hemisphere

Introduction

The approach of the Sierra Nevada Network Weather & Climate Protocol is to acquire data from existing national climate monitoring programs that provide consistent, long-term, and high-quality climate records for our regions and provide the critical steps of summarizing, reporting, and interpreting status and trends in climate at the park scale. Here, we describe the purpose, type of data, and strengths and limitations of each of these national climate monitoring programs including the National Weather Service Cooperative Observer Network, the Parameter-elevation Regressions on Independent Slopes Model, CA Snow Survey program, Snow Data Assimilation System (SNODAS), USGS Streamflow, drought indicies, and atmospheric and oceanic indices, and the station operated jointly by the Scripps Institution of Oceanography (Scripps) and California Department of Water Resources (CDWR) at DEPO.

Cooperative Observer Program

The National Weather Service (NWS) daily Cooperative Observer Network (COOP) Stations have been a foundation of the U.S. climate program for decades and has long served as the main climate observation network in the United States. NWS-COOP stations are established, supervised, and inspected by NWS personnel. Manual stations require recording of climate observations on a daily basis. Readings are usually made by volunteers using equipment supplied, installed, and maintained by the federal government. The observer in effect acts as a host for the data-gathering activities and supplies the labor; this is truly a “cooperative” effort. The U.S. Historical Climatology Network (http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn.html) is a subset of the cooperative network but contains longer (80 years or more) and more complete records. The mission of the COOP, created in 1890, is: (1) to provide observational meteorological data required to define U.S. climate and help measure long-term climate changes; and (2) to provide observational meteorological data in near real-time to support forecasting and warning mechanisms and other public service programs of the NWS (National Weather Service 2009).

The COOP provides national coverage with more than 11,000 volunteers, which in some cases include National Park Service employees. Observers record temperature and precipitation data daily. Data are reported daily or monthly (depending on the station) to the National Climatic Data Center (NCDC) or an NWS office. Typical observation days are morning to morning, evening to evening, or midnight to midnight. By convention, observations are ascribed to the date the instrument was reset at the end of the observational period. For this reason, midnight observations represent the end of a day. Observations include: daily maximum and minimum temperature, daily observation-time temperature, daily liquid precipitation, snowfall and snow depth, and pan evaporation (at some stations). Additional measurements may include river stage and special phenomena, such as hail and damaging winds.

Although some COOP stations have electronic instrumentation, they lack automated transmission capability. Daily observations are obtained by personnel directly reading the instruments (e.g., min-max thermometers and rain gauges) or by reading digital displays connected to electronic sensors. Procedures for reading instrumentation as well as for the maintenance and calibration of equipment performed by the cooperative observer are described in “Observing handbook no.2: Cooperative station observations” (NWS 1989). Cooperative observers also report damaged or defective equipment and instruments to the NWS representative, who informs the observer about arrangement of repair or replacement.

Data from COOP stations are transmitted to designated regional offices of the NWS immediately after the 24-hour temperature and precipitation observations are recorded. Observations are transmitted by phone to a recording device where the observer is prompted by a machine for specific information. The NCDC receives all monthly reports from the observer or the NWS, transcribes the data from paper copies to electronic databases where necessary, archives the data, and distributes error-checked daily values to cooperating agencies (e.g., regional climate centers).

Strengths of the COOP data include: long-term records at most sites (i.e., decades to a century); a widespread national coverage with thousands of stations; excellent data quality when stations are well maintained; it is cost-effective, and manually-taken measurements can eliminate problems with equipment and remote communications. Limitations of the COOP include: daily rather than hourly observation, uneven exposures, many stations are not well-maintained, a dependence on schedules of volunteer observers, slow entry of data from many stations into national archives, data are subject to observational methodology, which is not always documented, and manual measurements (there is an increase in variation among stations caused by observers, as opposed to automated recorders).

Snow Survey

The Snow Survey program in California is primarily managed by the California Department of Water Resources (CDWR); the data are posted on the California Data Exchange Center (CDEC) http://cdec.water.ca.gov/snow/current/snow/. Some California snow survey sites are managed by the Natural Resouces Conservation Service (NRCS) although none of them are in the SIEN. Sampling sites in this program may be snow courses (about 1,000 ft long) which are manual sites, where trained observers measure only snow depth and snow water content or snowpillows which are automated sites. At snow courses, measurements are taken one to two times per month during the winter and spring. Data records for these sites often extend back to the 1920s or 1930s, and the data are generally of high quality. The data collection process can be viewed at http://cdec.water.ca.gov/snow/hwy50/. The purpose of the snow survey network is to collect snowpack and related climate data to assist in forecasting water supply.

U.S. Geological Survey (USGS) Streamgages

The USGS has long operated a collaborative national network of stream gaging stations to meet federal, state, and local user needs for information on streamflows. The network is currently funded in partnership with over 800 agencies. Approximately 7,500 streamgages are used to provide long-term, accurate, and unbiased information on streamflow. Across the nation, there are approximately 25,000 sites with gages that report a summary of daily values for streamflow and over 8,500 stations report real-time, time-series data on flow (recorded at fixed intervals) from automated equipment. These data represent the most current hydrologic conditions. Measurements are recorded at 5–60 minute intervals and transmitted to the National Water Information System (NWIS) database every 1–4 hours. Real-time data are available online for a period of 31 days. On an annual basis, USGS publishes daily streamflow data in a series of water-data reports. Time-series data describe: streamflow (discharge), stream levels, reservoir and lake levels, precipitation, and surface-water quality.

Strengths of the USGS Gaging Stations include: long-term records at many sites, a widespread national coverage with thousands of stations, accurate and unbiased information, and real-time data that can be are easily obtained online.

Parameter-elevation Regressions on Independent Slopes Model (PRISM)

Parameter-elevation Regressions on Independent Slopes Model (PRISM; Oregon State University 2007) is a climate mapping system that uses point measurements of precipitation, temperature, and other climatic factors to produce continuous, digital grid estimates of monthly, yearly, and event-based climatic parameters. Data extend back to 1895 and are offered at numerous spatial scales with the 800-meter grid being both free and of reasonably high resolution. The greatest utility of PRISM is that it presents the spatial distribution of temperature and precipitation, which single point observations are unable to provide. The model was originally developed to provide climate information at scales matching available land-cover maps to assist in ecological modeling and to address the extreme spatial and elevation gradients exhibited by the climate of the western United States. (Daly et al. 1994, 2002, Gibson et al. 2002, Doggett et al. 2004). The PRISM technique accounts for the scale-dependent effects of topography on mean values of climate elements. Elevation provides the first-order constraint for the mapped climate fields, with slope and orientation (aspect) providing second-order constraints. The model has been enhanced gradually to address inversions, coast/land gradients, and climate patterns in small-scale trapping basins.

Monthly climate fields are generated by PRISM to account for seasonal variations in elevation gradients in climate elements. These monthly climate fields then can be combined into seasonal and annual climate records. Since PRISM maps are grid maps, they do not replicate point values but rather, for a given grid cell, represent the grid-cell average of the climate variable in question at the average elevation for that cell. The model relies on observed surface and upper-air measurements to estimate spatial climate fields. Data include: precipitation, maximum temperatures, minimum temperatures, dew point temperatures, and percent of normal precipitation.

PRISM incorporates point data, a digital elevation model, and expert knowledge of climate extremes. Although PRISM data sets were developed through projects funded, in part, by the federal government, there is not much funding for the maintenance and expansion of the data sets. Data may be available for a limited time only.

Atmospheric Indices

Long-term and large-scale atmospheric and ocean variations play a key role in understanding and predicting intra-seasonal and inter-annual variations in climate. Below we describe the indices we may use to determine if there are correlations among them and long-term climate patterns in the SIEN parks. In all cases, the indices are calculated by another entity. We will acquire the monthly or annual data and correlate these global scale indices to other regional or local climate.

Southern Oscillation Index (SOI)

The SOI is calculated by the NWS Climate Prediction Center from the monthly or seasonal fluctuations in the air pressure difference between Tahiti and Darwin, Australia. The index is most commonly computed on a monthly basis and ranges from about -3 to +3. The anomaly is based on the period of 1951–1980 and calculated as the difference between a standardized Tahiti and Darwin pressure (for more details see: http://www.cpc.ncep.noaa.gov/data/indices).

Prolonged periods of negative SOI values coincide with abnormally warm ocean waters across the eastern tropical Pacific typical of ENSO episodes. Prolonged periods of positive SOI values coincide with abnormally cold ocean waters across the eastern tropical Pacific typical of La Niña episodes. ENSO episodes are associated with wetter than normal conditions during June–August in the intermountain regions of the United States and most of regions experience abnormally warm conditions during December–February.

Madden-Julian Oscillation (MJO)

The Madden-Julian Oscillation (MJO) is the dominant component of the intraseasonal (30–90 days) variability in the tropical atmosphere. It consists of large-scale coupled patterns in atmospheric circulation and deep convection, with coherent signals in many other variables, all propagating eastward slowly (~5 m s^{-1}) through the portion of the Indian and Pacific oceans where the sea surface is warm. It constantly interacts with the underlying ocean and influences many weather and climate systems (Zhang 2005).

Pacific Decadal Oscillation (PDO)

The PDO is a pattern of climate variability based on sea surface temperatures in the North Pacific Ocean (Zhang et al. 1997, Mantua et al. 1997). The monthly mean global average SST anomalies are removed to separate this pattern of variability from any "global warming" signal that may be present in the data. The PDO shows 20–30 year phases of cool or warm sea temperatures. Positive PDO values, or warmer sea temperatures, are usually associated with wetter conditions in the southwestern United States, while negative PDO values are suggestive of persistent drought in the southwest (Webb et al. 2000).

Pacific/North American Pattern (PNA)

The Pacific/North American teleconnection pattern (PNA) describes the atmospheric circulation in the Pacific and North America (Climate Prediction Center 2005). The PNA pattern is associated with strong fluctuations in the strength and location of the East Asian jet stream. The positive phase is associated with an enhanced East Asian jet stream and with an eastward shift in the jet exit region toward the western United States. The positive phase of the PNA pattern is associated with above-average temperatures over western Canada and the extreme western United States, and below-average temperatures across the south-central and southeastern U.S. The PNA tends to have little impact on surface temperature variability over North America during summer. The associated precipitation anomalies include above-average totals in the Gulf of Alaska extending into the pacific northwestern United States, and below-average totals over the upper Midwestern United States.

Although the PNA pattern is a natural internal mode of climate variability, it is also strongly influenced by the ENSO phenomenon. The positive phase of the PNA pattern tends to be associated with Pacific warm episodes (El Niño-Southern Oscillation), and the negative phase tends to be associated with Pacific cold episodes (La Niña).

Scripps Institution of Oceanography at the University of California San Diego (Scripps) and California Department of Water Resources (CDWR) Soda Springs Meadow Meteorological Station at Devils Postpile (DEPO)

On August 17, 2005, the National Park Service and a team of partner agencies began construction of a new weather and river stage monitoring station at Soda Springs Meadow, located in the Middle Fork floodplain near the DEPO ranger station. The multi-agency collaboration included NPS staff from DEPO and SIEN, and representatives from the Scripps, and the California Cooperative Snow Surveys (part of the CDWR). The station features two separate arrays of sensors, data loggers, and transmission equipment, operated separately by Scripps and CDWR. Sensor maintenance and data delivery to public websites are also separate. (The operator is sometimes listed as "National Park Service." DEPO personnel are not, however, involved in station maintenance).

Scripps invested in the installation to augment its regional network of stations supporting long-term climate research through its Hydroclimate Weather Observation Program. Regular maintenance and data quality control procedures ensure a high standard of completeness and precision of the data coming from the Scripps suite of sensors. Real-time data access offers important efficiencies but is not the highest priority in this research context; high temporal resolution and data quality are of greater priority, although the data has not undergone rigorous

QA/QC procedures at the time of the publication of this protocol. Douglas Alden is the Scripps researcher responsible for the station.

The CDWR Cooperative Snow Surveys office uses station data as a weather forecasting and early warning tool in the San Joaquin River system as the river flows west toward a series of dammed reservoirs. This station is part of a regional network of real-time stations providing snowpack, streamflow, and meteorological data to forecasters. High temporal resolution data and real-time access are of highest value to CDWR. At the time of this document's publication, Frank Gehrke was responsible for the CDWR investment. The full period-of-record uncorrected data from CDWR sensors on the Soda Springs Meadow weatherstation are available online at the California Data Exchange Center (CDEC). CDEC is a clearinghouse for weather and water data and does not conduct any quality checks on the data it serves. Real-time viewing tools help with visualization, but downloading historical data is tedious and slow because of the way the database query is structured. Sensor error values are inconsistently recorded (some are confusingly recorded as "0"; see air temperature during the summer of 2008) which demands some attention by raw data users. Precipitation values are cumulative, requiring careful manipulation of the raw data before analysis. As a whole, the data are relatively complete, of high quality, and available in real-time making them useful not only to the CDWR but to the monument's staff and visitors

Appendix A. Literature Cited

Climate Prediction Center. 2005. Description of the Pacific / North American Pattern. National Weather Service, Camp Springs, Maryland. Available from: http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/pna.shtml (accessed 09 April 2012).

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., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris. 2002. A knowledge-based approach to the statistical mapping of climate. Climate Research 22:99–113.

Doggett, M., C. Daly, J. Smith, W. Gibson, G. Taylor, G. Johnson, and P. Pasteris. 2004. High-resolution 1971–2000 mean monthly temperature maps for the western United States. Paper 4.3 *in* Combined preprints of the proceedings of the 14th American Meteorological Society Conference on Applied Climatology, 84th American Meteorological Society Annual Meeting, 13–16 January 2004, Seattle, Washington, USA. American Meteorological Society, Boston, Massachusetts.

Gibson, W. P., C. Daly, T. Kittel, D. Nychka, C. Johns, N. Rosenbloom, A. McNab, and G. Taylor. 2002. Development of a 103-year high-resolution climate data set for the conterminous United States. American Meteorological Society 13th Annual Conference on Applied Climatology, Portland, Oregon. 2002:181-183.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society 78:1069–1079.

National Weather Service. 1989. Observing handbook no.2: Cooperative station observations. Online. (http://www.nws.noaa.gov/om/coop/Publications/coophandbook2.pdf). National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland.

National Weather Service. 2009. Cooperative observer program – Resources for observers. Available from: http://www.nws.noaa.gov/om/coop/ (accessed 1 June 2011).

Oregon State University. 2007. PRISM climate group. Online. (http://www.prism.oregonstate.edu). Accessed 1 June 2011.

Rocky Mountain Climate Working Group. 2010. Rocky Mountain climate protocol: Climate monitoring in the Greater Yellowstone and Rocky Mountain Inventory and Monitoring Networks, Version 1.0. Natural Resource Report NPS/IMRO/NRR—2010/222. National Park Service, Fort Collins, Colorado.

Webb, R., H, R. Hereford, and G. J. McCabe. 2000. Climatic fluctuations, drought, and flow in the Colorado River. Pages 57–68 *in* The state of the Colorado River ecosystem in Grand Canyon, a report of the Grand Canyon Monitoring and Research Center 1991–2004. U.S. Geological Survey Circular no. 1282.

Zhang, C. 2005. Madden-Julian Oscillation. Review of Geophysics 43:1-36.

Zhang, Y., J. M. Wallace, and D. S. Battisti. 1997. ENSO-like interdecadal variability: 1900–93. Journal of Climate 10:1004–20.

Appendix B: Weather Stations, Snow Courses, SNOTEL Stations, and USGS Streamgages in and near the Sierra Nevada Network Parks

Acronyms

CARB	California Air Resources Board
CASTNet	Clean Air Status and Trends Network
COOP	Cooperative Observer Network
CDWR	California Department of Water Resources
CDEC	California Data Exchange Center
DRI	Desert Research Institute
GPMP	Gaseous Pollutant Monitoring Program
NADP	National Atmospheric Deposition Program
NPS	National Park Service
NWS	National Weather Service
POMS	Portable Ozone Monitoring System
RAWS	Remote Automated Weather System
SIO	Scripps Institution of Oceanography
UCSB	University of California – Santa Barbara
USACE	U. S. Army Corps of Engineers
USGS	U. S. Geological Survey
ENSO	El Niño-Southern Oscillation

Introduction

This appendix presents a comprehensive list of weather stations in the SIEN as of June 2011 and contains the most complete weather station metadata currently available. This list is a revision of that presented in the SIEN climate inventory report (Davey et al. 2007). While information on weather stations is bound to change over time and although there may be some inaccuracies in the metadata below, we endeavor to document the state of weather and climate monitoring in the network at the time of protocol publication. SIEN maintains communication with the operators of all weather stations in the network parks and will update this list as necessary. Also presented here are lists of snow survey stations and streamgages in and surrounding the SIEN.

NWS-COOP Weather Stations

There are many NWS-COOP weather stations located within 40 miles of the SIEN parks that are no longer in operation or are in a geographic area with conditions that are not comparable to those seen in the parks. The seven stations within the park that are currently in operation and will be used for annual status reports are listed in bold text in Table B1, along with other stations in the park that will not be used for reports. Data from these seven stations will be archived annually and stored in the climate database. A list of additional stations that are outside the boundaries of the SIEN parks can be seen in Davey et al. (2007) or by visiting the National Climatic Data Center website http://www.ncdc.noaa.gov/oa/climate/stationlocator.html, typing in a station from Table B1 and clicking on “List stations within 5, 10, 25 or 50 miles of this station”. The climatologist preparing the trend report may elect to incorporate data from stations in areas surrounding the parks in order to provide an overview of climatic trends in the region with better spatial coverage; the goal of the status reports is to provide a concise summary which necessitates limiting the number of stations included.

Snow course and SNOTEL stations

Currently, there are 38 snow courses (some with snow pillows) and five snow pillows in the SIEN parks and an additional 55 snow courses within the major SIEN watersheds. Snow pillow stations typically have some meteorological (met) equipment, such as wind speed and temperature sensors. A small number of snow courses also have these sensors. Table B2 includes all snow survey stations in the SIEN watersheds, their elevation, watershed, and whether or not they record other meteorological data.

The CDWR records the SWE on the first of each month from January through June for each active snow course. The yearly course report lists the April 1 SWE at each station as well as the percent of the historic average for each course and each major watershed. SIEN will acquire the SWE data from CDWR, format it, and upload it for storage in the Climate Database. For reporting purposes, SIEN will primarily rely upon the watershed averages calculated by the CDWR and posted on their website. The annual snow report can be downloaded from the CDEC website yearly at http://cdec.water.ca.gov/cgi-progs/rpts1/COURSES. Only data from snow courses are included in the course report. Snow courses are manually surveyed multiple times each year. Some snow courses have automated snow pillows with daily data that is sent by satellite and updated to the CDEC website. There are also 5 snow pillows located in the SIEN that are not surveyed as snow courses. These pillows are not included in the course report or the watershed summaries.

There are multiple SNOTEL stations located northeast of Yosemite National Park. Data from these stations will not be acquired because the stations are not located in the SIEN parks and would provide data for just one point that could not be easily correlated to current SIEN conditions due to the patchiness of snow throughout the Sierra Nevada.

Streamgages

The USGS operates four streamgages in the SIEN and reviews the records for additional gages that are operated by the electric company Southern California Edison (SCE). The USGS posts the records for these gages on their website, the National Water Information System (NWIS). SIEN will include some or all of the gages from Table B3 in reports.

Table B1. Weather stations in the Sierra Nevada Network parks. Stations included in annual status reports are in bold font.

PARK	NAME	LAT	LONG	NETWORK	START_DATE	END_DATE	ELEV_M	ELEV_FT
SEKI	Lodgepole Lift Station	36.6041	-118.7330	APRSWXNET	11/30/2010	Present	2039	6690
SEKI	Sequoia NP - Lower Kaweah	36.5620	-118.7690	CARB	M	Present	1900	6234
SEKI	Sequoia & Kings Canyon NP	36.4880	-118.8270	CARB	M	Present	561	1841
YOSE	Yosemite NP-Turtleback Dome	37.7110	-119.7060	CARB	M	Present	1611	5285
YOSE	Yosemite Village-Visitor Cntr	37.7490	-119.5870	CARB	M	Present	1213	3980
YOSE	Yosemite NP-Merced River	37.7430	-119.5940	CARB	M	Present	1220	4003
SEKI	Sequoia/Kings Canyon NP	36.4292	-118.7625	CASTNet	2/1/1997	12/31/2004	1225	4019
SEKI	Ash Mountain	36.4894	-118.8269	CASTNet	7/1/1999	Present	457	1499
YOSE	Turtleback Dome	37.7133	-119.7061	CASTNet	1/1/1994	Present	1605	5266
YOSE	WAWONA	37.5333	-119.6667	COOP	1/1/1934	6/30/1941	1190	3904
YOSE	OSHAUGHNESSY DAM	37.9500	-119.7833	COOP	11/1/1946	5/5/1948	1220	4003
YOSE	MIGUEL MEADOWS	37.9667	-119.8333	COOP	11/1/1946	6/30/1948	1617	5305
SEKI	BULLFROG LAKE	36.7667	-118.4000	COOP	7/1/1948	7/31/1955	3264	1070
SEKI	SUGARLOAF MEADOW	36.7167	-118.6667	COOP	7/1/1948	8/31/1957	2196	7205
SEKI	LEWIS CREEK KINGS CN	36.8000	-118.6833	COOP	11/1/1945	7/25/1961	1418	4652
SEKI	CEDAR GROVE	36.7833	-118.6667	COOP	12/15/1940	5/23/1963	1418	4652
SEKI	EAST VIDETTE MEADOW	36.7333	-118.3833	COOP	4/28/1949	8/31/1964	3172	1040
SEKI	GRANITE BASIN	36.8667	-118.6000	COOP	7/1/1948	8/31/1964	3050	1000
SEKI	GIANT FOREST	36.5667	-118.7667	COOP	6/6/1921	11/8/1968	1955	6414
SEKI	MINERAL KING	36.4333	-118.5833	COOP	8/1/1956	7/31/1969	2434	7986
SEKI	PEAR LAKE	36.6000	-118.6667	COOP	8/1/1956	9/30/1969	2959	9708
YOSE	PARADISE MEADOW	38.0500	-119.6667	COOP	8/1/1948	9/30/1971	2349	7707
YOSE	BEEHIVE MEADOW	38.0000	-119.7833	COOP	7/1/1948	9/30/1971	1983	6506
SEKI	DUSY BENCH	37.1000	-118.5833	COOP	7/1/1948	11/30/1972	2888	9475
SEKI	CHAGOOPA	36.5000	-118.4500	COOP	7/1/1964	11/30/1972	3154	1034
YOSE	LAKE ELEANOR	37.9667	-119.8833	COOP	10/19/1909	2/28/1973	1421	4662

Table B1. Weather stations in the Sierra Nevada Network parks (continued).

PARK	NAME	LAT	LONG	NETWORK	START_DATE	END_DATE	ELEV_M	ELEV_FT
YOSE	GRACE MEADOW	38.1500	-119.6000	COOP	7/1/1948	6/30/1973	2715	8907
SEKI	VIDETTE MEADOW	36.7500	-118.4167	COOP	8/1/1964	9/30/1974	2898	9508
SEKI	MORAINE CREEK	36.7167	-118.5667	COOP	8/1/1964	9/30/1974	2696	8845
YOSE	BADGER PASS	37.6667	-119.6667	COOP	7/1/1948	6/30/1976	2227	7306
SEKI	MITCHELL MEADOW	36.7333	-118.7167	COOP	8/1/1957	9/30/1976	3020	9908
SEKI	STATE LAKES	36.9333	-118.5833	COOP	7/1/1955	9/30/1976	3142	1030
SEKI	RATTLESNAKE CREEK	36.9833	-118.7167	COOP	6/1/1961	9/30/1976	3020	9908
SEKI	GIANT FOREST RADIO	36.5667	-118.7667	COOP	9/1/1965	9/30/1976	2028	6654
SEKI	CRABTREE MEADOW	36.5667	-118.3500	COOP	7/1/1948	9/30/1976	3264	1070
SEKI	HOCKETT MEADOWS	36.3667	-118.6500	COOP	8/1/1959	9/30/1976	2593	8507
YOSE	OSTRANDER LAKE	37.6333	-119.5500	COOP	7/1/1948	9/30/1976	2623	8606
YOSE	SNOW FLAT	37.8333	-119.5000	COOP	7/1/1948	9/30/1976	2654	8707
SEKI	ATWELL	36.4667	-118.6667	COOP	6/24/1948	10/1/1976	1976	6483
YOSE	HODGDON MEADOW	37.8000	-119.8667	COOP	6/1/1967	6/30/1978	1281	4203
YOSE	WAWONA RS	37.5400	-119.6522	COOP	10/1/1940	6/15/2006	1215	3986
SEKI	**GRANT GROVE**	**36.7394**	**-118.9631**	**COOP**	**7/1/1940**	**Present**	**2012**	**6601**
SEKI	**LODGEPOLE**	**36.6044**	**-118.7325**	**COOP**	**2/1/1951**	**Present**	**2053**	**6736**
SEKI	**ASH MTN**	**36.4914**	**-118.8253**	**COOP**	**1/1/1927**	**Present**	**521**	**1709**
YOSE	**SOUTH ENTR YOSEMITE**	**37.5075**	**-119.6336**	**COOP**	**7/1/1941**	**Present**	**1566**	**5138**
YOSE	**YOSEMITE PARK HQ**	**37.7500**	**-119.5897**	**COOP**	**1/12/1905**	**Present**	**1209**	**3967**
YOSE	**HETCH HETCHY**	**37.9614**	**-119.7831**	**COOP**	**10/1/1910**	**Present**	**1180**	**3871**
YOSE	**TUOLUMNE MEADOWS**	**37.8833**	**-119.3500**	**COOP**	**10/21/2008**	**Present**	**2638**	**8655**
YOSE	K6IXA-2 Yosemite Park	37.7230	-119.5750	CWOP	M	Present	2469	8100
YOSE	Gin Flat TC tower	37.7669	-119.7731	DRI	10/1/2003	Present	2149	7051
YOSE	Turtleback Dome GPMP	37.7133	-119.7061	GPMP	1/1/1992	12/1/1992	1605	5266
YOSE	Yosemite Valley GPMP	37.7503	-119.5869	GPMP	12/1/1989	11/30/1994	1219	3999
SEKI	Grant Grove GPMP	36.7397	-118.9611	GPMP	12/1/1989	12/31/1995	2012	6601

Table B1. Weather stations in the Sierra Nevada Network parks (continued).

PARK	NAME	LAT	LONG	NETWORK	START_DATE	END_DATE	ELEV_M	ELEV_FT
SEKI	Ash Mountain GPMP	36.4936	-118.8286	GPMP	6/1/1982	10/1/1996	610	2001
YOSE	Wawona Valley GPMP	37.5358	-119.6517	GPMP	4/1/1987	10/1/1996	1220	4003
YOSE	Merced River GPMP	37.7431	-119.5939	GPMP	8/1/2002	12/31/2005	1219	3999
SEKI	Lower Kaweah GPMP	36.5658	-118.7772	GPMP	6/1/1984	Present	1890	6201
YOSE	YOSE Valley Ozone schoolyard	37.7503	-119.5869	GPMP	6/7/2006	Present	1219	3999
SEKI	Sequoia NP-Giant Forest	36.5669	-118.7772	NADP	7/22/2003	Present	1902	6240
YOSE	Yosemite NP-Hodgdon Meadow	37.7961	-119.8581	NADP	12/8/1981	Present	1408	4619
SEKI	Elk Creek	36.5133	-118.8092	NPS	1/1/1983	12/31/2000	0	0
SEKI	MEWSS	36.5538	-118.7520	NPS	M	M	0	0
YOSE	Snow Flat	37.8270	-119.4970	NPS-YOSE	1/1/1988	Present	2652	8701
YOSE	Tioga Pass	37.9108	-119.2586	POMS	7/20/2005	Present	3037	9964
SEKI	Milk Ranch	36.4867	-118.7800	RAWS	8/1/1997	8/31/1999	1897	6224
SEKI	Sugarloaf	36.7267	-118.6750	RAWS	7/1/1992	Present	2475	8120
SEKI	Cedar Grove	36.7878	-118.6561	RAWS	9/1/1999	Present	1439	4721
SEKI	Park Ridge	36.7242	-118.9425	RAWS	7/1/1997	Present	2298	7539
SEKI	ASH MOUNTAIN	36.4914	-118.8253	RAWS	12/1/2004	Present	527	1729
SEKI	Wolverton	36.4450	-118.7033	RAWS	6/1/1996	Present	1597	5240
SEKI	Rattlesnake	36.4069	-118.4217	RAWS	7/1/1992	Present	2621	8599
YOSE	White Wolf	37.8511	-119.6500	RAWS	8/1/1988	Present	2446	8025
YOSE	Gaylor Meadow	37.8683	-119.3183	RAWS	8/1/1988	Present	2825	9268
YOSE	Mariposa Grove	37.5128	-119.6047	RAWS	9/1/1988	Present	1951	6401
YOSE	Crane Flat Lookout	37.7617	-119.8247	RAWS	11/1/1991	Present	2025	6644
YOSE	Wawona	37.5400	-119.6517	RAWS	12/1/2004	Present	1235	4052
DEPO	**Soda Springs Meadow**	**37.6293**	**-119.0847**	**SIO**	**9/1/2006**	**Present**	**2307**	**7569**
SEKI	Topaz Lake	36.6253	-118.6394	UCSB	10/1/1996	Present	3221	1056
SEKI	Emerald Lake	36.5978	-118.6744	UCSB	7/1/1985	Present	2808	9213
SEKI	M3	36.6103	-118.6472	UCSB	3/1/1994	Present	3232	1060

Table B1. Weather stations in the Sierra Nevada Network parks (continued).

PARK	NAME	LAT	LONG	NETWORK	START_DATE	END_DATE	ELEV_M	ELEV_FT
SEKI	Mitchell Meadow	36.7369	-118.7119	USACE	8/1/1988	Present	3018	9902
SEKI	State Lakes	36.9269	-118.5739	USACE	8/1/1988	Present	3139	1029
SEKI	Upper Tyndall Creek	36.6500	-118.3969	USACE	8/1/1988	Present	3475	1140
SEKI	Giant Forest	36.5619	-118.7650	USACE	8/1/1988	Present	2027	6650
SEKI	Atwell Camp	36.4640	-118.6310	USACE	10/01/1949	Present	1951	6400

Table B2. Snow survey stations in the major SIEN watersheds. There are 98 stations (94 courses and 4 stations with snow pillows only) within the SIEN watersheds; 43 stations are within the parks.

In Park?	Has Met sensors	RIVER	NAME	CDEC_ID	ELEV (FT)	AGENCY
Yes	Yes (+pillow)	KAWEAH R	FAREWELL GAP	FRW	9500	Sequoia and Kings Canyon National Parks
Yes	No	KAWEAH R	PANTHER MEADOW	PTM	8600	Sequoia and Kings Canyon National Parks
Yes	Yes	KAWEAH R	HOCKETT MEADOWS	HKM	8500	Sequoia and Kings Canyon National Parks
Yes	No	KAWEAH R	MINERAL KING	MNK	8000	Sequoia and Kings Canyon National Parks
Yes	No	KAWEAH R	GIANT FOREST	GFR	6400	Sequoia and Kings Canyon National Parks
Yes	No	KERN R	BIGHORN PLATEAU	BGH	11350	CA Dept of Water Resources
Yes	No	KERN R	SIBERIAN PASS	SIB	10900	CA Dept of Water Resources
Yes	Yes (+pillow)	KERN R	CRABTREE MEADOW	CBT	10700	CA Dept of Water Resources
Yes	No	KERN R	GUYOT FLAT	GYF	10650	CA Dept of Water Resources
Yes	No	KERN R	SANDY MEADOWS	SDM	10650	CA Dept of Water Resources
Yes	No	KERN R	TYNDALL CREEK	TND	10650	CA Dept of Water Resources
Yes	No	KERN R	ROCK CREEK	RCR	9600	CA Dept of Water Resources
Yes	No	KERN R	QUINN RANGER STATION	QRS	8350	Sequoia and Kings Canyon National Parks
Yes	Yes	KINGS R	BISHOP PASS	BSH	11200	CA Dept of Water Resources
Yes	No	KINGS R	CHARLOTTE RIDGE	CLT	10700	CA Dept of Water Resources
Yes	No	KINGS R	BULLFROG LAKE	BLF	10650	CA Dept of Water Resources
Yes	No	KINGS R	BENCH LAKE	BNH	10600	Kings River Water Association
Yes	No	KINGS R	RATTLESNAKE CREEK BA	RTT	9900	Pacific Gas and Electric Company, Auberry
Yes	No	KINGS R	SCENIC MEADOW	SCE	9650	Kings River Water Association
Yes	No	KINGS R	RIDGE TRAIL	RGT	7500	Sequoia and Kings Canyon National Parks
Yes	No	MERCED R	SNOW FLAT	SNF	8700	Yosemite National Park
Yes	Yes (+pillow)	MERCED R	OSTRANDER LAKE	STR	8200	Yosemite National Park
Yes	Yes (+pillow)	MERCED R	TENAYA LAKE	TNY	8150	Yosemite National Park
Yes	No	MERCED R	PEREGOY MEADOWS	PGM	7000	Yosemite National Park

Table B2. Snow survey stations in the major SIEN watersheds (continued).

In Park?	Has Met sensors	RIVER	NAME	CDEC_ID	ELEV (FT)	AGENCY
Yes	No	MERCED R	GIN FLAT (COURSE)	GFL	7000	Yosemite National Park
Yes	No	SAN JOAQUIN	EMERALD LAKE	EML	10600	Southern California Edison Company, Big Creek
Yes	No	SAN JOAQUIN	COLBY MEADOW	CBM	9700	Southern California Edison Company, Big Creek
Yes	Yes (+pillow)	TUOLUMNE R	DANA MEADOWS	DAN	9800	Yosemite National Park
Yes	No	TUOLUMNE R	RAFFERTY MEADOWS	RFM	9400	Yosemite National Park
Yes	No	TUOLUMNE R	NEW GRACE MEADOW	NGM	8900	Yosemite National Park
Yes	Yes (+stage)	TUOLUMNE R	TUOLUMNE MEADOWS	TUM	8600	Yosemite National Park
Yes	No	TUOLUMNE R	WILMA LAKE	WLW	8000	Yosemite National Park
Yes	No	TUOLUMNE R	SACHSE SPRINGS	SAS	7900	Summit Ranger District
Yes	No	TUOLUMNE R	SPOTTED FAWN	SPF	7800	Summit Ranger District
Yes	Yes (+pillow)	TUOLUMNE R	PARADISE MEADOW	PDS	7650	Yosemite National Park
Yes	No	TUOLUMNE R	VERNON LAKE	VNN	6700	Yosemite National Park
Yes	No	TUOLUMNE R	BEEHIVE MEADOW	BHV	6500	Yosemite National Park
Yes	Yes	SAN JOAQUIN R	Soda Springs Meadow	DPO		Devils Postpile National Monumnent
No (>20 miles)	No	KERN R	BONITA	BNM		
No (>20 miles)	No	KERN R	DEAD HORSE MEADOW	DHM		
No (>20 miles)	No	SAN JOAQUIN	NELLIE LAKE	NLL		
No (>20 miles)	No	SAN JOAQUIN	Tamarack Creek	TMK		
No (>20 miles)	No	SAN JOAQUIN	HUNTINGTON LAKE	HTT		
No (<20 miles)	No	KERN R	COTTONWOOD PASS	CWP	11050	CA Dept of Water Resources
No (<20 miles)	No	KERN R	BIG WHITNEY MEADOW	BWH	9750	CA Dept of Water Resources
No (<20 miles)	No	KERN R	RAMSHAW MEADOWS	RMM	8700	CA Dept of Water Resources
No (<20 miles)	No	KERN R	LITTLE WHITNEY MEADOW	LWM	8500	CA Dept of Water Resources
No (<20 miles)	Yes (+pillow)	KERN R	CASA VIEJA MEADOWS	CSV	8300	Cannell Meadow Ranger District

Table B2. Snow survey stations in the major SIEN watersheds (continued).

In Park?	Has Met sensors	RIVER	NAME	CDEC_ID	ELEV (FT)	AGENCY
No (<20 miles)	No	KERN R	BEACH MEADOWS	BHM	7650	Cannell Meadow Ranger District
No (<20 miles)	Yes (+pillow)	KINGS R	BLACKCAP BASIN	BCB	10300	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	BEARD MEADOW	BMD	9800	Pacific Gas and Electric Company, Auberry
No (<20 miles)	Yes (+pillow)	KINGS R	UPPER BURNT CORRAL	UBC	9700	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	ROUND CORRAL	RDC	9000	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	ROWELL MEADOW	RWM	8850	Kings River Water Association
No (<20 miles)	No	KINGS R	WOODCHUCK MEADOW	WDH	8800	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	LONG MEADOW (KINGS R)	LMD	8500	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	COURTRIGHT	CUR	8350	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	STATUM MEADOW	SMD	8300	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	HELMS MEADOW	HLM	8250	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	POST CORRAL MEADOW	PRM	8200	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	DODSONS MEADOW	DSM	8050	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	KINGS R	HORSE CORRAL MEADOW	HCM	7600	Kings River Water Association
No (<20 miles)	No	KINGS R	BIG MEADOWS	BMS	7600	Kings River Water Association
No (<20 miles)	No	KINGS R	FRED MEADOW	FDM	6950	Pacific Gas and Electric Company, Auberry
No (<20 miles)	No	SAN JOAQUIN	MONO PASS	MNP	11450	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	PIUTE PASS	PPS	11300	CA Dept of Water Resources
No (<20 miles)	No	SAN JOAQUIN	PIONEER BASIN	PNB	10400	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	HEART LAKE	HRT	10100	Southern California Edison Company, Big Creek
No (<20 miles)	Yes (+pillow)	SAN JOAQUIN	VOLCANIC KNOB	VLC	10050	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	ROSE MARIE	RMR	10000	Southern California Edison Company, Big Creek
No (<20 miles)	Yes (+pillow)	SAN JOAQUIN	AGNEW PASS	AGP	9450	Southern California Edison Company, Bishop
No (<20 miles)	No	SAN JOAQUIN	DUTCH LAKE	DTL	9100	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	KAISER PASS	KSR	9100	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	COYOTE LAKE	CYT	8850	Southern California Edison Company, Big Creek

Table B2. Snow survey stations in the major SIEN watersheds (continued).

In Park?	Has Met sensors	RIVER	NAME	CDEC_ID	ELEV (FT)	AGENCY
No (<20 miles)	No	SAN JOAQUIN	CORA LAKES	CRA	8400	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	BADGER FLAT	BDF	8300	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	LAKE THOMAS A EDISON	THE	7800	Southern California Edison Company, Big Creek
No (<20 miles)	No	SAN JOAQUIN	CHILKOOT LAKE	CKT	7450	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	FLORENCE LAKE	FLC	7200	Southern California Edison Company, Big Creek
No (<20 miles)	Yes (+pillow)	SAN JOAQUIN	CHILKOOT MEADOW	CHM	7150	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	CLOVER MEADOW	CLM	7000	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	JACKASS MEADOW	JCM	6950	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	CHIQUITO CREEK	CHQ	6800	Minarets Ranger District
No (<20 miles)	No	SAN JOAQUIN	POISON MEADOW	PMD	6800	Minarets Ranger District
No (<20 miles)	No	TULE R	QUAKING ASPEN	QKA	7000	Tule River Ranger District
No (<20 miles)	No	TULE R	OLD ENTERPRISE MILL	OEM	6600	CA Dept of Forestry
No (<20 miles)	No	TUOLUMNE R	BOND PASS	BNP	9300	Summit Ranger District
No (<20 miles)	Yes (+pillow)	TUOLUMNE R	HORSE MEADOW	HRS	8400	Summit Ranger District
No (<20 miles)	No	TUOLUMNE R	HUCKLEBERRY LAKE	HCL	7800	Summit Ranger District
No (<20 miles)	No	TUOLUMNE R	KERRICK CORRAL	KRC	7000	Summit Ranger District
No (<20 miles)	No	TUOLUMNE R	UPPER KIBBIE RIDGE	UKR	6700	Summit Ranger District
No (<20 miles)	No	TUOLUMNE R	LOWER KIBBIE	LKB	6700	Summit Ranger District
No (<20 miles)	No	TUOLUMNE R	BELL MEADOW	BEM	6500	Summit Ranger District
Yes	yes	KERN R	Chagoopa Plateau	CHP		Not a snow course, snow pillow only
Yes	yes	KINGS R	Charlotte Lake	CRL		Not a snow course, snow pillow only
Yes	yes	MERCED R	Merced Lake	MLK		Not a snow course, snow pillow only
Yes	yes	MERCED R	Dog House Meadow	DGH		Not a snow course, snow pillow only
Yes	yes	TUOLUMNE R	Slide Canyon	SLI		Not a snow course, snow pillow only

Table B3. USGS streamgages in or near SIEN that will be used in SIEN climate reports.

Station Name	Station Number	Period of Record	Watershed	In Park?
Merced River at Happy Isles Bridge	11264500	1916-Present	Merced	Yes – YOSE
Merced River at Pohono Bridge	11266500	1917-Present	Merced	Yes - YOSE
Tuolumne River above Hetch Hetchy Reservoir	11274790	2006-Present	Tuolumne	Yes – YOSE
Middle Fork San Joaquin in DEPO	11224000	2009-Present	San Joaquin	Yes – DEPO
Kern River near Kernville	11186000	1911-Present	Kern	No – Near south border of SEQU

Appendix B. Literature Cited

Davey, C. A., K. T. Redmond, and D. B. Simeral. 2007. Weather and climate inventory, National Park Service, Sierra Nevada Network. Natural Resource Technical Report NPS/SIEN/NRTR—2007/042. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?Code=649248 (accessed 09 April 2012).

Appendix C: Administrative Record

History of protocol development

The purpose of the Administrative Record is to provide a history of protocol development and refinement. A timeline of events is presented in Table C1. The protocol lead should update the record as major changes are made or milestones are achieved. The original protocol development summary follows the timeline of events.

Table C1. Timeline of protocol development.

Date	Event
October 2006	Work group meeting to discuss protocol development strategy.
Dec 2007	Weather and Climate Inventory Report Complete.
August 2008	Redmond and Edwards complete draft SIEN Climate Assessment.
Sept 2009	Workgroup meeting to discuss goals and objectives.
May 2010	Jeff Balmat on detail to SIEN. Start of protocol writing.
June 2010	Jennie Skancke assumes responsibility for protocol. Final Protocol Objectives sent to work group for approval.
Nov 2010	SIEN begins cooperation with IMD and ROMN/YELL on Climate Database. IMD SOPs in development and review by ROMN/YELL and SIEN.
March 2011	Final drafts of Climate Protocol narrative and SOPs are complete and submitted to Program Manager Alice Chung-MacCoubrey for review. Significant revisions occur based on discussion among and editing by Jennie S., Alice C. and the SIEN data manager, Les Chow.
July 2011	Submission of protocol to PWR peer review process.
January 2012	Peer review returned to SIEN. Moderate revisions needed.
April 2012	Submission of peer review response to PWR review coordinator
May/June 2012	Protocol returned to SIEN with minor revisions needed and made. Kittel (2009) added to protocol as an appendix.

Early stages of protocol development included inventories of relevant data sources. The Western Regional Climate Center (WRCC) developed a weather and climate inventory report for SIEN as part of a national level collaboration with the NPS I&M program (Davey et al. 2007). The report provides a nearly complete inventory of point-based monitoring within and around each park and provides an overview of climate within the Sierra Nevada region.

The WRCC also collaborated with SIEN to develop a draft report assessing the status of Sierra Nevada climate monitoring, analysis of the quality of existing weather datasets, and recommendations for monitoring (Edwards and Redmond 2011). A workshop was convened in October 2006 to discuss park and network needs and give direction to this report as well as SIEN's protocol development strategy. The expert opinion supplied by the report provides justification for much of the approach we are taking toward monitoring and reporting. Additional direction has been provided along the way by the authors of the Northern Coast and Cascade Network Climate Protocol (Lofgren et al. 2010), Rocky Mountain Climate Protocol (Rocky

Mountain Climate Working Group 2010), and Northern Colorado Plateau Network Climate Protocol (Garman et al. 2004) . Significant portions of these existing protocols were utilized in the development of this protocol.

In 2009, SIEN began working in earnest on the development of this protocol. The climate protocol working group, along with experts from weather monitoring agencies and partners, met in 2010 to approve monitoring goals and objectives. This protocol closely follows the Rocky Mountain Climate Protocol (Rocky Mountain Climate Working Group 2010) which went through a lengthy development process and relied upon multiple climate experts. In late 2010, the NPS Inventory and Monitoring Division (IMD) staff began exploring tools to assist networks with their climate monitoring, including an Enterprise Climate Database. At that time, it was determined that such tools would be extremely helpful to multiple networks and parks. SIEN decided to utilize the Climate Database and other tools in development by IMD. In July 2011, SIEN sent the protocol to the Pacific West Region (PWR) peer review process. The protocol was returned in January 2012 with peer review comments and a list of topics that needed to be addressed. SIEN responded to peer review comments in April 2012.

Appendix C. Literature Cited

Davey, C. A., K. T. Redmond, and D. B. Simeral. 2007. Weather and climate inventory, National Park Service, Sierra Nevada Network. Natural Resource Technical Report NPS/SIEN/NRTR—2007/042. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?Code=649248 (accessed 09 April 2012).

Edwards, L. M., and K. T. Redmond. 2011. Climate assessment for Sierra Nevada Network Parks. NPS/SIEN/NR-2011/482. National Park Service, Fort Collins, Colorado. Available from: http://science.nature.nps.gov/im/units/sien/monitoring/Reports/SIENClimateAssessment_EdwardsRedmond_Final_20111230.pdf (accessed 26 March 2012).

Garman, S., M. Beer, M. A. Powell, and R. DenBleyker. 2004. Climate monitoring protocol for the park units in the Northern Colorado Plateau Network. Version 1.00, 15 December 2004. U.S. Department of the Interior, National Park Service, Inventory and Monitoring Program, Northern Colorado Plateau Network, Moab, Utah.

Lofgren, R., B. Samora, B. Baccus, and B. Christoe. 2010. Climate monitoring protocol for the North Coast and Cascades Network (Mount Rainier National Park, Olympic National Park, North Cascades National Park, Lewis and Clark National Historical Park, Ebey's Landing National Historical Reserve, San Juan Island National Historical Park, Fort Vancouver National Historic Site): volume 1. narrative and appendices, version 5/26/2010. Natural Resource Report NPS/NCCN/NRR—2010/240. National Park Service, Fort Collins, Colorado.

Rocky Mountain Climate Working Group. 2010. Rocky Mountain climate protocol: Climate monitoring in the Greater Yellowstone and Rocky Mountain inventory and monitoring networks, Version 1.0. Natural Resource Report NPS/IMRO/NRR—2010/222. National Park Service, Fort Collins, Colorado.

Appendix D: Climate Database Schema, Table and Field Definitions

This document was created by Brent Frakes and Simon Kingston of the Inventory and Monitoring Division. It is available through IRMA at: https://irma.nps.gov/App/Reference/Profile/2170262 and is provided here for reference. The most recent version of this document will always be available from IRMA. We may not revise the protocol each time the database schema is updated.

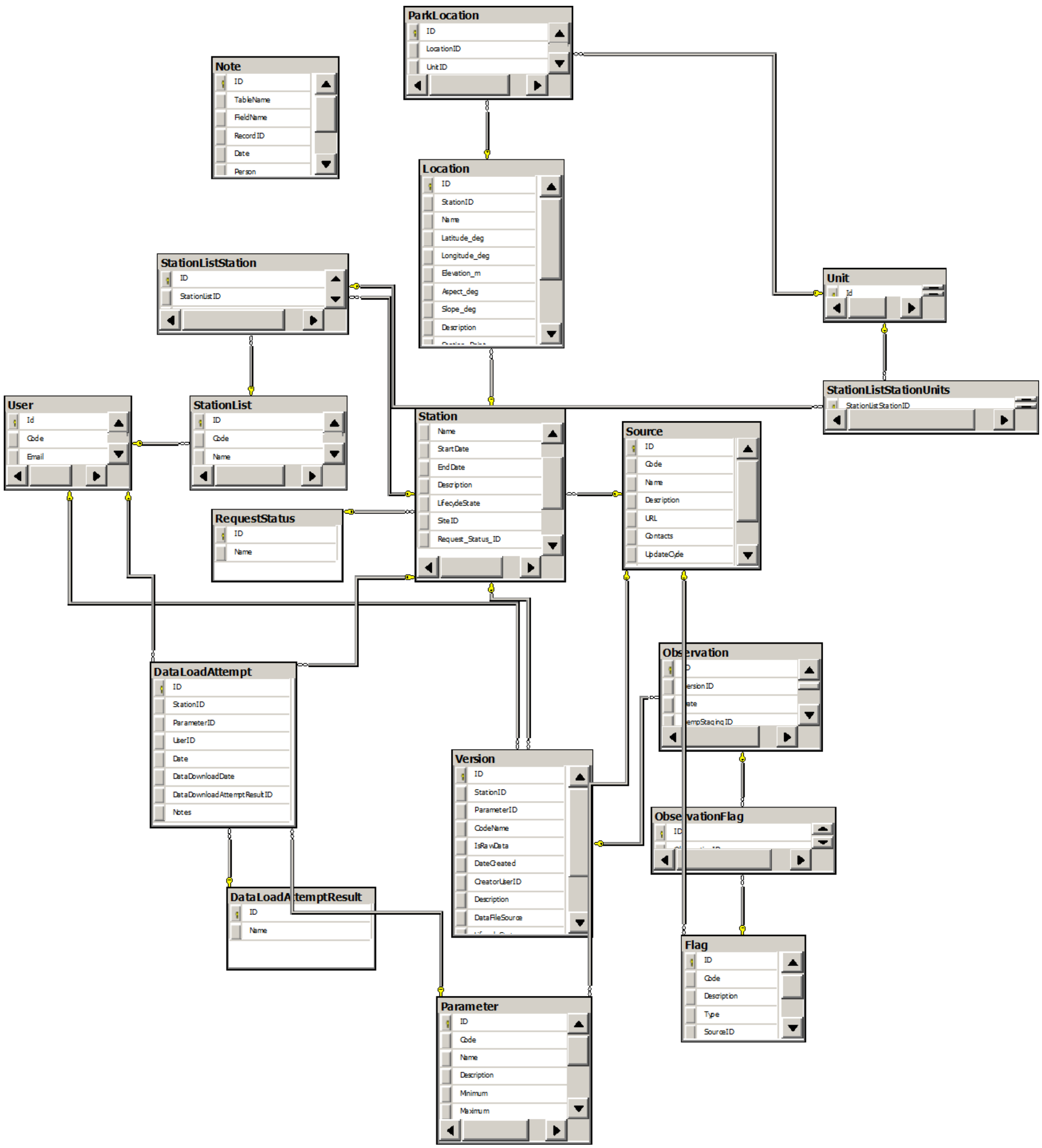
Authors: Brent Frakes; Simon Kingston (Revision date 31 March 2011)

Table of Contents

1. Introduction

This document provides an overview of the database schema and includes both table and field definitions.

2. Database Schema



3. Table and Field Definitions

3.1. DataLoadAttempt

This table contains all information about attempted data loads.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique identifier
StationID	bigint	Yes. FK		Station ID
ParameterID	bigint	Yes. FK		ID of the parameter
UserID	bigint	Yes.FK		ID of the User requesting the data load
Date	datetime	Yes		Date/time the user requested the data be loaded
DataDownloadDate	datetime	No		Date the attempt was made to download the data
DataDownloadAttemptResultID	tinyint	Yes. FK		
Notes	varchar(255)	No		Notes specific to the download attempt

3.2. DataLoadAttemptResult

Lookup table of data load attempt result values that describe success or failure to load data into the climate database.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique identifier
Name	varchar(50)	Yes		Short description of the reason for success or failure at loading data.

3.3. Flag

This table holds all of the flags for each observation. Flags can be related to measurements, quality and or processing.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique identifier
Code	varchar(50)	Yes		A code representing the flag. Flags should generally not be more than 4 characters
Description	varchar(MAX)	Yes		Full description of the flag
Type	varchar(50)	Yes	Data Measure Flag, Data Quality Flag, Data Processing Flag	Measure – Pertaining to the actual collection of the data by the source provider. Includes whether there were issues with the instrumentation or other associated atmospheric events. Examples include there was a severe storm which damaged some of the instrumentation or caused them to go offline. Quality – Pertaining to any quality control applied to the data following collection. Quality control may include modifications to the original observations by the source provider Processing – Pertaining to the processing of the data by the NPS Inventory and Monitoring Division during loading of the data.
SourceID	bigint	Yes. FK		ID from Source table. Indicates which source defined the flag.
Error	bit	Yes	Yes/No	Indicates whether the source considers the flag an error.

3.4. Location

A single station can, over time, be located at one or more locations. This table provides information on each of those locations.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK.		Unique record ID
StationID	bigint	Yes. FK	StationID in the Station table	The station that this location describes.
Name	Varchar(50)	Yes	N/A	Name of the station location. If there is no name provided by the data provider, then create a simple descriptive name.
Latitude_deg	float	Yes	-90 to 90	Latitude in decimal degrees. Points in the northern hemisphere are positive; those in the southern hemisphere are negative(e.g., 40.75).
Longitude_deg	float	Yes	-180 to 180	Longitude in decimal degrees. Points in the western hemisphere are negative; those in the eastern hemisphere are positive (e.g., -100.82).
Elevation_m	float	No	-86.0 to 6194.0	Elevation of the station measured in meters. If the station is not measured in meters, you will need to convert it.
Aspect_deg	smallint	No	0 to 359	Aspect is the direction of the slope where the station is located. North facing slopes have an aspect of 0. If the aspect is not measured in degrees, you will need to convert it.
Slope_deg	tinyint	No	0 to 90	Slope is the gradient of the slope where the station is located. A zero degree slope indicates a completely flat surface. If the slope is not measured in degrees, you will need to convert it.
Description	varchar(max)	No	N/A	Description of the specific station location. Ideally, location description distributed by the data provider. Otherwise, create your own description.
Station_Point	geometry	No		The location of the station as a geographic feature
StartDate	date	No	Between 1/1/1700 and today	This is the date the station began recording data at this particular location. If station didn't move this is the length of record from "StationInfo"
EndDate	date	No	Between 1/1/1700 and today	Use 12/31/9999 if still active location.
StateName	varchar(50)	No	US State Names or other national political boundaries	All full state names or canadian province names
Shape	geography	No		
DatumID	tinyint	No		The datum for the coordinates.
IsActive	bit	No		1 if the station is still active at this location, 0 if it is not, null if indeterminate.

3.5. Note

This table provides a flexible means of allowing any extra notes to be applied to any single record and field within the database.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK	Identity	Unique ID for each note
TableName	varchar(50)	Yes	Table names of climate database	Name of the table for which this note applies
FieldName	varchar(50)	Yes	Field names of respective tables	Name of the table field for which this note applies
RecordID	bigint	Yes		Unique record ID for which this note applies
Date	date	Yes		The date this note was generated
Person	varchar(50)	Yes		The person who created the note
Note	varchar(50)	Yes		The actual note

3.6. Observation

This table houses all of the individual weather/climate observations.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK	N/A	Unique identifier for each observation
VersionID	bigint	Yes. FK	VersionIDs	VersionID from the Version table
TempStagingID	bigint	No		ID from Temp.DataStaging table that is only used as a utility column for importing new data.
Date	date	Yes	N/A	The date of the weather/climate observation
ParameterValue	Float	Yes		The observed value for a given parameter
Comment	nvarchar(max)	No		Any other information that informs a specific observation for a given parameter

3.7. ObservationFlag

Contains all flags specific to an single observation

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique record ID for each flag
ObservationID	bigint	Yes. FK		The observation that the flag pertains to
FlagID	bigint	Yes. FK		The specific flag

3.8. Parameter

Each observation pertains to a particular parameter, whether it is minimum temperature, precipitation or river flow. Each parameter is identified by a parameter code, which includes the units of measure.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique ID for each parameter
Code	varchar(50)	Yes	N/A	The parameter code used by the data provider. In general, we recommend that you do not modify the parameter values if you are getting information from a third party. If the provider does not use codes, you will need to create one.
Name	varchar(50)	Yes	N/A	Name of the parameter being measured
Description	varchar(500)	Yes	N/A	Text description of what parameter is being measured
Minimum	int	Yes		Minimum range of acceptable values for this particular parameter
Maximum	int	Yes		Maximum range of acceptable values for this particular parameter
Unit	varchar(50)	Yes	N/A	The measurement unit described as precisely as possible
SourceID	bigint	Yes. FK		The data source which provided this parameter

3.9. ParkLocation

This table describes how a particular station relates to a specific park.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique ID for each station
LocationID	bigint	Yes. FK		Location of the station
UnitID	int	Yes. FK	NPS Units	Distinct NPS Unitcodes for which this station should be affiliated. Generally, this is a 4-letter code (e.g., ROMO).
InPark	bit	No	True or False	Indicate whether the stationlocation is within the park boundary.
DistanceToParkBoundary_m	smallint	No	Between 0 and 32,767	If not within the park, indicates the distance to the park boundary in meters.

3.10. Request Status

Lookup table containing the request satus codes.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique ID for each station
Name	varchar(50)	Yes. CK		'Can Request' – Indicates that a network can request data for a given status. By default, all national-level stations are given a 'Can Request' status. 'Is Requested' – Indicates that one or more networks have requested that IMD actively synhcronize data aqusition for this station 'Data Not Available' – Indicates that it is not possible for IMD to actively obtain data for this station. This status does not preclude a network from obtaining the data from alternate sources.

3.11. Source

Most weather observations are coordinated by a data provider who is the ultimate source of the weather/climate data. This table houses all of the source information.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique ID for each source data provider
Code	varchar(50)	Yes. CK		Unique source code
Name	varchar(50)	Yes	N/A	The name of the data provider. In many cases, a data provider may have more than one station. An example is “COOP”.
Description	varchar(50)	Yes	N/A	Description of the data provider. Indicate who they are beyond what the name would suggest.
URL	varchar(100)	No	N/A	If there is a URL for the data provider’s web site, indicate that here.
Contacts	varchar(50)	No	N/A	Many data providers will likely have a point of contact, either as a contact name and phone number or perhaps as a URL.
UpdateCycle	varchar(50)	No	N/A	This indicates how often the observed data is updated by the provider. For example, some providers do annual updates while others update the observations monthly. You can approximate the times if you are not sure.
Agency	varchar(50)	Yes	N/A	In a few cases, the data provider may be identified within a larger agency. For example, the data provider may be the climate group of a larger entity known as the agency. An example is “NWS” for the National Weather Service.

3.12. Station

This table contains all of the unique information for a particular weather station.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes. PK		Unique index
CDB_Code	varchar(50)	Yes. CK	N/A	This is the unique code given to station when loaded to the database.
SourceID	bigint	Yes. FK		Code for entity who is the data provider for the climate data.
Code	varchar(50)	Yes		Code given to the station by the data provider.
Name	varchar(255)	No	N/A	This is the full station name given by the data provider. In cases where there is not a name, staff submitting the data are encouraged to create a meaningful name of no more than 255 characters (spaces are allowed).
StartDate	date	No	Between 1/1/1700 and today. Times are ignored.	Known start date when station began recording (this may be different than the available data).
EndDate	date	No	Between 1/1/1700 and today. Times are ignored.	Known end date when station ended recording (this may be different than the available data). Between 1/1/1700 and today. Use 12/31/9999 if still active.
Description	nvarchar(max)	No	N/A	This is a more thorough description of the station. When possible, use the description used by the data provider or create your own.
LifecycleState	varchar(50)	No	Active, Inactive	Indicates whether a station is active, as of this current date, actively collecting climate data
SiteID	varchar(10)	No		Alternate identifier given to the station by the data provider.
RequestStatus	varchar(50)	FK. Yes	1,2,3	Indicates whether it is possible for a user to request the

				IMD automate the pulling and loading of the data to the database. See RequestStatus table.
IsActive	bit	No		1 if the station is still active at this location, 0 if it is not, null if indeterminate.

3.13. StationList

This table contains metadata about the user-defined list of stations described in the StationListStation table. StationList is used for creating lists of stations of interest for an NPS unit.

Field	Type	Required?	Domain	Definition
ID	int	Yes. PK		Unique index
Code	varchar(20)	Yes. CK		Unique value used to refer to a StationList record.
Name	varchar(100)	Yes		Short description of the station list that includes the name of the requestor.
OwnerUserID	bigint	Yes. FK		ID from the User table identifying the individual who owns (requested) the list of stations.
DateCreated	date	Yes		Date that the StationList record was created.

3.14. StationListStation

This table contains the list of stations that were requested in a StationList.

Field	Type	Required?	Domain	Definition
ID	int	Yes. PK		Unique identifier.
StationListID	int	Yes. FK		ID from the StationList table.
StationID	bigint	Yes. FK		ID from the Station table.
DateCreated	date	Yes		Date that the StationListStation records was created.

3.15. StationListStationUnits

This table identifies the NPS unit(s) with which a Station on a Station List has been associated. This association is defined by the creator of the Station List and does not necessarily imply geographical overlap.

Field	Type	Required?	Domain	Definition
StationListStationID	Int	Yes. PK. FK		ID from the StationListStation table.
UnitID	Int	Yes. PK. FK		ID from the Unit table.

3.16. TemperatureBins

This table is a lookup that defines the temperature bins (groups) used by functions that group temperature values.

Field	Type	Required?	Domain	Definition
ID	tinyint	Yes. PK		Unique identifier.
Name	varchar(50)	Yes		Name of the range of the bin.
MinValue	float	No		Minimum value included in the bin.
MaxValue	float	No		Maximum value included in the bin.

3.17. Unit

This table contains a list of all NPS park units.

Field	Type	Required?	Domain	Definition
ID	int	Yes		Unique ID for each NPS unit
UnitCode	varchar(10)	Yes	NPS Park Unit Codes	The accepted NPS unit code as defined by NPS NRInfo Unit Service (See https://nrinfo.nps.gov/Unit.mvc/Search).

3.18. User

This table records all users of the database. Users generally pertain to the technical support staff (i.e., those doing the data loading and management) as opposed to those using the database.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes		Unique ID of each staff
Code	varchar(50)	Yes		Unique code for each staff member. Recommendations are to use the initials
Email	varchar(50)	Yes		Valid email address at time the staff was working
FirstName	varchar(20)	Yes		First name of staff member
LastName	varchar(20)	YEs		Last name of staff member

3.19. Version

A version is defined as a series of observations for a given parameter and station.

Field	Type	Required?	Domain	Definition
ID	bigint	Yes		The unique ID for each version
StationID	FK – bigint	Yes	Station IDs found in the Stations table	The unique stationID
ParameterID	bigint	Yes	Parameter IDs found in the parameter table	The specified parameter (e.g., minimum temperature F, maximum temperature F, etc.)
CodeName	varchar(50)	Yes		
IsRawData	bit	Yes	0/1	Indicates whether the data was loaded directly from the authoritative data source or if it was provided by a third party. A value of 1 indicates a direct load.
DateCreated	DateTime	Yes		The date and time the version was created
CreatorUserID	bigint	Yes	Valid IDs in the User table	The individual creating the new version of the data
Description	varchar(max)	Yes		For non-authoritative source data, the description provides a complete lineage of how the version was obtained and/or processed. For authoritative data, it indicates only that.
DataSourceFile	varchar(255)	No		Provides the name of the original text file used to load the database.
LifecycleState	varchar(50)	Yes	Provisional, Final	Indicates the lifecycle state of the version
DateUpdated	DateTime	No		Date the version was loaded to the database
CDB_Code	varchar(20)	No		Auto-generated code based on the ID value for new records.

Appendix E: The Development and Analysis of Climate Datasets for National Park Science and Management: A Guide to Methods for Making Climate Records Useful and Tools to Explore Critical Questions

Report prepared for National Park Service Inventory and Monitoring Program

Final Report
10 December 2009
(minor revisions as of 23 March 2010)

Author:
Timothy Kittel
Institute of Arctic and Alpine Research (INSTAAR) CB450
University of Colorado
Boulder, CO 80309
kittel@colorado.edu

Please note: The original version of this document contained numerous figures from copyrighted publications. Here, copyrighted figures have been removed, but their captions with citations remain in place. The original document is not for public distribution and can be found at: https://irma.nps.gov/App/Reference/Profile/2169763.

Please cite this report as:
Kittel, T. 2012. Appendix E: The development and analysis of climate datasets for national park science and management: a guide to methods for making climate records useful and tools to explore critical questions. In Skancke, J. R., A. L. Chung-MacCoubrey, and L. S. Chow. 2012. Sierra Nevada Network climate reporting protocol: Version 1.0. Natural Resource Report NPS/SIEN/NRR—2012/543. National Park Service, Fort Collins, Colorado.

Acknowledgements

Many thanks to John Gross, Stephen Gray, Mark Williams, Brent Frakes, Stacey Ostermann-Kelm, Cathie Jean, Chris Daly, Chris Knud-Hansen, Mark Otto, and anonymous reviewers for suggestions and comments. Thanks also to Michael Hartman and Todd Ackerman who contributed to Niwot Ridge climate analyses presented in some of the figures. This report was funded by the National Park Service, Rocky Mountains Cooperative Ecosystem Studies Unit (RM-CESU) in support of the Inventory and Monitoring Program. This publication is in addition a contribution of the Niwot Ridge Long-Term Ecological Research (LTER) project funded by the National Science Foundation.

1.0 Introduction

1.2 Objectives

Climate sets the stage for a landscape's species and ecosystems, albeit a moving stage. Temporal variability over a wide range of scales – from hourly events, year-to-year variation to decadal shifts and centennial trends – is a strong determinant of the state of populations and systems now and where they are headed. In addition, a site's climate variability tends to show strong correspondence with what is happening regionally and to dynamics at hemispheric scales. Consequently, understanding historical and current patterns in local climate and its links to regional to global processes is crucial for park monitoring, management, and research goals.

However, answering climate-related questions can be problematic. This can be because available climate datasets are often not ready 'off the shelf' for such applications without careful preprocessing or because appropriate tools for analyzing climate data are not well known or understood by users. Problems can also arise because nearby climate stations do not exist, or the necessary variables were not recorded for the period required. The goal of this report is to guide the development, analysis, and interpretation of climate data to address questions relating park resources and climate.

The objectives of this report are then:

(1) To layout a methodology for developing research-grade climate datasets appropriate for resource management science.

(2) To introduce the array of analysis techniques available for answering questions we often ask regarding climate dynamics and their interaction with landscape processes.

1.3 Approach

My approach is to identify common, key issues encountered when working with climate data and then suggest approaches and resources for pursuing specific solutions most likely to be useful to NPS personnel. The report is more guide than cookbook: for the specifics of any given process, I point to references to serve as an entry into the literature. In Figure 1, I lay out a workflow diagram that follows the overall components of this guide – the diagram portrays generic elements and decision points that are common to climate observation-based studies.

I start off with an outline for laying out a project's or program's data requirements as this will establish what problems are worth fretting over and what methods may or may not be appropriate (section §2.0, Figure 1a).[1] Second, I discuss methods for handling problems that typify station records (§3.0, Figure 1b). Next, I go through common analysis techniques and their interpretation (§4.0, Figure 1c). I wrap up with an overview of key considerations in the implementation of these processes (§5.0) – you may wish to review that synopsis early on to see the take-home messages.

For most situations, I draw on established statistical methods that are relatively common in environmental sciences and generally found in statistical software packages, and provide references where this is not so much the case. I cover caveats that go with these solutions and analyses. I give examples of some techniques in the figures, and make use of their captions to reveal detailed considerations needed to implement such methods.

[1] In the digital version of this report, cross-references to sections (§'s), tables, figures, and footnotes are hyperlinked. Websites linked to in the report were last accessed December 2009.

2.0 Establish Goals, Identify Requirements

Different research questions exploring climate processes have, of course, different data requirements (Figure 1a). Clearly establishing your research questions and then corresponding data requirements is key. Some requirements may be obvious, such as temporal and spatial specifications.[2] Others may be more subtle, such as what data problems need to be dealt with or can be ignored and which analytical techniques are appropriate. Creating a technically-fancy infilled daily dataset could be a lot of effort for little payoff if the goal is to look at centennial trends. On the other hand, such work can give greater confidence in evaluating trends in day-specific variables such as growing-season onset and termination.

2.1 Data requirements from hypotheses

Best practice starts with a clear *a priori* statement of hypotheses to be evaluated.[3] This is preferred over undertaking a large number of analyses in an attempt to find significant relationships. The downside of undirected numerous comparisons is there is a reasonable probably of one or more tests having statistically-significant results due to random chance alone.[4]

Your hypotheses will dictate the specific data needs of a research problem. Use their statement as a guide to identify:

- Variables of interest – what are the system drivers of consequence?
- Spatial scales of interest – single site vs. landscape or regional analysis, for example. (§2.2)
- Temporal scale – e.g., are daily event structure and the occurrence of extremes important, or just seasonal means? (§2.2)
- Record length and/or spatial density – to assure sufficient observations in time or space to reveal the patterns hypothesized.

2.2 Issues of time and space scale

Through this process, some forethought needs to be given to capture not only the multivariate but also multiscale nature of climatic controls on organisms and ecosystems. Selecting relevant temporal and spatial scales for analyzing meteorological data to reveal climate's impact can be guided by two principles:

Characteristic scales. Ecological processes tend to operate at characteristic temporal and spatial scales (Figure 2c, d; Delcourt et al. 1983, Urban et al. 1987). These initially prescribe climate analysis scales for given ecological processes. Note, however, there is not a 1:1 correspondence

---

[2] For example, spatial data specifications include spatial domain and station density or grid interval. Temporal specifications include record length requirements and timestep.

[3] See Schumm (1991: Chapter 2) for discussion of the scientific method in practice.

[4] Regarding such 'fishing expeditions' or 'data dredging' – If analyses use, for example, a probability threshold for Type I Errors[22] of $\alpha = 0.05$, then there is a theoretical likelihood that one out of 20 tests will give statistically-significant results due to random chance alone. If such multiple comparisons are part of a directed design, then this error can be controlled by setting an acceptable *experiment-wise error rate* and calculating a corresponding probability threshold for each comparison. This raises the bar for any comparison to be significant (i.e., lowers the probability threshold for each individual comparison). For more on multiple comparisons and experiment-wise error rate, see: Yandell (1997: §6.1) – available in part on Google Books (see References). Other multiple comparisons techniques are *familywise error rate* and *false discovery rate*. Another approach to avoid spurious results in any analysis is *cross-validation*, where analysis is re-tested on a random subset of data that were withheld from the original analysis.

between characteristic temporal/spatial scales for climatic and ecological processes (Figure 2b; see caption).

Scale interactions. While climatic and ecological processes have characteristic scales (Figure 2), factors controlling them operate across the range of scales. Climate forcing, for example, occurs through integration of finer effects and constraints of coarser ones.[5] A mismatch in temporal scales of aggregating station climate data (e.g., monthly values) can miss important relationships. Such is the case when the critical impact of weekly weather events on, for example, population processes is not captured by monthly data (e.g., Hallett et al. 2004; see also §4.4.2: *Process timescales differ, Mode 3*).

Spatially, a broader view of climate variability provides regional context to local climate forcing (§4.7) or, broader yet, an understanding of how hemispheric processes (e.g., El Niño) set the stage for local ecology year to year (§4.8). For the latter, Stenseth and Mysterud (2005) demonstrate the benefits of the span of spatiotemporal scales implicit in analyses linking hemispheric climate to local ecology (see §4.8.3: *Circulation indices broadly integrative*).

There are some common pitfalls in identifying suitable scales for aggregating climate data to reveal forcings on ecological systems. These include selecting the level of aggregation based on:

- What data are readily available – e.g., for temporal aggregation: monthly or annual means, when daily or hourly data may be more appropriate

- A poor understanding of ecological processes and their controls – either because a study is exploratory (‘a fishing expedition’) or established notions of how a system works are not well vetted

Such pitfalls are often difficult to avoid, but recognizing such limitations from the start can aid in hypothesis formulation, analysis choice, and interpretation. Timescale mismatching and aggregation issues are further discussed in the context of timeseries analysis, §4.4.2.

2.3 Data requirements of analytical methods

The next step is to identify what classes of analytical methods are needed to address your key questions and their data requirements. Are you interested in solely descriptive statistics (means, variances, frequency distributions; §4.1)? Do you as well aim to statistically test relationships either within the climate data or with other site variables (e.g., trend, correlation, and spatial coherence analyses; §4.3-4.8)? You may need to explore analysis options covered in these sections to identify corresponding requirements. Some general points on data requirements:

- Descriptive statistics require sufficient and unbiased sampling of the record to give representative results

- Statistical tests expect that certain assumptions about the data be met – such as having observations that are independent, normally (Gaussian) distributed, and identically distributed

[5] Scale linkage through integration of finer-scaled effects and constraints of courser ones reflects the hierarchal nature of biological and geophysical systems (Allen and Starr 1982, Holling 1992).

In regards to the latter bullet, spatial and temporal climate data are characteristically not independent and some variables, such as precipitation, not normally distributed. I talk about these issues as they come up, yet keep in mind that some statistical methods are more forgiving than others in their requirements.

In dataset development and data analysis sections (§3.0-4.0), I point out key requirements for many of the techniques I discuss.[6] Consult statistical references for specifics on these requirements, tests for these requirements, and possible workarounds (e.g., transformations). Some of this material can be found in on-line statistical texts and statistical software documentation (e.g., Helsel and Hirsch 2002, Schreuder et al. 2004, McDonald 2009, Garson 2009).[7] A highly-regarded reference for standard (parametric) tests (e.g., linear regression) is Sokal and Rohlf (1994) and for non-parametric tests (e.g., rank statistics) is Conover (1999), though there are many other useful texts along these lines.[8] Books focused on statistical methods in geophysical sciences include von Storch and Zwiers (2001),[9] Helsel and Hirsch (2002),[10] and Wilks (2006).[10] 'R' is a powerful statistical package which includes many of the techniques discussed – it is available free online.[11] On-line calculators for some tests include Kirkman (1996).[12]

2.4 Requirements guide what is possible with what is available

Together, data requirements dictated by your research hypotheses and by methods to test these will guide station selection and screening of the data (§3.0). For some research questions, available station data may not be up to the task. Alternatively, regional high-resolution gridded temporal climate data may do well for certain purposes (e.g., PRISM:[13] Daly et al. 2002, 2008, Di Luzio et al. 2008; Hijmans et al. 2005). Baron (2006) illustrates the use of a gridded dataset for point applications. Surrogate variables are also options for exploring climate processes of interest, such

---

[6] Of the techniques discussed in this report, linear regression is a one employed in many situations. I introduce it in the context of station change and missing data (§3.4.2-3.4.3, §3.6) and later in trend analysis and bivariate comparisons (§4.3.1, §4.4). If you use regression in other contexts, still refer to these sections (especially §3.4.3) regarding requirements, implementation, and pitfalls.

[7] General online resources include:

- Helsel and Hirsch (2002): http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html
- Schreuder et al. (2004): http://www.fs.fed.us/rm/pubs/rmrs_gtr126.html
- Lane (2007) – *HyperStat Online Statistics Textbook*: http://davidmlane.com/hyperstat/index.html
- McDonald (2009) – *Handbook of Biological Statistics*: http://udel.edu/~mcdonald/statintro.html
- Garson (2009) – *Statnotes:* http://www2.chass.ncsu.edu/garson/PA765/statnote.htm
- SPSS tutorial: http://www.stat.tamu.edu/spss.php

[8] Also:

- For linear regression, Draper and Smith (1998)
- Crawley (2002), a stats text with S-Plus applications
- Crawley (2007) is a reference manual for the R statistical language.[11] Also, see Verzani (2004) – available in part on Google Books (see References).

[9] von Storch and Zwiers (2001) available in part on Google Books (see References).

[10] Wilks (2006) available online on "Scribd:" http://www.scribd.com/doc/7128720/Statistical-Methods-in-the-Atmospheric-Sciences-Daniel-Wilks- [No longer available. 2/1/10]

[11] R site: http://www.r-project.org/. A USGS course on R has useful materials online: http://www.fort.usgs.gov/brdscience/LearnR.htm, including resources listed at: http://www.fort.usgs.gov/brdscience/LearnR.htm#References.

[12] Kirkman (1996): http://www.physics.csbsju.edu/stats/; see also http://www.physics.csbsju.edu/stats/Index.html. Also: Wessa (2009) – http://www.wessa.net/rwasp_spectrum.wasp. These and other on-line tools are presented as possible resources, not as an endorsement or reflecting an assessment.

[13] PRISM is a 'smart interpolation system' that uses location-specific relationships among elevation, aspect, basin configuration, etc. and minimum and maximum temperature and precipitation to develop spatial interpolation functions.

as surface hydrology to reflect watershed climates (e.g., Stohlgren et al. 1998) or station variable correlates (e.g., MTCLIM: Thornton et al. 2000, Running et al. 1987).[14]

3.0 Methods for Making Climate Records Useful

3.1 Types of problems

3.1.1 Station issues

Once climate data requirements are laid out and station datasets identified as appropriate to these needs, the next task is to check for and deal with problems in these stations' records (Figure 1b). Nearly all station records have such problems, unless they have been extensively processed by another party (e.g., NOAA U.S. Historical Climate Network dataset, USHCN: Easterling et al. 1996, 1999, Endoe 2009)[15] – issues associated with processed datasets are discussed in the next section (§3.1.2).

Problems that typify station records are:

- Data errors – resulting in physically implausible values and questionable outliers (§3.3.1, §3.3.2)
- Collection biases (§3.3.3)
- Station changes – changes in location and instruments, changes in station environs (§3.4)
- Record length – whether sufficient for detecting temporal patterns (§3.5)
- Missing observations (§3.6)

These issues and common solutions are laid out in the sections indicated. I also discuss important tasks to do throughout, namely, to track data changes (§3.2) and document and evaluate the effects of your processing (§3.7). As you see from this list of issues to work through, creating a credible dataset suitable to your research questions requires an integrated, multi-stepped quality control (QC) process – some stages may be automated, but ultimately the process entails hands-on decisions (Peterson et al. 1998b: §5). Decisions are facilitated by involving those who are familiar with the region's climate and with the workings and limitations of QC tests.

As you evaluate the results of your efforts to create a usable dataset, keep in mind that individual station data problems are sometimes severe enough – too many missing values, intractable station changes – that estimated (corrected, infilled) values overwhelm original information, and it becomes more prudent to drop a station and rely on a nearby record.

3.1.2 Processed dataset issues

In the case of processed 'cleaned-up' data, your job is not over quite so easily. Understand what methods were used to handle these problems and evaluate if these methods are consistent with your needs. Judge if their techniques altered the data in a way that obscures a key process of interest, such as infilling missing days in mountain stations based an adjustment of valley records when understanding elevation contrasts are your research goal. Processing centers regularly flag modified datapoints and may offer versions with original data and data after different stages of

[14] MTCLIM estimates solar radiation and relative humidity based on their physical relationships with minimum and maximum temperature and precipitation.

[15] Easterling et al. (1996, 1999) constitute USHCN Version 1: http://www.ncdc.noaa.gov/oa/climate/research/ushcn/ushcn.html. Endoe (2009) provides a summary of the current research version (Version 2): http://www.ncdc.noaa.gov/oa/climate/research/ushcn/. See also §3.4.6 and footnote 57.

processing so you can backtrack to the level of adjustment that matches your requirements (e.g., in USHCN Version 1).[16]

One class of processed climate data are area-averaged summaries. Area averages present their own issues. This is especially the case for climatically heterogeneous domains, such as U.S. Climate Divisions (Guttman and Quayle 1996, CLIMAS 2002).[17] Division dataseries are unweighted means of available stations at a given time. This means that available stations are not incorporated in proportion to how well they represent the domain's climate. Resulting problems include:

As stations come into and leave the averages, the spatial representation within a domain shifts through the record. This is especially a concern over topographically-diverse areas.

Averages of station records blend signals – the resulting signal's temporal variance is reduced and no longer realistically represents a region's climate. The degree of such signal blurring changes as stations come and go.

For these reasons, it is preferable to base analyses on individual station records. With these caveats in mind, however, areal averages can facilitate regional monitoring objectives (CLIMAS 2002, Wolter and Allured 2007). To reduce some issues with U.S. Climate Divisions, Wolter and Allured (2007) propose a revision with divisions based on climatic similarity.[18]

3.2 Version control and change flags

As you develop a workable dataset, document your process and provide means to undo your changes. Best practices dictate that:

- Adjusted, removed, or infilled observations be indicated with a flag
- Version control be exercised so that changes can be recovered, back to the original data if need be

Such tracking is crucial throughout the process of developing datasets and allows for evaluation of implemented changes. Include these elements in your documentation (§3.7).

3.3 Data errors, outliers, and biases

Data errors and biases come from observation collection, coding, and processing errors[19] and instrument and collection-protocol biases. These problems can be identified through screening for nonsense values (discussed next) and outliers (§3.3.2) and testing for known biases (§3.3.3).

3.3.1 Screening for reasonable values

Common quality checks for meteorological data are outlined in *NPS Climate Data and Monitoring Options* (Redmond et al. 2008); also refer to Peterson et al. (1998b: §2, App. A).[20] A good, generic guide to data checking is provided by Chatfield (1995: §6.4).[21] Screening can catch data transcription errors and common instrument errors such as dropouts, drift, biases, and other glitches

[16] Four levels are offered for USHCN v. 1 http://www.ncdc.noaa.gov/oa/climate/research/ushcn/ushcn.html#DATA

[17] U.S. Climate Divisions are laid out more along watershed, economic, and political boundaries than by climatic (and so general ecological) similarity. Interactive maps of current conditions by U.S. Climate Division are at: http://gis.ncdc.noaa.gov/website/ims-cdo/div/viewer.htm

[18] Wolter and Allured (2007): http://wwa.colorado.edu/IWCS/archive/IWCS_2007_Jun.pdf, and see http://www.esrl.noaa.gov/psd/people/klaus.wolter/ClimateDivisions/.

[19] Be aware that some processing errors may include previous attempts to correct errors and you may need to 'undo' these (Peterson et al. 1998b).

[20] For a detailed review of errors and correction techniques for hourly data from automated networks, see Wade (1987).

[21] Chatfield (1995) – available in part on Google Books (see References).

if of sufficient magnitude (Figure 3). Such quality checks include nonsense and plausibility checks, such as a day's minimum temperature greater than the maximum and values beyond physical limits (see Table 1 for an example layout). Other checks look for questionable outliers and unusual behavior, such as spikes, step changes, or flat lines (§3.3.2).

Some data screening is done by data archive centers; otherwise, such processes need to be implemented as part of your own data protocol. You may decide that additional custom filtering of the data should be employed to meet specific needs of a project (cf. Redmond et al. 2008).

3.3.2 Outliers

Checks regarding unusual behavior can be among the most challenging to design. This is because errors can be hard to discriminate from real extremes and other meteorological dynamics. Any scheme to identify outliers as 'bad data' runs the risk of rejecting real events that have important ecological or hydrological consequences. Formally, we discuss this as a Type I error, that of rejecting good data.[22] Generally, climatologists are adverse to this error in favor risking a Type II error (accepting bad data) and so of retaining all outliers (provided they lie within physically-plausible limits). For your application, weigh the relative consequences of these errors to your analysis to decide how lax or aggressive to be in filtering out extreme values. NOAA regional climate centers, for example, differ in balancing these errors depending on their geography and prevailing user needs.[23] Outlier effects on regression analysis are discussed in §3.4.3.1.

Detection and rejection decisions for outliers or other unusual behavior are usually applied to daily or hourly data. These rely on a variety of methods that check for:

- Multivariate physical consistency – similar dynamics in physically-related parameters (e.g., among minimum and maximum temperatures, cloud cover, relative humidity, and precipitation)
- Climatological consistency – staying within the range of climatological norms, based on longterm record statistics
- Temporal consistency – rate of change tests looking for spikes and step changes
- Spatial consistency – similar dynamics in records of nearby stations

Detection processes can be automated – however, manual inspection is advised before flagged outliers are rejected. Meek and Hatfield (1994) present basic methods for (1) physical limits, (2) climatological filters, and (3) rate of change tests for daily and hourly records for an array of climate variables.

Multivariate physical consistency. For multivariate checks, Gandin (1988) gives examples of numerical tests and Redmond et al. (2008) of logic-based ones.

---

[22] More formally, a Type I error is the error of rejecting a null hypothesis (H_0) when it is *true* and should not be rejected. In this context, H_0 is that an outlier is from the same population as other observations (i.e., 'good data'). A Type II error is not rejecting a H_0 when it is *false* and should be rejected.

[23] For example, the Western Regional Climate Center (WRCC; Table 2) leans toward retaining all station data, including outliers which may or may not be true extremes. The High Plains Center (HPRCC), on the other hand, tends to eliminate likely outliers, with some possibility of tossing true extremes. These practices relate to common uses of the data in the respective regions – HPRCC data uses include crop models, which may be highly sensitive to daily extremes and so thrown off by outliers. WRCC's philosophy is that end-users can best determine criteria for handling outliers based on their specific application. Where these centers' regions overlap, data for the same station can differ depending on which center's dataset is accessed. (Stephen Gray, personal communications)

Climatological limits. Meek and Hatfield (1994) present a variety of climatological tests generally based on absolute ranges from the record (Figure 4c). This creates a rather strict test. More common climatological filters use some multiple of the standard deviation (SD) from the longterm mean to set limits. This multiple:

- Is usually large (e.g., >3 SD) – so more liberal than record limits
- Depends on climate variable – e.g., 5 SD for temperature (Figure 4d). For non-normally distributed data, such as precipitation, a multiple of the coefficient of variation (CV = SD/mean) is can be used; alternatively, data can first be normalized with an appropriate transform (§3.4.3.1) and then SD limits determined and applied in transformed space.
- May be asymmetrical – greater limits on one side of the mean than the other

Climatological limits also typically vary by season (Figure 4c, d) and can depend on site characteristics (for example, whether a site is maritime vs. continental, at high vs. low elevation, or in an air-shed prone to thermal inversion or cold-air drainage).

Peterson et al. (1998b: §4) review a variety of SD-based methods for detecting temperature outliers. SD thresholds may be staged – with a moderate limit (e.g., 2.5 SD) calling for additional evaluation, and a higher limit (e.g., 5 SD) outright rejection (Figure 4d). Box-and-whisker plots, described in §4.6.1, provide another frequency distribution-based method for identifying outliers. Longterm trends should be removed before implementing distribution-based techniques as trends can substantially contribute to a record's variance – it's more appropriate to judge outliers against a detrended record variance or vs. the local (shortterm) variance (Peterson et al. 1998b).

Temporal consistency. Some temporal inconsistencies, such as spikes and step changes, can be detected visually (Figure 4a) or numerically with rate of change checks (Meek and Hatfield 1994) (Figure 4b). The opposite case, too much temporal consistency – a flat line – suggests data dropout due to observer, instrument, or coding error.[24] Outliers and unusual behavior found in temporal tests can be further scrutinized with spatial consistency checks relying on neighbor comparisons (e.g., Eischeid et al. 1995; and Figure 4d caption). Used in tandem, these consistency tests bring both temporal and spatial perspectives to bear on judging outliers.[25]

Spatial consistency. In spatial checks, a primary consideration is how far away a neighboring station can be and still be relied on. This *spatial correlation length scale* depends on:

- Climate variable – spatial correlation generally drops off less rapidly for temperature than precipitation, especially for regions and seasons prone to convective rainfall (with a patchy distribution on hourly and daily bases)

[24] Data flatlines can be identified as a series of observations with no or little change in value. These show up as an outlier in plots of short-period running standard deviations (like a running mean). Dropped observations should be marked as missing data. As an example test, implemented for PRISM input quality control for daily maximum and minimum temperature, a flatline period is identified as:

- ≥ 2 consecutive days with observations of 0.0ºC
- ≥ 10 consecutive days with observations within ± 0.4ºC.

5-10 consecutive days with observations within ± 0.4ºC are seen as a potential flatline period. These periods are statistically evaluated based on whether the 5-day moving standard deviation of the target station differs from that of surrounding stations; if differs and lower, then it is a flatline period. (Chris Daly, personal communication)

[25] A method that uses spatial comparisons to detect temporal inconsistencies is *double mass analysis*, discussed in §3.4.6 in the context of detecting station changes.

- Time aggregation interval – spatial correlation generally extends farther for longer time averages (hourly vs. daily vs. monthly)
- Terrain uniformity – spatial length scale diminishes more quickly over heterogeneous terrain (e.g., regions that are mountainous, with significant landcover changes, or with varying degrees of lake/maritime influence)

Spatial tests are more readily applied in regions under a relatively uniform climate regime, such as the Great Plains, than one with sharp physiographic contrasts (e.g., the Rocky Mountains) or extremely sparse precipitation (e.g., the Southwest). These factors are discussed further in §3.6.5. Hubbard et al. (2007) and You et al. (2007) evaluate spatial methods for detection of outliers in daily surface temperature and precipitation data, respectively; Eischeid et al. (1995) implemented a multiple-test spatial approach.

3.3.3 Collection biases

Data collection biases include those from:

Instrument calibration and design. Poor or drifting calibration issues can be accounted for if a close-by instrument can be used to develop a correction function or if a later re-calibration is well documented. Well understood biases associated with specific instrument changes are discussed in §3.4.5.

Time of observation for minimum and maximum temperature. For stations using minimum and maximum thermometers, time of observation other than midnight but rather at the more usual early morning or late afternoon times creates a bias in monthly mean temperatures. The bias can exceed 2ºC (Karl et al. 1986; Figure 5). This inconsistency arises largely from the time thermometers are reset relative to passage of warm and cold fronts.[26] This timing leads to cold biases in monthly means for morning reading stations and warm biases in means for late afternoon reading stations (Vose et al. 2003).

This is particularly troublesome for evaluating trends if observation time has changed during a station's history (see Figure 5 caption). A correction for this bias is presented by Karl et al. (1986) and further evaluated by Vose et al. (2003).

Missing observations that are time dependent (e.g., by weather, season, decade). Dataseries with dropped days that are pervasive, for example, at the start of a record or for periods of inclement weather will incur undersampling biases with time (problematic for trend analysis) or for bypassed synoptic conditions (problematic for daily event and seasonal analyses). Keep in mind the possibility that such temporal biases will interfere with questions being evaluated with these data. In general, these biases are not formally accounted for in studies – though they should be.

While missing days can potentially be infilled, the bias will persist in terms of lower variances: infilling estimation reduces variance in timeseries (see §3.4.3.3: *Reduction in variance*). Judging temporal biases with respect to infilling missing observations is discussed further in §3.6.2.

[26] For stations with a morning observation time (e.g., 0700h), the passage of a warm front soon after thermometers are reset can result in the *minimum thermometer* retaining that 0700h temperature through the next 24 hours rather than capturing the next morning's pre-dawn temperature, the diurnal-cycle's actual minimum. For stations with late afternoon observations (e.g., 1700h), the passage of a cold front after thermometers are reset can result in the *maximum thermometer* holding that 1700h temperature through the next day, rather than catching that day's post-cold front max.

Multiday-accumulated totals for precipitation. Daily precipitation series suffer from the accumulation of precipitation over several days due to:

- Observer absence for manual precipitation gauges, resulting in several zero days followed by an accumulated value. These are typically flagged as an accumulated total.
- Sub-threshold input for tipping bucket gauges – If minor precipitation during a day doesn't accumulate enough to tip the collector but does with more inputs in the next day(s), then correct precipitation amounts will not be recorded on corresponding days. The only indication of an error will be if a nearby observer-read gauge shows distributed precipitation events during the same period.

When the data goal is monthly totals, multiday totals can be summed in. That is, unless the accumulation period crosses a month boundary. Then totals must be parsed. An accumulated total is often parsed out by day by apportioning the total guided by day-to-day amounts observed at a nearby station clear of these issues. This approach is likely to meet with success if the stations have good correspondence in daily structure (e.g., with tests in §4.6.1) and are likely to experience the same precipitation events (see §3.3.2: *Spatial consistency*, and §3.6.5: *Spatial heterogeneity*, regarding spatial correlation length scale). Such corrected data should be omitted in daily analyses (§4.6). If the total cannot be parsed in a defensible manner, then the total is dropped and all days included in the total are counted towards the months' missing day tolerances (§3.6.1).

An overview of manual station collection issues, such as for NWS cooperative network stations,[27] is presented by Leffler and Redmond (2004);[28] Daly et al. (2007) look at observer precipitation biases in detail.

3.4 Accounting for station changes

Station histories are rarely boring and tell of changes that give little consistency to observations. Troublesome station changes can as varied as:

- Station relocations to vastly different sites (even if nearby)
- New instruments with different response curves and accuracy (with no overlapping record)
- Dramatic adjacent landuse change – such as the introduction of irrigation, pavement, or an overshadowing building

Such changes introduce sharp, discontinuities into the record (Figure 6a). These changes are for the most part artificial, that is, 'non-climatic change points.' Temporal inhomogeneities arise as well from gradual changes, such as instrument drift (§3.3.3) or slowly changing conditions in the vicinity of stations (§3.4.7): for example, as landscaping or natural vegetation grows up and closer, and rural areas become urbanized. Peterson et al. (1998a) review many of these issues and numerical techniques to detect and handle them. Correction of temporal inconsistencies is important – ignoring them can have a major impact on climate change assessment (Figure 7). If, however, the intended use is for climate regime change detection, some detection and correction strategies are best avoided (see cautions in §4.5.3).

[27] National Weather Service (NWS) Cooperative Observer Program (COOP): http://www.weather.gov/om/coop/what-is-coop.html. Also referred as 'co-op' stations in this document.
[28] Online tutorial version: http://www.weather.gov/om/csd/pds/PCU6/IC6_2/tutorial1/Factors.htm

In this section, I discuss developing station histories as the first step toward accounting for station changes (§3.4.1) and then present several approaches for inhomogeneity correction depending on if:

- The station has overlapping records spanning the change point (§3.4.2)
- There is no overlap nor appropriate nearby station to guide adjustment (§3.4.4)
- The change is due to well-documented instrument issues (§3.4.5)
- There are reliable, nearby stations to guide adjustment – including cases where change points may be present but not sufficiently known from the station history (§3.4.6)
- Changes are gradual due to environmental changes (§3.4.7)

Section 3.4.3 gives a primer on regression – re implementation, interpretation, and cautions – as many of these approaches rely on this statistical method.

3.4.1 Building station histories

We have more confidence in numerical approaches for detecting and correcting temporal inhomogeneities if they are directed by knowledge of a station's history. Sussing out station changes can be an arduous process, however, and is one usually undertaken only if:

(1) Only a few stations are targeted.
(2) Their records are of high value with respect to a programmatic goal – such as for an important location.

– and –

(3) They bear a high probability of forming a high-quality record – that is, there's enough original information to form a strong foundation for developing a credible correction protocol.

To build a history for critical stations[29] –

1. Consult station documentation to identify potential sources of temporal inhomogeneities, recognizing that not all changes get into such metadata.
   a. For NWS co-op stations,[30] *Station History Reports* (B-44 forms) are available from NOAA regional climate centers and state climatologists (Table 2).[30] These reports describe station moves, instrument changes and major maintenance, and changes to the surrounding environment – though the reports are not necessarily complete.

2. Review repeat photography, which can provide additional insights especially for gradually changing environs.

3. Visit the site and its operator before undertaking a lengthy data protocol. This can provide an understanding not gleaned from station history reports and save effort in the long run. Intuition can be gained from:
   a. Seeing station siting, instrument condition, and obvious recent environmental changes
   b. Talking with operators, who very often have a longterm association with a station and can relate details that do not get into formal paperwork

If a location or instrument change is designed with an overlap between the old and new records, adjustments can be relatively straightforward, using, for example, linear regression (§3.4.2). Without overlapping records, a common method is mean offset and ratio adjustment (§3.4.4).

[29] For a guide to problematic station siting and observation practices to look out for, refer to Peterson et al. (1998), Leffler and Redmond (2004),[28] and Daly et al. (2007).

[30] Metadata describing station histories for NWS Cooperative and other U.S. stations are available from NCDC (http://www.ncdc.noaa.gov/oa/metadata/metadataresources.html) (see also State Climate Offices, Table 2).

While introduced here for simple cases, these approaches form the basic elements of more complex record processing for well-documented changes and other inhomogeneities (§3.4.5–3.4.7), merging sequential records of nearby stations (§3.5), and infilling missing observations (§3.6).

3.4.2 Correcting known change points. I: The simple case with overlapping records – Regression

When there is sufficient overlap between outgoing (*x*) and new (*y*) station records, we can develop a linear regression conversion equation:

$$y'(t) = b_0 + b_1\, x(t) \quad , \qquad \textbf{(1)}$$

where a value from the outgoing record $x(t)$, x at time (t), is adjusted to a value $y'(t)$ consistent with the new record.[31] The mark (′) indicates that the value is estimated; b_0 and b_1 are the regression *y*-intercept and slope, respectively, determined from the observed (x,y). An example correction based on overlapping instrument records is shown in Figure 7 for a precipitation gauge change – the correction had a substantial impact on a longterm trend analysis.

Details in implementing regression are covered next (§3.4.3).

3.4.3 Regression analysis primer – Implementation, interpretation, caveats

While I'm presenting this primer on regression embedded in the section on correcting station changes, these guidelines apply to using regression with climate series in general. I'll refer back to this section for other applications later in the report.[32]

Regression analysis implementation details –

I discuss in this section considerations in implementing regression analyses – these are with respect to:

Timescale considerations
Numerical method
y-intercept method

and the means to deal with:

Autocorrelation
Outliers
Data distribution requirements
Nonlinearity

Time-related considerations: Record overlap, timestep, and season dependence. There will be more confidence in developed conversion equations:

- For cases with longer overlap periods
- If the conversion timestep is at the original collection temporal resolution – e.g., daily rather than on temporally aggregated values. Conversions at daily or weekly rather than monthly or longer time intervals are better able to capture nonlinear relationships. On the

[31] It doesn't matter which of the early or current record is adjusted – in the example implementations here and §3.4.4, the earlier series is adjusted to match the current record.
[32] Other sections that deal with regression in some detail are §3.6 (*Missing observations*), §4.3 (*Trend analysis*), and §4.4 (*Covariation among variables*)

other hand, there is little expectation that hourly data are correlated between sites except at the closest distances.[33]

- If conversion relationships are season or weather-regime dependent. Conversions that account for seasonal shifts in overall weather regime better reflect relationships between locations. One technique is to base analyses on a moving window (in time, e.g., spanning 28 days) to avoid sharp temporal boundaries in conversions.[34]

Numerical implementation. Overlapping observed records (x, y) are used to calculate b_0 and b_1 in eq (1) using best-fit techniques:

- *Ordinary least squares* (OLS) is the most common method for determining the equation that best fits the data – and is the primary technique provided in most statistical packages and presented in statistics books (cf. §2.3).[35] OLS assumes that observations are *independent*, *normally distributed*, and *variance constant* (homoscedastic).[36]
- *Robust regression* methods, while less powerful than ordinary least squares, are alternate techniques that are 'robust' with respect to the limitations of non-normal or heteroscedastic data, as well as 'resistant' to the effect of outliers (discussed later in this section).[37]

Ordinary least squares regression is the method of choice (in terms of testing power) when data distribution assumptions are reasonably met directly or with data transformations (discussed shortly), while robust regression techniques are more general and can be applied with fewer restrictions (Wilks 2006). In the heteroscedastic (but still normal) case, an appropriate robust regression method is *weighted least squares* (WLS; Helsel and Hirsch 2002: §10.3).[38] In WLS, observations (y) are weighted by the inverse-square root of their local variance[39] – this accounts for variance changes along the range of y.

Zero intercept. Depending on the nature of the relationship between two records (x,y), you may force the y-intercept b_0 to zero in setting up the regression analysis. While this may make sense logically, in practice it generally does not produce the best results. This is because, when not set to zero, b_0 is the regression line's y-offset determined across the full range of data and so is not so much about what's happening at $x = 0$.([40])

---

[33] Recall that station moves are not necessarily over short distances (§3.4).
[34] In this approach, the conversion for any given day of the year is based on all daily data in the record that fall within, for example, a 28-day window centered on the day.
[35] von Storch and Zwiers (2001: §8.3.15-18)[9] review pros and limitations of the least squares method.
[36] See Helsel and Hirsch (2002: §4.4)[7] re tests to evaluate whether data are normal. See also Steinskog et al. (2007). Residual plotting techniques for evaluating heteroscedasticity are discussed in §3.4.3.2.
[37] Re: robust methods –

- Wilks (2006: §3.1.1)[10] presents a general discussion on robustness and resistance.
- Helsel and Hirsch (2002)[7] cover robust and resistant methods throughout their book; robust regression techniques are treated in Chapter 10.
- Software packages – robust techniques are offered in R and SPSS, for example.

[38] Helsel and Hirsch (2002: §10.3)[7] present a technique for performing WLS using OLS linear regression software.
[39] *Local* variance refers to variance of y determined for an interval along the range of y.
[40] If a regression line substantially misses the mark for observations near $x = 0$, a nonlinear transformation may be in order (covered later in this section: *Transformations for modeling nonlinear relationships*).

Serial correlation. Climate data are typically serially correlated – that is, not independent. I discuss tests and corrections for serial correlation in §4.3.1 in the context of trend regression analysis – these are applicable here as well.

Outliers and end members. Outliers can have an undue influence on a regression analysis, especially for ordinary least squares regression. Outliers near the end of a regression line (end members) have significant leverage in determining the slope (b_1) of the line.[41] To check for these issues:

(1) Graph *y* vs. *x* (scatter plot) to reveal suspect points
(2) Check that outliers are not errors (§3.3.2)
(3) Test the sensitivity of the regression to outliers by re-running the analysis omitting them
(4) If the regression is sensitive, apply a robust technique[40]

Transformations – to adjust data distribution. Transformations may be used to adjust data to normality (or at least to symmetry) and constant variance.[42] Precipitation data, for example, are rarely normally distributed, as are variables reported in relative proportion or percentage units (relative humidity, % sunshine). Common precipitation transformations include logarithmic, square root, and cubic root.[43] A transformation of proportional data is arcsin[square root(*x*)], where $0.0 \leq x \leq 1.0$ (McDonald 2009).[44] Wilks (2006: §3.4.1)[13] lays out a process for determining the most appropriate power transformation for a given dataset.[45] For issues arising from transformations in regression analysis, see §3.4.3.3.

Transformations – for modeling nonlinear relationships. If you have reason to believe that the *x, y* relationship is nonlinear, examine the data initially with a *y* vs. *x* scatter plot. Based on the shape of the plot, explore straightforward nonlinear transformations of either or both *x* and *y*.[45,46] Re-graph to see if the conversion had the desired effect. The resulting 'linearized' (*x*, *y*) is then used in the linear equation (eq 1). As just noted, for concerns re transformations in regression, see §3.4.3.3.

Judging results –

To decide if the regression equation you've developed is sufficient to the task:

(1) Check the regression significance level (*p* value)[47] – Reject the result if *p* is not within an acceptable level (e.g., $p < 0.01$, or at least < 0.05).

---

[41] See Helsel and Hirsch (2002: §9.5.1)[7] re: outliers.
[42] Resources re transformations:
- Helsel and Hirsch (2002: §1.7, §9.3, §9.6)[7]
- Wilks (2006: §3.4.1) for more specifics[10]
- von Storch and Zwiers (2001: §8.6.2)[9] present complex functions to resolve common issues.
- McDonald (2009: pp. 160-164) = http://udel.edu/~mcdonald/stattransform.html.

[43] Examples – (a) to apply a square-root transformation in eq (1), both *x*(t) and *y'*(t) would first be converted to SQRT (precipitation). (b) see eq (5) in §3.4.4.
[44] For percentage data, first convert *x* to a proportion: *x*/100.
[45] Power transformations, as some of those just noted for precipitation, include power of *x* (e.g., square-root $x^{1/2}$, x^2, inverse x^{-1}), exponential (e.g., e^x, e^{-2x}), and logarithmic [$ln(x)$].
[46] Many statistical and spreadsheet packages have nonlinear line fitting routines that can aid in exploring possible *x*-transformations, such 'trendlines' in (Microsoft Office) Excel (for trendlines options, select: display equation, display R-squared).
[47] The *p*-value is the probability that the relationship found is due to chance alone, rather than to a hypothesized relationship. More formally, this is the probability of committing a Type I Error (see footnote 22). von Storch and Zwiers (2001: §4.1.7, 4.1.9-.11)[9] review the interpretation of statistical tests.

(2) Judge the predictive power of the regression – The regression coefficient (= coefficient of determination), R^2, can be interpreted as the proportion of the dependent variable's (y's) variance that is explained by the regression equation. Ask if it is high enough to be useful (e.g., R^2 >0.60 = 60% variance explained). *N.B.*, it is not sufficient to have high significance (a high p), as this does not necessarily mean that a large amount of the observations' variance is explained.[48]

(3) Visually check a scatter plot of the regression's residuals (y'-y) vs. x – where y' and y are the predicted and observed value of the dependent variable in eq (1). A residual plot can reveal problems in the data or regression model used. If the residual plot shows a pattern other than a random normal, homoscedastic dispersion from a zero-slope line, then it is likely that transforms to adjust for data distribution issues or nonlinearities are needed or that those already applied need to be reassessed.[49] Wilks (2006: §6.2.6)[13] lays out an approach for diagnosing residual plots.

If the significance test's p, R^2, or residual plot are not satisfactory, review implementation details (§3.4.3.1) and pitfalls (discussed next, §3.4.3.3); see also Helsel and Hirsch (2002: §9.5)[10] re regression diagnostics. Other caveats in interpreting regression results are given in §4.4.2 (for the case of regressing timeseries of two different variables)[50] and §4.9 for statistical results in general.

Pitfalls in linear regression results –

Before relying on regression results, take into consideration two possible analysis artifacts: (1) variance reduction arising from statistical estimation and (2) inconsistencies arising from transformations.

Reduction in variance. A problem stemming from using regression-estimated values in a dataset is that these estimates reduce the variance of a timeseries. This is because the conversion equation delivers the best estimate, but without accounting for randomness associated with observation. Such reduced variance will cause standard statistical tests you implement in later analyses to overstate the significance of your results. If a fair portion of the data is estimated (from this and other preprocessing steps), keep in mind the consequences of this artifact when interpreting your final results. More sophisticated regression methods include an 'error' term in equation (1), which stochastically adds back in such noise to predicted values.

Transformations and R^2. In regression, transforming the *dependent* variable y – whether to adjust data distributions or to linearize the model – complicates the interpretation of R^2. This is because the transformation alters the variance structure of the underlying data, such that R^2 becomes the variance explained relative to the *variance of the transformed y*. As a result, R^2's for regressions with linear and different transformations of y are not on comparable scales (Scott and Wild 1991) – that is, you cannot judge whether a transformation produced a better regression equation based on R^2's.[51]

---

[48] The interpretation of p vs. R^2 is also discussed in the context of teleconnections (§4.8.3: *Prediction*).
[49] Helsel and Hirsch (2002: §2.3.3)[7] note difficulties in visually judging residual plots for heteroscedasticity, and present a smoothing method to avoid these problems.
[50] under: *Interpretation of correlation and regression results* in §4.4.2
[51] See also Helsel and Hirsch (2002: §9.6).[7] This is an issue only when the dependent variable y is transformed, not if just independent x's are transformed.

Kvålseth (1985) and Scott and Wild (1991) suggest that regressions with transformed *y* are best evaluated on how well predicted *y′* match observed *y* back in their original scale, rather than by the regression R^2 in transformed space. This is done by back-transforming predicted *y′* and observed *y*, and regressing (*y′*, *y*) pairs now in linear space to determine the appropriate R^2 – one that relates to the original data and is comparable across models. The resulting overall strategy is then: the best-fit *regression equation* (and its statistical significance, *p*) is determined in *y*-transform space, but the R^2 used to evaluate % variance explained is calculated in the original (linear) space. Willett and Singer (1988) discuss this issue with respect to weighted least squares robust regression (WLS, §3.4.3.1).

For other scale disconnect issues regarding back transformation of data distribution parameters (e.g., mean, SD), see §4.6.1.

3.4.4 Known change points. II: The simple, single station case, without overlapping records – Offset and amplitude adjustment

The regression method is obviously not possible if there are no overlapping records associated with a documented station move or instrument change. If, on the other hand, there is a relatively clean step change in a climate variable with the station change (e.g., Figure 8), then an easy method is to adjust one part of the record or the other by such a step offset.[34]

Mean offset adjustment. For the case of adjusting the early record (*x*) to be consistent with the more recent record (*y*), the offset is simply based on the difference in the record's *shortterm means* on either side of the change point, $\Delta_{y\text{-}x} = (\bar{y} - \xi)$. Each observation in the early record, *x*(t), is then shifted by the offset:

$$x'(\mathrm{t}) = x(\mathrm{t}) + \Delta_{y\text{-}x} \ , \qquad \textbf{(2)}$$

where *x′* (t) is the adjusted value of the early record consistent with the more recent record *y*. The means ($\bar{y}$, ξ) should be for a period long enough to represent the two parts of the record, but not so long that their difference is affected by decadal shifts in climate regime or longterm trends.

Variance adjustment. If the variance of the record also changes at the change point, then the early record (to continue the case just used) can be *amplitude* adjusted by the ratio of the standard deviations of the two records. Amplitude adjustment, however, is applied to *deviations* of the early record from its mean, $\delta\, x(\mathrm{t}) = x(\mathrm{t}) - \xi$. This creates an anomaly timeseries.

The amplitude of the anomalies is then adjusted by the ratio of standard deviations after vs. before the change point:

$$\delta\, x'(\mathrm{t}) = \delta\, x(\mathrm{t}) \times (\mathrm{SD}_y \,/\, \mathrm{SD}_x) \qquad \textbf{3)a}$$

$$x'(\mathrm{t}) = \delta\, x'(\mathrm{t}) + \bar{y} \qquad \textbf{b}$$

In eq (3a), $\delta\, x'$ (t) is the adjusted anomaly series and SD_x and SD_y are standard deviations over the same shortterm periods used to calculate ξ and $\bar{y}$. The corrected timeseries is constructed from anomalies by adding back in the mean: for the adjusted early record (*x′*) to be fully consistent with the recent period, we add in the mean for the recent period, $\bar{y}$ (eq 3b). The process of taking out ξ at the start and then adding $\bar{y}$ at the end incorporates the offset adjustment in eq (2).

Considerations for specific variables –

Minimum and maximum temperature. Offset adjustment is often used for temperature series. Where both minimum and maximum temperatures (T_{min}, T_{max}) are being adjusted, the physical requirement that $T_{min} < T_{max}$ can end up being violated when eq (3) is independently applied to these linked variables. A common approach for avoiding this issue is to recombine them into the generally independent variables mean temperature (T_{mean}) and diurnal temperature range ($DTR = T_{max} – T_{min}$). We then apply record adjustment techniques to the derived variables, and subsequently restore the adjusted timeseries to T_{min} and T_{max}.[52]

Precipitation. Because the frequency distribution of precipitation is highly skewed, a nonlinear adjustment is more appropriate. A simple technique is scaling the earlier record x(t) by the ratio of the means:

$$x'(\mathrm{t}) = x(\mathrm{t}) \times (\bar{\mathrm{y}}/\xi) \tag{4}$$

In a more generalized approach, we use a nonlinear transform to try to make the precipitation frequency distribution function roughly symmetric about the median. Select the transform – such as a log, square root, or cubic root transform – that best suits the data (§3.4.3.1: *Transformations*). Using a natural-log transformation, eq (2) would be:[53]

$$ln[x'(\mathrm{t})] = ln[x(\mathrm{t})] + [ln(\bar{\mathrm{y}}) – ln(\xi)] \tag{5}$$

3.4.5 Known change points. III: Well-documented instrument changes
Some instrument changes are so common they have been well studied and correction practices recommended in the literature. Precipitation gauges vary in catch dynamics and their biases, such as for wind-driven undercatch of snow, are well documented. Corrections methods are described by Legates (1995) and Yang et al. (2005).[54] For NWS co-op station temperatures, a change over in the mid-1980's from liquid-in-glass thermometers to the electronic sensor Maximum/Minimum Temperature System (MMTS) introduced a bias in these records (Figure 8). The bias and its adjustment are examined by Quayle et al. (1991; see also Doesken 2005). Biases due to change in time of observation at manual sites are also well documented – this bias and its correction are discussed in §3.3.3.

3.4.6 Seeking and correcting undocumented inhomogeneities – Role of neighboring stations
Most station changes, however, are not accompanied by overlapping records nor are well documented. In this case, the common method relies on examination of records of nearby stations to detect and correct timeseries discontinuities in the station of interest (Karl and Williams 1987, Peterson et al. 1998a).[55] Such techniques can also be used to guide and confirm correction of known heterogeneities (§3.4.2–3.4.5).

The general procedure is:

---

[52] Adjusted records of T_{min} and T_{max} are reconstructed from T_{mean} and DTR as: $T_{min} = T_{mean} – ½\,DTR$, and $T_{max} = T_{mean} + ½\,DTR$. Note that, alternatively, DTR and T_{mean} can be analyzed in place of T_{min} and T_{max} (§4.2.1).

[53] Note that eq (5) can be re-expressed as: $ln\ [\ x'(\mathrm{t}) = x(\mathrm{t}) * (\bar{\mathrm{y}}/\xi)\ ]$, which reveals the relationship in eq (4).

[54] Photographs showing the variety of precipitation gauge types: http://www.uaf.edu/water/faculty/yang/bcp/photos.htm. Gauge biases are illustrated in online images from Yang et al.: e.g., Northern Hemisphere January measured and corrected precipitation: http://www.uaf.edu/water/faculty/yang/bcp/pm_jan.gif and http://www.uaf.edu/water/faculty/yang/bcp/pc_jan.gif, respectively.

[55] Another approach uses breakpoint analysis (Haimberger 2007, Christy et al. 2009). Another application of breakpoint analysis is discussed later re regime shifts (§4.5.3).

(1) Locate potential record shifts using station histories as well as possible (§3.4.1)
(2) Detect and evaluate potential inhomogeneities by comparison to a climate reference timeseries
(3) Adjust the station record if needed, guided by the reference series

A *reference series* can be based on a single neighboring station or a network of stations strongly correlated to the station being evaluated. For such corrections to be successful, a key criterion for a reference series is that it is 'clean' (temporally homogeneous) so that changes in its record only reflect the climate. As this is rarely the case, Peterson and Easterling (1994) give a method for optimizing the creation of reference series. Because metadata are often incomplete, Menne and Williams (2005) present more sophisticated methods that identify inhomogeneities in records without station history guidance. Enloe (2009) presents an overview of these protocols as used in USHCN Version 2.[18]

Temporal inconsistencies at a station relative to neighboring stations can also be identified and corrected with *double mass analysis.* This technique plots accumulated station values over time against the accumulation for a reliable neighboring station or the average accumulation across a set of stations. The line should be straight with little variation if records are consistent; any breaks in slope point to a systematic change in a station's collection regime (Kohler 1949).[56]

3.4.7 Station environment changes

Once non-climatic change points have been accounted for in a dataseries, some local human-induced gradual changes in climate may still remain, such as from changes in environs. For example, an urban heat island effect is found in timeseries for stations once in rural environs and now urbanized. Accounting for such warming effects is explored by Karl et al. (1988; Figure 9).[57] The climatic effects of other land use change can also be imbedded in the records of stations near land converted from, for example, forest to pasture and then to irrigated crops. Such changes can locally affect mean temperature, diurnal temperature range, atmospheric moisture, and precipitation – Hale et al. (2006) explores some of these impacts in station records.

If these changes might hamper detecting the signal you're interested in, it is important to remove or at least document these effects. However, keep in mind that these gradual climate changes are real even if they only reflect highly localized effects and not regional dynamics. Removal is warranted only if they interfere with addressing your research questions.

3.5 Concatenating timeseries of nearby stations to extend the record

To extend a site's climate history farther back in time than its current record, we can concatenate (link) its record with that of an older, similarly behaving station in the region (Figure 6). Similarly, if a station of great value (e.g., because of location and longevity) has been dropped from the observation network, we can substitute in an ongoing record from a nearby station. Concatenation can be completed much in the same way as we handle known station discontinuities (§3.4.2-3.4.4).

Generally, first evaluate quality and comparability of the two records. This entails:

---

[56] NWS/NOAA user's guide for a Unix program for double mass analysis provides an overview of the technique: http://www.nws.noaa.gov/oh/hrl/idma/html/dma_home_frame.htm

[57] Urban heat island effects are by their nature automatically treated in the scheme of Menne and Williams (2005), referred to in the previous section (§3.4.6); see Endoe (2009).[15]

- Ensuring that both records are temporal homogeneous (§3.4)
- Testing for a strong correlation between stations over a reasonable period of overlap – If the overlap is sufficiently long, similarity in station behavior at multiyear and longer timescales can be evaluated using spectral analysis (see Figure 6 caption and §4.5.2).

Second, concatenate the two series by:

- Adjusting one of the two series to conform to the other (§3.4.2-3.4.4)
- Choosing a point to switch from series to the other – This is often broadly based on judgment as to which original series is of higher quality and most important to preserve and fine tuned so as not to create a local discontinuity at the switch-over point.

3.6 Missing observations

When data are missing, three options are:

(1) Decide its ok to ignore in aggregations (§3.6.1)
(2) In the analysis phase, choose a statistical technique that allows for missing values (such as the Mann-Kendall trend test, §4.3.2)
(3) Infill them (§3.6.2)

3.6.1 Aggregation and missing day tolerances

If you have daily data and will be analyzing monthly values, missing a few days per month can be acceptable. An approach is to allow some number of days of missing precipitation data (e.g., 3 days/month) and of missing temperatures (e.g., 5 days/month) for the month to be considered complete. Generally, missing temperature values are of less consequence than missing precipitation because a few days of precipitation can account for most of a month's accumulation. For some climate regimes, such missing-tolerance levels may need to be far lower (e.g., sites prone to rapid Arctic frontal passage, summer rain primarily as convective storms) than for others (e.g., maritime climates with low day-to-day variability).

3.6.2 Infilling – Initial considerations

Infilling missing data should be processed last, after data errors, biases, and inhomogeneities are taken care of.[58] First, evaluate how much of the record is missing – and if these gaps are random across the record or temporally biased (discussed earlier, §3.3.3: *Missing observations that are time dependent*). It is often sufficient to judge such biasing visually (e.g., Figure 10).

The importance of missing values and temporal biasing depends on analysis requirements. If you plan to use a statistical technique that permits missing data, then infilling may not be needed provided that gaps are not huge and their temporal distribution not particularly biased (e.g., Hirsch and Slack 1984). In deciding whether to infill, also consider previously mentioned cautions re the effect of estimated data on variance structure (§3.3.3, §3.4.3.3: *Reduction in variance*).

Also keep in mind you may need to treat missing values differently for different analysis goals. Infilled dailies may be required for creating a reliable monthly dataset – e.g., for calculating trends or as input to a resource model. On the other hand, infilled values will interfere with daily frequency distribution and extreme value analyses (§4.6). In this case, conflicting requirements can be handled using the same dataset for both purposes by flagging infilled data for omission when

[58] Concatenation (§3.5) can be done either before or after infilling. Concatenating before gives the benefit of having a longer record on which to base infilling, but the disadvantage that infilling relationships are based on a derived record for part of the period.

called for (§3.2); in other cases, a few versions of the dataset may be called for to meet diverse analysis goals.

In the next sections, I present several different infilling approaches – including statistical and non-statistical spatial models (§3.6.3–3.6.6) and temporal statistical models (§3.6.7).

Infilling using a single or few neighboring stations – Simple regression models

Spatial methods (covered here and §3.6.4) are powerful techniques for infilling missing values. Much like spatial techniques for correcting inhomogeneities, these methods evaluate and utilize correlations between a target station (with missing data) and its neighbors. The most straightforward technique uses standard linear regression analysis. In a simple implementation, the steps are:

(1) Select candidate nearby stations based on:

   1. Proximity
   2. Similar environmental controls over climate
   3. Completeness of record at times when values are missing at the target station
   4. Sufficient record length on which to build a statistically significant regression

(2) Run separate, simple regressions[59] (eq 1; §3.4.2) of the target record on each candidate station record covering a specified period. The specified period can be defined using various criteria – the objective is to optimize the predictive power of the regression model for a given situation. Examples are:

   1. Use the entire record (all years) for the corresponding season or month (e.g., all daily data across all Januaries)
   2. Use a moving window about a missing value's date (such as daily data from 2 weeks on both sides of the date for that year) – An advantage of this approach is that the regression is tuned to weather conditions happening around the missing date. Note that which neighboring station is selected as the best predictor can change as the window is moved to each missing point.

(3) From these regressions, select the single station with the greatest predictive power (highest statistically significant R^2). Follow the guidelines for accepting a regression model laid out in §3.4.3.2.

Refer to caveats and considerations regarding regression-based estimates in §3.4.3.3 and re spatial models in §3.6.5. If this method does not produce a workable model, explore using:

1. A multivariate regression model – Following the single-station procedure, but with terms for additional stations in eq (1).
   a. Note: Add in a limited number of other stations only to the extent that R^2 is significantly improved. Models with too many predictor terms, each adding small improvements in R^2, are prone to being 'overfit' (where noise is being modeled as if it was true signal) giving false confidence in the model's predictive power.

2. Spatial autocorrelation models – §3.6.4
3. Alternative, non-statistical spatial schemes – §3.6.6

[59] Simple regression is with only one independent variable, *x* (in this case, a neighboring station), in comparison to multiple regression with two or more independent variables.

4. Temporal models – §3.6.7

3.6.4 Spatial autocorrelation models

More sophisticated infilling techniques involve modeling a correlation surface through space and using this surface for prediction at a point – the target station. The spatial autocorrelation surface is created based on a relatively large network of stations (e.g., 10-100 stations) that are locally well correlated with the target station. Geostatistical techniques include kriging (Haas 1995) and thin-plate spline prediction models (Hutchinson 1995, 2004); the development and evaluation of a kriging model is illustrated in Figure 11 and Figure 12 (discussed in the next section, §3.6.5). A classic reference for spatial statistics is Cressie (1993a).[60]

In applying these methods, it is usually more effective to work with anomaly (means removed) or standardized anomaly fields (anomaly divided by standard deviations) than with original values.[61] This is because temperature and precipitation anomalies tend to be regional in scope and so more spatially coherent than original fields. Additional techniques simultaneously evaluate spatial and temporal autocorrelation to infill missing data (e.g., Kondrashov and Ghil 2006, 2007).[62] Interpretation of spatial autocorrelation analyses is also discussed in §4.7.2.

3.6.5 Considerations in implementing spatial statistical infilling models

The value of spatial statistical models (§3.6.3–3.6.4) is that they permit a statistical evaluation of the process – to be able to say how good a technique is in terms of the strength and significance of relationships used to build the infilling model. The following considerations will help design a successful protocol, while recognizing model limitations:

Spatial heterogeneity. As noted earlier (§3.3.2: *Spatial consistency*), spatial predictive power drops off with distance as a function of climate variable, time aggregation interval, and heterogeneity in climate-controlling factors. The effect of climate heterogeneity on a spatial model is shown in Figure 11 – while this illustration is for a kriging model, the issue pertains just as well to simple regression schemes.

Because of heterogeneity in climate-controlling factors, the best predictor may not be the closest station. Consider a number of neighboring stations as candidates for simple regression models (as in the implementation strategy in §3.6.3). One approach is to stratify a regional domain by controlling factors (e.g., topography) and look for stations in a stratum matching the target station (this is the approach in PRISM: Daly et al. 2002, 2008).

Correlations dynamic. Spatial correlations among stations are dynamic, e.g., with strong seasonal dependence. Allow for station selection and regression models in your protocol to vary at least on a monthly basis or with a moving window.

Timestep choice. Infilling can be applied at different timesteps with varying results (Figure 13). As noted (§3.3.2: *Spatial consistency*), spatial correlations are smoother at longer timesteps. However, if, for example, daily missing values are relatively limited in number and scattered (rather than forming long gaps), a better monthly record may be obtained infilling at the daily scale because more of the original record is retained (see Figure 13 caption).

---

[60] Another text is Clark and Harper (2000); the 1979 edition (by Clark) is online at: http://www.kriging.com/PG1979/ (as pdf: http://www.kriging.com/PG1979/PG1979_pdf.html).

[61] Following the notation in §3.4.4, the anomaly series $\delta\, x(\mathrm{t}) = [x(\mathrm{t}) - \xi]$ and the standardized anomaly = $[\delta\, x(\mathrm{t}) / \mathrm{SD}_x]$.

[62] Advanced methods for spatial and *spatiotemporal* data are covered in Banerjee et al. (2003).

It is best not to use regressions based on monthlies to infill daily values – monthly regressions will not properly capture daily variability.

Station density effects. Station density has a direct effect on how well we can model spatial relationships. For the kriging model discussed in Figure 11, model error is larger during periods of low station density (Figure 12a, b). A consequence of this is that as station density drops off, interannual variability is artificially diminished. This is because with lower station density, the scheme reaches farther out to less well correlated stations, so that poorly related signals are blended to create the infilled record (Figure 12c vs. d). This variance loss can be accounted for by adding a stochastic term into the spatial model (Cressie 1993b).

3.6.6 Alternate methods – When regression techniques do not work

Regressions can fail for stations where proximity and similarity in climate otherwise suggest they should work. Reasons for such failure include:

1. Period of overlap between target and nearby station is too short on which to base a reliable regression relationship
1. While both stations have comparable climate dynamics, daily events or even weekly values are not in phase over moderate spatial scales – e.g., under convective precipitation regimes.

The following alternate techniques, while practiced, are only advised in the case when statistical methods have failed. Generally speaking, this is because they are not statistically testable and because they can introduce undesirable features into data timeseries.

Longterm mean substitution. Insert the longterm (or some other period) day-of-the-year or monthly mean for missing days. Undesired effects are that (i) this technique under-represents temporal variability and (ii), for precipitation, it inserts spurious precipitation events into the record.[63]

Neighbor substitution. Insert the day's value from a neighboring station (or grid point) with offset or amplitude adjustment (§3.4.4). An advantage is that this process maintains some variability in the record based on the dynamics of a nearby station. A disadvantage is that there is no test as to whether the variability at the nearby station or grid point is related to that of the target station. In practice, this means that strong justification of station selection (or gridded dataset) is needed. Baron (2006) used this method to scale a gridded dataset's grid point timeseries to a station.

Distance-related interpolation. Use an inverse-distance (or other distance-dependent) interpolation of values from multiple neighboring stations (Wilmot et al. 1985, Chen et al. 2008). These spatial methods are distance-weighted in contrast to the previously presented statistical spatial models which are correlation-optimized. An undesirable effect is that this technique blends signals of the nearest stations – if these signals are not in phase, the result is a dampened signal.

[63] By using a mean value, an artificial, low magnitude 'event' is inserted in the record regardless of the synoptic weather conditions at the time – such as in the middle of a dry spell.

A variant of the distance-based scheme is to interpolate station anomalies (as suggested for statistical spatial models, §3.6.4). One implementation is to use a variance-scaled anomaly (standardized anomaly):[64]

$$\delta(t) = \langle \text{INTERPOLATION}_{(i=1,n)} \{ [x(t,i) - \xi(i)] / SD_x(i) \} \rangle \quad (6)a$$

$$y'(t) = \bar{y} + [\, \delta(t) \times SD_y \,] \quad b$$

where $\delta(t)$ is an interpolated standard deviation-normalized anomaly at time (t), derived from $x(t,i)$, $\xi(i)$, and $SD_x(i)$ which are observations, mean, and standard deviation for each of the i=1,n number of *neighboring* stations used in the interpolation.[64] In eq (6b): $y'(t)$, $\bar{y}$, and SD_y are the target station's infilled value at time (t), mean, and standard deviation, respectively. This technique adjusts for signal mean level and variance difference not accounted for in the basic distance interpolation scheme; however, signals still have the potential for being smoothed.

3.6.7 Temporal models – Serial correlation simulation

Data simulation based in temporal autocorrelation is an alternate, statistical technique. If a station's serial correlation is high, a temporal autoregressive (AR) model can be created to infill data using values in a station record on either side of missing points.[65] Regression terms in an AR model are observations at lag/lead times – most of the power in AR models tend to be at a lag/lead of 1 timestep so that a first-order model is sufficient. This approach can be valuable when there are no surrounding stations on which to base spatial methods as it uses only the target station's record. The disadvantage of this technique is that while it infills a record with an event structure that is characteristic of the station, the estimated values are synthetic (the values do not reflect actual events). Techniques that combine temporal and spatial correlations are mentioned at the end of §3.6.4.

More advanced versions of these models look for correlations with other variables at a station to guide this infilling. While used in other contexts, WGEN (Richardson 1981, Parlange and Katz 2000) is a temporal model that uses daily autocorrelation and cross-correlation[66] to stochastically simulate daily minimum and maximum temperatures and transition probabilities to model precipitation events.[67]

3.7 Document data changes and evaluate consequences

When you have a processed dataset ready to address your research questions, finalize the documentation of your data protocol. Layout your data clean-up methods, catalog dataset versions, record flag coding for adjusted, infilled, or omitted data (§3.2), and report tests that support your choice of methods. It is crucial to state assumptions, limitations, and caveats that accompany the techniques applied.[68] As mentioned earlier, keep in mind that estimated fields can in some cases artificially reduce variance in the data and lead to an overstatement of significance in statistical tests. Such frank discussion will be an important reference for others using the dataset, but also for you in interpreting your analyses. Do not neglect the documentation process.

---

[64] The INTERPOLATION function can be any distance-related or other reasonable spatial interpolation scheme.
[65] Wilks (2006: §8.3.1)[10] discusses temporal autoregression and AR models.
[66] Serial correlation (autocorrelation) is the lead-lag correlation for a single variable and cross-correlation lag-lead between variables.
[67] WGEN is designed to generate synthetic month-long and longer daily time series using a first-order Markov chain-exponential model for precipitation and a first-order autoregressive model for daily temperature that was conditional on precipitation. (Figure 18 shows the frequency distribution of WGEN-simulated dailies compared to station records)
[68] For one approach, see Kittel et al. (2004: §7.1)'s caveats.

Evaluate the effect of data adjustments and infilling as they carry into the statistical analysis phase (§4.0). Run different levels of unmodified and corrected data through to see how critical your decisions were, how robust your analysis results are. Be convinced that your data preprocessing choices had effects you are comfortable with. As part of documentation, do not neglect the evaluation process.

4.0 Analysis – Tools to Explore Critical Questions

With a credible, well-documented dataset tailored to planned analyses, you can pursue two complementary lines of study: discovery and hypothesis testing (Figure 1c). Discovery is aided by visualization and descriptive statistics, which help to develop an intuitive sense of the data, generate new hypotheses, and relate information to others. Hypothesis evaluation advances, with confidence, our understanding of a system and is accomplished through rigorous statistical tests.

Both discovery and hypothesis testing are components of analyses commonly applied to climate data, reviewed in the following sections. After discussing descriptive methods (§4.1), I introduce common derived variables that can help in such exploration (§4.2) and present temporal and spatial analytical techniques. Temporal analyses include those for:

- Longterm trends (§4.3)
- Covariation among timeseries (§4.4)
- Interannual variability and regime shifts (§4.5)
- Daily event structure and extremes (§4.6)

And spatial analyses for:

- Regional coherence (§4.7)
- Hemispheric teleconnections (§4.8)

I end with caveats re interpretation (§4.9) and a summary of key lessons for dataset creation and analysis (§5.0).

4.1 Description

Before proceeding with statistical tests, take advantage of descriptive methods to understand your data and visually explore your ideas. This process can reveal problems not already caught in dataset development (§3.0) and suggest the best analytical approach for testing your *a priori* hypothesized relationships (§2.0).

4.1.1 Descriptive statistics

Key descriptive statistics are:

2. Data distribution parameters – e.g., mean, SD, quartiles, skewness, outliers
3. For monitoring, diagnostic measures of recent observations relative to the historical record – e.g., departures from longterm mean, historical rank for extreme events. Departures can be expressed in original units, percentiles, or standard deviations.

4.1.2 Graphic methods

Helsel and Hirsch (2002: Chapters 1 and 2)[10] and Wilks (2006: Chapter 3)[13] present a wide variety of quantitative and graphic exploratory methods tailored to different discovery objectives. I list common graphic methods here, given by visualization goal and with examples from figures in upcoming sections (§4.3-4.8):

Data distribution (Univariate) –

1. Histograms – to show the frequency distribution of a variable across:
   a. Its own data range (e.g., Figure 17a,c)

   b. Range of a second variable influencing the first's occurrence (e.g., by month: Figure 17b)

1. Cumulative distribution graphs – to compare data distribution patterns of
   a. Two datasets (Figure 20c)
   b. A dataset vs. a hypothesized distribution

2. Box-and-whisker plots (§4.6.1) –
   a. To display data distribution features such as median, quartiles, outliers (Figure 18)
   b. Using multiple plots, to compare these features for different sites (Figure 18), seasons (Toews et al. 2007), or decades or by any other discriminating attribute influencing the variable's frequency distribution.

Multivariate: Relationships with other system variables –

- *x-y* scatter plots – to explore:
   a. How two variables co-vary
   b. How data from different categories of observations (different regions, seasons, etc.) break out in *x-y* space (e.g., with data domains delineated as in the style of Figure 2d).[69]

Temporal dynamics: To explore trends and interannual variability patterns –

- Timeseries plots – with values plotted on an absolute scale (Figure 15a) or as deviations from the longterm mean (Figure 15c), with or without smoothing (§4.5.1).
- Complemented by plots in the *frequency domain* (§4.4.1) to indicate periodic behavior (Figure 16a, Figure 22).

Spatial display: To explore geographic relationships –

- Mapped variable fields – a variable's distribution in space revealing gradients or distinct domains, such as with contoured fields (Figure 19c) or symbols (arrows in Figure 23a). Maps can be in absolute values or as anomalies from a spatial mean (e.g., contours and arrows in Figure 23a).

- Correlation maps – mapping temporal correlations of a spatially-distributed variable (such as temperature timeseries) with another, single timeseries.

    - The single series can be of the same variable at one key location – to illustrate spatial autocorrelation (§4.7.2; Figure 21a).

    - Alternatively, the single series can be of a second variable – to show their cross-correlation, such as with a hemispheric circulation index illustrating teleconnections (§4.8.2; Figure 24a).

Space-time: Evolution of spatial patterns with time –

- Animations – to display a time sequence of maps

[69] See also "polar smooth plots" in Helsel and Hirsch (2002: §2.3.2, Figure 2.28).[7]

- Time-longitude/Time-latitude section plots – two-dimensional display of spatiotemporaldata, where one spatial dimension is collapsed by averaging (e.g., Figure 14).[70]

4.2 Insights through derived variables

Analysis of composite and other derived variables can provide additional insights into site climates in terms of their thermal and drought regimes.

4.2.1 Thermal regime measures

Three commonly derived thermal variables are:

Diurnal temperature range (DTR). Diurnally, minimum and maximum temperatures (T_{min}, T_{max}) are strongly controlled by different local processes, those affecting daytime heating vs. nighttime cooling. However, day to day, they tend to be correlated because of multiday- to month-timescale effects – e.g., air mass advection and seasonal heat storage. This high correlation tends to obscure differences in T_{min} and T_{max} dynamics. On the other hand, DTR and mean temperature (T_{mean})[71] are generally orthogonal[72] and can be used to segregate processes controlling a climate's thermal regime – with T_{mean} showing multiday synoptic and seasonal effects and DTR the daily offsetting of local heating and cooling.

In terms of observed climatic trends, T_{min} and T_{max} have been changing at different rates over the recent record. We see that minima are often rising more strongly than maxima, so that DTR has narrowed with time (Easterling et al. 1997). DTR is a simple index of this change.

Accumulated growing degree days (AGDD). Accumulated growing degree days is a frequent measure of growing season conditions. By 'growing season,' we are generally referring to terrestrial plant phenology, but recognize that other groups of organisms have their own environmental cues. AGDD is the sum of mean daily temperatures that exceed a critical base temperature. Base temperatures (T_{base}) are set depending on application or ecosystem. A T_{base} of 5ºC has been used in general applications for temperate natural systems (e.g., Rehfeldt et al. 2006). However, where literature can support it, the limit should be one fitting the ecosystem or organisms studied. In montane and alpine environments, for example, 0ºC is considered a lower limit for plant growth (Billings and Bliss 1959, Kimball et al. 1973). The formula is:

---

[70] In addition to animations, another means to display the evolution of spatial processes with time are *time-longitude* and *time-latitude section plots* (also known in meteorological applications as Hovmöller diagrams). In this method, three-dimensional spatiotemporal data – two dimensions in space (e.g., latitude by longitude) and the third in time – are represented in 2-D by collapsing one of the spatial dimensions into an average. These plots cross-section the data by time and one spatial dimension, and so reveal how a variable's spatial pattern evolves over time along a longitude or latitude transect:

- In a *time-longitude section plot*, averages over a latitudinal range are calculated by each longitude position and timestep, and then presented as a contour plot with longitude on the *x*-axis and time on the *y*. This shows how a variable generalized for a latitude range changes with time and longitude (e.g., Figure 14).
- In a *time-latitude plot*, longitude averages are contoured on a plot with time on the *x*-axis and latitude on the *y*.

In sector plots, spatial dimensions need not be latitude/longitude. A method is at http://locust.mmm.ucar.edu/episodes/episodes_paper_technote.html#_Toc516370405 – their technique is highly detailed, specific to their data, but the figures illustrate the process. For additional examples, see the interactive site for tropical climate variables: http://www.pmel.noaa.gov/tao/jsdisplay/.

[71] Both calculated from T_{min} and T_{max}: DTR = ($T_{max} - T_{min}$) and, of course, $T_{mean} = (T_{max} + T_{min})/2$. DTR was discussed earlier in the context correcting station records (§3.4.4: Considerations for specific variables).

[72] I.e., independent. This is generally speaking the case for T_{mean} and DTR at daily timescales; on the other hand, DTR can vary seasonally with T_{mean} and regionally, for example, maritime areas with low seasonality in T_{mean} tend also to have low DTR compared to continental climates.

$$AGDD = \Sigma[T_{mean}(t) - T_{base}], \text{ for days (t) when } T_{mean}(t) > T_{base}, \quad (7)$$

summed over a year, and where $T_{mean}(t)$ is the average of daily minimum and maximum temperatures for day (t).

Frost-free period timing and length. The frost-free period is another indicator of growing season conditions. We can ask not only how the length of frost-free period changes year to year, but also about spring onset and fall termination date changes – as shifts in these two dates are not necessarily linked. Different ecosystems (and their different components) have different sensitivities to freezing temperatures – so we can use different cold temperature thresholds, T_{freeze}, depending on the application. A natural threshold for T_{freeze} is 0ºC, but a lower threshold such as –2 or –3ºC can be used to represent a 'hard' frost.[73] Example seasonal markers for spring and fall dates are:

- Last and first frost at *night* – based on $T_{min} \leq T_{freeze}$
- Last and first day with freezing *daytime* temperatures – based on $T_{max} \leq T_{freeze}$
- Last and first *run* of, for example, 3 days with nighttime frost – based on $T_{min} \leq T_{freeze}$

4.2.2 Drought indices

The purpose of drought indices is to capture the occurrence and duration of wet and dry spells. Heim (2002) reviews commonly applied drought indices; two well-known ones are:

Palmer Drought Severity Index (PDSI). The Palmer Drought Severity Index is a standardized measure of soil moisture supply, typically evaluated on a monthly basis (Palmer 1965).[74] PDSI is a common metric for determining when a dry or wet spell begins and ends, integrating the effects of both precipitation and temperature (through its control over evaporative demand) on surface water balance (e.g., Figure 14). While reported for a given point in time, the index includes antecedent soil moisture conditions and so reflects the accumulative effects of water deficit or surplus. Alley (1984) presents a method for calculating monthly PDSI based on monthly precipitation and temperature data.

While PDSI's soil water budget allows it to integrate effects of temperature and precipitation and to accumulate moisture deficits or surpluses, key limitations come from the budget's shortcomings. These include (1) difficulty in applying it over terrain with heterogeneous soils and topography and (2) its lack of runoff generation lags, of snow and frozen-ground lag effects, and of seasonally in the role of vegetation (Alley 1984, Heim 2002). Relative to SPI (discussed next), PDSI may be slow to identify an emerging drought and may underrate the magnitude of prolonged drought (Karl and Knight 1985). Heim (2002) further explores the utility and limitations of PDSI.

Standardized Precipitation Index (SPI). The Standardized Precipitation Index focuses solely on the precipitation component of drought and wet periods (McKee et al. 1993, Guttman 1999). SPI is positive for wet conditions, negative for dry. To give temporal context to current drought or water surplus conditions, the index is determined for retrospective timescales from most

[73] Note that many poikilotherms and plants (once hardening has begun) are not susceptible to freezing at temperatures a few degrees below 0ºC because of tissue solute levels and physiological adaptations (cf. Marchand 1996).
[74] A global, coarse-resolution PDSI dataset is available at: http://www.cgd.ucar.edu/cas/catalog/climind/pdsi.html .

immediate (proximate month and season) to sustained (multiyear) durations (Heim 2002).[75] SPI compares the current period's cumulative precipitation to the historical probability of reaching that amount of precipitation.[76] The SPI is the number of standard deviations that current precipitation totals are away from the historic median. The probabilities can be easily backed out of SPI using rules for normal distributions, e.g., ±2 SD correspond to roughly to the 2nd and 98th percentiles, respectively – so SPI values of –2 and +2 represent extreme dry and wet conditions relative to the historical record for that site and timescale.[77] Guttmann (1999) lays out the calculation method.[78]

Benefits of SPI include: (1) that it is relatively straightforward to calculate (requiring only precipitation records and no site attributes), (2) its short to longterm perspectives on current drought conditions, and (3) advantages over PDSI noted earlier. However, a key limitation is that it only evaluates the role of precipitation anomalies in the occurrence and intensity of drought, while temperature anomalies can be an equally strong contributing factor (Hu and Willson 2000).

4.3 Trend analysis

4.3.1 Regression analysis

A method for statistical evaluation of trends in climate and other environmental variables is regression analysis.9 Regression trend analysis applies eq (1) with x(t) now representing time t itself:

$$y'(t) = b_0 + b_1 t \qquad \textbf{(8)}$$

where the regression slope, b_1, is the calculated trend and y'(t) traces the variable's change due to the longterm trend alone.

Regression model assumptions. While regression is frequently used to study trends, it is not always an appropriate method. Referring back to earlier discussion re regression (§3.4.3.1), the standard implementation (OLS), while powerful, requires that observed data are independent (discussed shortly) and their distribution normal and homoscedastic (variance constant).[39] As noted previously, climate data do not always meet these criteria. Transformations to adjust data to normality and constant variance are noted in §3.4.3.1, along with corresponding issues.

[75] Current SPI maps for the conterminous U.S. (by Climate Division) for 1- to 72-month durations are at: http://www.wrcc.dri.edu/spi/spi.html; see also: http://lwf.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html. A historical, global gridded SPI data viewer is at: http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/SOURCES/.IRI/.Analyses/.SPI/SOURCES/.IRI/.Analyses/.SPI/SOURCES/.IRI/.Analyses/.SPI/html+viewer? – zoomable, interactive.

[76] This probability is taken from the *cumulative distribution function* (CDF) for the location's precipitation record. This function is derived (accumulated) from a gamma *probability density function* (PDF) fitted to monthly precipitation data. (A gamma distribution is often a good portrayal of precipitation data.) The relationship between cumulative and probability density functions is illustrated in Figure 20a; see also §4.6.3. An inverse-normal transform of the gamma-derived CDF nicely puts these probabilities in terms of standard deviations of a normal distribution – the value of SPI is the number of standard deviations from the longterm median precipitation

[77] As noted in footnote 76, the probabilities behind the SPI are transformed to a normal distribution. Conversions from standard deviations (and so, SPI) to percentiles are based on one-tailed probabilities. Other rough conversions are: ±1 SD = 16th and 84th percentiles, ±1.6 SD = 5th and 95th percentiles, ±2.3 SD = 1st and 99th percentiles, ±3 SD = 0.1th and 99.9th percentiles (see also Guttmann 1999). Note that distribution function fitting does not always do well at extreme tails of a distribution, so there is high uncertainty in extreme SPI values. (See related discussion: http://www.cpc.noaa.gov/pacdir/NFORdir/INTR.html – "Caution Required for the Tails of the Curves".)

[78] PACN (2008) gives a procedure for calculating SPI in Excel. Inclusion here is not an endorsement – before using, validate output against results from other methods.

If the two distribution assumptions are violated (even after transformation), use a robust technique –

- If non-normal – use a non-parametric method, such as Mann-Kendall test for trends (next section, §4.3.2)
- If normal, yet heteroscedastic – use a robust regression which models the changes in variance, such as WLS regression (§3.4.3.1)[41] and quantile regression (§4.3.3).

Adjusting for serial correlation. In timeseries, independence means that the observations are not serially correlated (no temporal autocorrelation). However, it is not unusual for geophysical data to be serially correlated. Needless to say, these assumptions are often ignored and regression trend statistics commonly reported in studies. A test for serial correlation is the Durbin-Watson test, which is applied to the detrended series[79] (Wilks 2006: p. 192; Helsel and Hirsch 2002: §9.5.4.1).

If serial correlation is significant and the correlation is positive at a timestep lag of minus 1 (r_1), then its effect can be accounted for in trend analysis by inflating the regression's mean sum of squares for error (*MSE*). This appropriately makes the regression's significance test more conservative. The inflation factor increases with r_1, such that:

$$MSE^{*} = MSE\,[(1 + r_1) / (1 - r_1)] \ , \qquad (9)$$

where MSE^* is the serial-correlation adjusted value of *MSE* (Wilks 2006: p. 194). From the adjusted MSE*, recalculate R^2, F-test statistic, and significance level p. [80]

4.3.2 Mann-Kendall trend test

The Mann-Kendall test for trends is a non-parametric method using ranked data. Advantages of this test are that it:

- Does not require normally-distributed observations
- Allows for missing data
- Is resistant to the effect of outliers
- Detects monotonic rather than strictly linear trends – with the added benefit of eliminating the need for nonlinear transforms
- Is often just as powerful as corresponding parametric tests such as regression (Lettenmaier et al. 1994)

While normality is not needed, constant spread in the data's distribution (homoscedasticity) is. However, with this relaxed set of requirements, workable transformations are more readily achieved

---

[79] A detrended timeseries is usually calculated as the residuals from linear regression (trends) analysis.

[80] To calculate the serial-correlation adjusted R^2* and F*:

- Back out the error sum of squares (SSE) and total sum of squares (SST) from adjusted MSE* using the definitions:
  - MSE* = SSE*/(n-2) $\rightarrow$ SSE* = MSE*$\times$(n-2)
  - SST* = SSR + SSE*

  where * indicates an adjusted parameter, n = number of observations, and SSR is the sum of squares for regression obtained from the un-adjusted regression analysis.
- Recalculate R^2* and F*, get new p
  - R^2* = 1 – (SSE*/SST*)
  - F* = MSR/MSE*
  - Determine the p from F* and degrees of freedom for MSR (df=1) and MSE (df=n-2) [or compare F* to $F_{critical}(\alpha)$]

  where MSR is the mean sum of squares for regression also obtained from the un-adjusted regression analysis.

than under the more restrictive assumptions of ordinary least squares regression.[81] In the Mann-Kendall test, climate observations are converted to ranks and the ranked correlation (climate rank vs. time) is tested with Kendall's *tau* statistic (Helsel and Hirsch 2002: §8.2, §12.2.1).[10] Hirsch and Slack (1984) give a method accounting for seasonal dependence in trends.

Because observations are ranked, this technique tests for monotonic trends (as already noted) – with the advantage of not needing to specify a linear or nonlinear model,[82] but with a notable disadvantage of not quantifying that trend. Hirsch et al. (1982) present a slope estimator, Kendall-Theil robust line, to accompany Mann-Kendall tests.[83] An alternative strategy is to use (1) the Mann-Kendall test for statistical assessment and (2) linear regression (with or without nonlinear transformations) to describe the trend.[84,85]

Accounting for serial correlation. A concern in applying Mann-Kendall is, like parametric tests, that it assumes independent observations. In evaluating this issue, Harcum et al. (1992) found that the test considerably overstated a trend's significance level when the detrended serial correlation at 1-year lag had a correlation coefficient $r_1 > +0.2$. They concluded that for $r_1 < +0.1$, the test was rigorous, with a grey area between +0.1 and +0.2.([86]) Evaluate this issue by (1) detrending the series[82] and (2) calculating the lag-correlation between this series and itself offset by one timestep. Kulkarni and von Stroch (1995) and Hamed and Rao (1998) offer methods for handling serial correlation in the Mann-Kendall test.

4.3.3 Quantile regression – Tracking heteroscedasticity

So what if a process is heteroscedastic with time – that the variance and other moments have trends? If we can see how a process varies over time in a manner far more complicated than just a changing mean, we gain deeper insights into how the system is actually working (Cade and Noon 2003, Beniston and Stephenson 2004). While standard regression considers the tendency of the mean, *quantile regression* follows the trends of quantiles (percentiles) of a variable's distribution evaluated for intervals along the time axis. It allows us to simultaneously track any part of the distribution, such as the median (50%-tile) and top vs. bottom 10%-tiles. The approach is robust as it models heteroscedasticity and has low sensitivity to outliers (Hao and Naiman 2007). Cade and Noon (2003) present examples of quantile regression from ecology. General references are Koenker (2005) and Hao and Naiman (2007).[87]

---

[81] If corrected by a power transform, the Mann-Kendall test statistic *tau* (discussed next) is variable-scale independent – that is, not affected by the transform and so is comparable across tests on original and variously transformed data (Helsel and Hirsch 2002: §12.2.1)[7] in contrast to scale-dependence in R^2 (§3.4.3.3: *Transformations and* R^2).

[82] vs. for regression techniques, §3.4.3.1

[83] See Helsel and Hirsch (2002: §10.1)[7] re the Kendall-Theil robust line

[84] In this approach, it is important to distinguish the roles of these two analyses when reporting results. While the Mann-Kendall test evaluates the statistical significance of an observed trend, it does not similarly assess the regression's quantification of the trend.

[85] Helsel and Hirsch (2002: §10.1.2)[7] compare Kendall-Theil robust line and ordinary least squares regression slope approaches.

[86] For the mirrored case where $r_1 < -0.2$, autocorrelation leads to an understatement of statistical significance (Hamed and Rao 1998).

[87] Koenker created an R-package for quantile regression: http://cran.r-project.org/web/packages/quantreg/index.html (http://cran.r-project.org/web/packages/quantreg/vignettes/rq.pdf). Koenker (2005) and Hao and Naiman (2007) are available in part on Google Books (see References).

4.4 Covariation among variables

4.4.1 Relationships in time and frequency domains

If two variables both vary in time, we can ask whether they co-vary in a manner that implies cause and effect (regression) or coordination (correlation). In park applications, these variables may be both climatic or one that is resource related, following research questions. Parametric and non-parametric regression methods just noted for trend analysis, and corresponding correlation techniques, can similarly be employed to compare two such timeseries: x(t) vs. y(t). For example:

- For regression –
  a. OLS and WLS regression (§3.4.3.1, §4.3.1)
  b. Kendall-Theil robust line (§4.3.2)
  c. Quantile regression (§4.3.3)

2. For correlation[88] –
  a. Pearson's *r*
  b. Rank correlation methods: Kendall's *tau*, Spearman's *rho*

Processes that vary in time lend themselves to additional questions and corresponding techniques:

- *Cross-correlation* and *Multi-lag regression* – To test for a *lag* or *lead* of one variable's dynamics relative to the other.[69,89,90]
- *Cross-spectra* – In addition to evaluating relationships in the *time domain* (as in methods just covered), we can ask how correlation is structured in the *frequency domain*.[91] Specifically, is their correlation (*spectral coherence*) concentrated in certain frequency ranges? And are their dynamics in these frequencies in *phase*, or do they lag/lead? Cross-spectral techniques are presented in §4.7.3, in the context of comparing station records of the same variable.

4.4.2 Cautions regarding time-series comparisons

In addition to taking care to adhere to these tests' assumptions (re data distribution, serial correlation, etc.) and regression implementation and interpretation guidelines presented earlier (§3.4.3), take note of the following cautions for timeseries comparisons:

Smoothed dataseries. Smoothing timeseries underrepresents observed variance (lost in temporal averaging) and increases serial correlation (information is blended among adjacent time points). Both effects compromise statistical analysis – they increase the risk of overstating significance in statistical tests (Type I error).[25] Consequently, smoothed data should be strictly avoided in statistical comparisons (discussed further in §4.5.1: Additional cautions).

[88] See Helsel and Hirsch (2002: Chapter 8 – Correlation)[7]

[89] A couple examples of studies using lag-lead cross-correlation between climate and biological variables are: Martinez-Yrizar and Sarukhan (1990), Braswell et al. (1997).

[90] Serial correlation in individual timeseries also interferes with cross-correlation tests. As an alternate approach to cross-correlation, Burnaby (1953) presented a test for comparing timeseries that are autocorrelated. Malmgren et al. (1998) used this technique to evaluate teleconnections.

[91] In the time domain, variables are paired by time index: $y(t)$ vs. $x(t)$. In the frequency domain, they are indexed along a range of possible oscillation frequencies, f: $y(f)$ vs. $x(f)$. In spectral techniques, f is generally treated as a narrow frequency band, rather than a single frequency. Univariate spectral analysis is presented in §4.5.2, bivariate (cross-spectral) analysis in §4.7.3.

If timesteps do not match. Trying to compare timeseries with differing timesteps can be troublesome. Some points:

- The high-frequency variance of a short-stepped timeseries is missing in long timestep data, and so cannot contribute to shared variance in correlation or explained variance in regression. Instead, this variance ends up in the error term, reducing the power of these tests.
- There are various approaches to force timesteps to match. If the longer timestep series is interpolated to the shorter step, the resulting series becomes strongly serially correlated (information is repeated at the more frequent time points) and the apparent number of observations (and so degrees of freedom) becomes inflated. If unaccounted for, these effects increase the likelihood of a Type I error (overstating significance).
- The more common approach is to rescale short timestep data to the longer step, such as through aggregation as by summing or averaging across fixed time-intervals[92] (giving, for example: AGDD, total annual precipitation, and mean monthly temperature) or binning by event categories (e.g., number of extreme cold events in winter).[93] If rescaling is done through smoothing, however, issues discussed above arise. With aggregation, keep in mind there is a loss of shorter-term information – a concern if scale interaction plays a role in the processes being compared (a topic covered shortly: *Process timescales differ*).

Processes out of phase. If one process has a lagged effect on the other, standard correlation and regression techniques are likely to miss or underestimate a relationship. Cross-correlation and multi-lag regression (as just noted, §4.4.1) can be used to explore this possibility.

Process timescales differ. Regression and correlation analyses are straightforward when the temporal scales of processes of interest reasonably match.[94] On the other hand, in looking for relationships between a slow-moving process and a fast one, the mismatch in scale complicates matters. This is because (1) processes may only interact in a narrow, shared range of frequencies or (2) they interact across scales (§2.2: *Scale interactions*). The three modes by which slow and fast-moving processes interact and corresponding analysis options are:

- *Mode 1: Processes interact only at shared frequencies.* Cross-spectral analysis (§4.7.3) can be used to show at what frequencies two processes correlate well and if this coherence is in phase or lagged. Note that coherence and phase do not prove there's a physical link between the two processes, but can be the basis for hypothesizing mechanisms.
- *Mode 2: Slow process constrains a fast process*. From the perspective of a fast process, a slow process operates between relatively stable states (phases) and potentially constrains the dynamics of the fast process. We can ask if these phases set the stage for the high-frequency process, resulting in its distinct behavior. An analysis strategy to assess such phase dependence is to:

  (1) Block the slow-moving dataseries into periods when the system is in more or less stable phases.[95]

[92] As distinguished from running averages, which constitutes smoothing.
[93] As another example, the Climate Extremes Index (Gleason et al. 2008) uses a season-aggregate index of extreme events (Figure 15a).
[94] Provided, needless to say, that variation at these scales are adequately captured by their datasets.
[95] Data blocking is also discussed in §4.8.2, re testing teleconnections. See also Helsel and Hirsch (2002: §7.3 – Blocking).[7]

(2) Contrast behavior of the fast process under the different blocks of the slow process. This can be evaluated using analysis of variance (ANOVA) or other parametric and non-parametric techniques to test whether a significant part of faster process's variance is explained by the slow process's phases.[96]

- *Mode 3: Fast process determines the state of a slow process.* Critical, threshold events in a fast process (e.g., extreme weather events) may control the outcome of a longer-acting process. The trick in this type of analysis is to select the appropriate fine-scale variable, such as crisis weather events – this selection is aided by having detailed knowledge of the system. For example, in an ungulate demographic study, Hallett et al. (2004) evaluated the effects of irregularly-timed winter weekly low temperature, high rainfall, and high wind events on interannual variability in mortality.[97]

These approaches will only work, of course, if processes of interest are represented in the data at timesteps corresponding to the timescales at which they constrain or impact each other.

Interpretation of correlation and regression results. Caution is needed in interpreting both significant and nonsignificant results in correlation and regression analyses. As we are always reminded, "correlation does not imply causation" – that is, be careful not to take significant results as validation of hypothesized mechanisms. See §4.9 for further discussion, especially in regards to nonsignificant results.

4.5 Interannual variability – Spectral analysis and regime shifts

An important temporal feature of climate is oscillatory patterns at interannual, multidecadal, and longer scales.[98] These often suggest a link to climate system processes that exhibit similar characteristic modes of behavior. Such oscillations tend to be quasi-periodic, that is, tending to fluctuate within a band of frequencies rather than with a strict return period. Their dynamics are often partly obscured by 'noise' and processes operating at other frequencies and so need to be examined using techniques that separate out signals by frequency bands. Signals of interest can be explored visually using low-pass filters (next section, §4.5.1) and evaluated numerically using spectral analysis (§4.5.2). Another key dynamic of climate is regime shifts at multidecadal and centennial timescales; I cover techniques for their analysis in §4.5.3.

Linkage of such local interannual variability to regional and global dynamics is explored later on in sections on spatial pattern analysis, §4.7 and §4.8, respectively.

4.5.1 Smoothing filters

Smoothing filters are low-pass filters, allowing only lower frequencies of a timeseries through. The simplest are moving averages (running means) using *unweighted averaging* (also, called 'rectangular filters') (e.g., Figure 15a). These are typically applied over an odd number of years

---

[96] The two-sample *t*-test is another parametric technique for comparing means of one variable by blocks of another in the case of when there are just two blocks. Parametric and corresponding non-parametric tests are discussed by Helsel and Hirsch (2002: Chapter 7).[7]

[97] This study was also discussed in §2.2: *Scale interactions*

[98] A primer on statistics related to climate variability is: http://www.nws.noaa.gov/om/csd/pds/PCU2/statistics/Stats/part1/SPrimer1.htm and .../part2/SPrimer2.htm. For accompanying glossaries for statistical terms: http://www.nws.noaa.gov/om/csd/pds/PCU2/statistics/glossary.htm and dynamical meteorology: http://www.nws.noaa.gov/om/csd/pds/PCU2/meteorology/glossary.htm

and time registered on the middle year of the moving window.[99] Other common low-pass filters use *weighted averages* and span a sufficient number of years to remove high frequency variation in the data (e.g., Figure 15b, c).

Example weighted filters across a range of filter widths, as applied to annual data, are:

- 3-year, (1-2-1) weighted filter – this simple 'triangular filter' is commonly used for removing highest (interannual) frequencies. The weight denominator is 4, so the weights are ¼ - ½ - ¼.
- 5-year, (1-3-4-3-1) filter – for removing interannual through 3-5 year variability (cf. Trenberth et al. 2007: Appendix 3.A). Weight denominator = 12.
- More complex functions for emphasizing decadal processes – such as, the 13-year scheme in Figure 15b and applied in Figure 15c.

Numerical issues with smoothing include two artifacts:

End effects. Because these filters are calculated at the center point of the moving window, they cannot be calculated for the first and last few points of a series. The resulting timeseries will then be shorter than the original. These points can be left blank or padded with a value that makes sense in the context to your application: such as the longterm mean (zero, if an anomaly series) or the proximate filtered value.

However, it may be desirable for these values to reflect information from the original series's end points. One method is to apply a filter with stepwise decreasing span length as the filter center approaches an end point – with the caveat that the ends will have higher frequency variability than in the main body of the smoothed series. Mann (2004) presents an alternate technique that optimizes a choice among three constraints on the smoothed line as it approaches the end points: these constraints are on the smoothed line's local departure from the longterm mean, slope steepness, and slope change (the second of these was applied in Figure 15c).[100]

Spectral side lobes. Some filters, including weighted filters, create 'side lobes' in power spectra (§4.5.2) outside of the frequency range intended to be captured. As a result, higher frequency variance inadvertently shows up in smoothed (low-pass filtered) signals.

Additional cautions are:

Not for statistical analysis. Smoothing is an exploratory tool, to aid visualizing temporal patterns. As mentioned earlier (§4.4.2), the resulting series has severely reduced variance and elevated serial correlation and so should not be used in statistical testing. They also should not be used in spectral analysis (§4.5.2) due to the introduction of side lobes and any end-effect corrections.

Keep these caveats in mind when obtaining processed dataseries from other sources, such as teleconnection indices (§4.8.2). To this end, check documentation to see if processing included smoothing and, if so, check for the availability of unsmoothed versions.

[99] Avoid using non-centered (e.g., prior-moving) averages as this offsets peaks and troughs with respect to actual timing. These are occasionally found in the literature and are what is provided by the moving average option in Excel's graph trendline function.
[100] Mann's Matlab routine is provided at: http://www.meteo.psu.edu/~mann/Mann/tools/tools.html (relocated from that given in Mann 2004).

Missing the details. Smoothing climate data may become so routine for monitoring mid- and longterm behavior that we miss or discount unusual, yet critical events. As an example, Verosub and Lippman (2008) note this caution with respect to tracking effects of single-year global climatic events, such as volcanic eruptions, in regional and local climate records.

4.5.2 Spectral analysis – A look in the frequency domain

Spectral analysis is a method for formally exploring time series for oscillatory behavior (Yiou et al. 1996, Ghil et al. 2002). A common spectral technique is Fourier analysis which identifies at what frequencies the data most strongly vary. Results are commonly portrayed as a *power spectral density function*, where oscillation strength (spectral power) is plotted against frequency (Figure 16a). Peaks show at what periods (= 1/frequency) a climate record varies.

Narrower peaks reflect oscillations with tighter return periods vs. quasi-periodic dynamics shown by broader peaks. Occurrence of certain characteristic patterns can suggest linkage to continental and hemispheric climate dynamics, such as El Niño (more on teleconnections follow in §4.8). Evaluating whether spectrum peaks are statistically significant is an important component of spectral analysis (Figure 16a).[101] Without statistical evaluation, it's too easy to place undue significance on frequencies corresponding to peaks in spectral power.

Data are commonly preprocessed for spectral analysis by removing trends and, if present, the seasonal cycle. This processing removes corresponding peaks in the spectrum, making results more straightforward. Trends can be removed by determining residuals from linear regression;[82] the seasonal cycle removed by subtracting corresponding longterm seasonal means.[102] As discussed earlier (§4.5.1), spectral analysis should not be run on smoothed data because of:

- Filter side lobes – frequency spillover into higher frequency bands
- End-effect corrections, if incorporated into smoothed series – also blurring filter frequency boundaries

– both of which significantly affect spectral density (von Storch and Zwiers 2001).

The utility of Fourier-based spectral analysis is limited (1) when periodic signals are nonstationary (their frequency changes) and (2) for short timeseries (Kestin et al. 1998). Wavelet analyses is an advanced technique for following changes in the frequency of oscillations and will work for short records. Torrence and Compo (1998), Gedalof and Smith (2001), Ghil et al. (2002), Chang et al. (2004), Gray et al. (2003), and Labat (2005) present this method as applied to climate studies.

4.5.3 Regime shifts

Shifts in features that make up regional climates are not uncommon, having occurred at times over the last century and longer timescales (e.g., Pederson et al. 2006, Diaz et al. 2008) (Figure 16b). These shifts, which run through both the abiotic and biotic environment, tend to be linked to regime changes in hemispheric atmosphere and ocean circulation (e.g., Hare and Mantua 2000, Gedalof and

[101] Spectral analysis software with confidence interval capabilities include –
- IDL: Coherence function
- on-line spectral calculator: *Spectral Analysis (v1.0.6) in* Wessa (2009) http://www.wessa.net/rwasp_spectrum.wasp
- *Singular Spectrum Analysis - MultiTaper Method (SSA-MTM) Toolkit* (http://www.atmos.ucla.edu/tcd/ssa/)

These packages are presented as examples, not as an endorsement or reflecting an assessment.

[102] Longterm seasonal means are longterm means by day, week, month, or season corresponding to the timestep of the observed data. 'Longterm' averaging is over station record, 30-year normals, or other set period.

Smith 2001; such cross-scale geographic linkages are discussed in §4.8). Regime shift step-detection techniques are presented by Box and Tiao (1975),[103] Biondi et al. (2002), and Rodionov and Overland (2005).[104]

Changes in regime are also expressed in records as breaks in trend slopes. Breakpoint analysis techniques in climate and ecological studies include piecewise-linear (segmented) regression (Tomé and Miranda 2004, Marlon et al. 2009)[105] and quantile regression (Koenker and Schorfheide 1994)

Regime analysis is best done on records that are from the start free of artificial temporal inhomogeneities (e.g., station changes) as these change points complicate shift detection. Note also that methods correcting data heterogeneities (§3.4) run the risk of removing true regime shifts. Such corrections should be (1) omitted from datasets bound for regime change studies or (2) applied carefully only for the clearest station change cases, and then well documented and followed up with a review of regime shift results for interference at correction change points.

4.6 Daily analysis – Structure and extremes

At finer temporal scales, analysis of daily and hourly records can reveal the characteristic structure of weather events and the frequency of extremes for a site. These features are generally a reflection of a region's climate. Shifts in their structure can reveal important climate changes as much as altered means and interannual variability do. Meehl et al. (2000) and Trenberth et al. (2003) review modes by which the character of weather events change under altered climate. Among recent observed shifts in daily structure are more days with warm temperature extremes, fewer cold extremes, and greater occurrence of extreme daily precipitation (Easterling et al. 2000, Trenberth et al. 2007).

This section presents means for characterizing events (§4.6.1) and ways to assess their change in terms of:

- Event structure parameters (§4.6.2)
- Frequency distribution functions (§4.6.3)
- Extreme value analysis (§4.6.4)

Records with corrected or infilled values (§3.6) should be omitted from these analyses.

4.6.1 Event structure characterization and display

Event structure can be evaluated graphically with frequency distribution plots (Figure 17). In this technique, occurrences of a given event are 'binned' by:

- Event magnitude (Figure 17a, Figure 20d)
- Timing (season, time of day; Figure 17b)
- Duration or prevalence (Figure 17c, Figure 19b)
- Other variable offering a perspective on event structure.

---

[103] See, for example, Pederson et al.'s (2006) implementation of Box and Tiao's (1975) intervention analysis and other techniques in the detection of regime shifts.
[104] Rodionov and Overland's model is available at: http://www.beringclimate.noaa.gov/regimes/index.html
[105] Tomé and Miranda (2004)'s method is implemented in Miranda and Tomé (2009). Marlon et al. (2009) used the R package 'Segmented' (Muggeo 2009). See also references re use of breakpoint analysis in testing record homogeneity (footnote 55).

A more analytical presentation is box-and-whisker or notch plots[106] which express the median, standard deviations (SD), quartiles, range of values, and outliers (Wilks 2006: §2.5)[13] (Figure 18). For variables with highly skewed frequency distributions (such as precipitation), values can first be transformed to normalize the data. As noted earlier, typical transformations for precipitation include natural log (Figure 18), square root, and cubic root (see §3.4.3.1 for methods; cautions re frequency-distribution transformations are covered shortly). Toews et al. (2007) present a box plot application for seasonal analysis of event structure.[107]

For precipitation and other discrete-event variables that have a reasonable likelihood of zero values (e.g., dry days), only non-zero values should be incorporated in box plots and related analyses.[108] These plots then show the structure of precipitation events just for when they occur. An additional graphic can be employed to show the frequency of whether or not there's an event (as in, number of wet vs. dry days).

Issues in the analysis of daily event structure include:

Observation biases. For precipitation, observer practices can lead to underreporting the smallest events and a biasing toward certain frequency bins over others (e.g., favoring recording daily values in multiples of 0.05 inches). Daly et al. (2007) evaluates the consequences of these biases.

Corrected data. Omit inserted precipitation daily values parsed from multiday accumulated totals (§3.3.3), infilled (§3.6), and other corrected data that may not adequately capture (and instead blurs) a site's true daily event structure. Be aware that frequency distribution parameters and estimation of frequency distributions (especially with respect to extremes, §4.6.4) can be highly sensitive to data problems and their correction.

Back transformation of mean, percentile, and interval parameters. When nonlinear transforms are applied to event data, caution is needed in back transforming structural parameters to the original linear scale. The mean calculated in transformed space generally has little correspondence to the variable's actual mean (in linear space), while the median (50%-tile) and other percentiles do.[109] Any *interval* – such as interquartile range (IQR, cf. Figure 18), standard deviation (SD), and confidence intervals (CI) – determined in transform space and back transformed also has little meaning. Instead, an interval should be represented in terms of its *upper and lower limits,* where these are determined in transform space and back transformed to linear values.

4.6.2 Changes in event structure parameters

Altered daily or hourly structure can be assessed by analyzing changes in *event structure parameters* such as those used to characterize events (§4.6.1). These can evaluated simultaneously using quantile regression (e.g., Beniston and Stephenson 2004; §4.3.3) or individually with trend and interannual variability analyses (e.g., Figure 19d; §4.3, 4.5). Candidate parameters include:

[106] Box-and-whisker plots are also known as 'box plots.' The difference between box and notch plots is described in Figure 18.
[107] This application is for the 'R' statistical computing environment. For box plot graphing in general, PACN (2008) gives a procedure for Excel.
[108] Using only non-zero values is also practical if a logarithmic transform is selected [because *log*(0) is undefined].
[109] Helsel and Hirsch (2002: §9.6.3)[7] discuss corrections for transformation biases in the mean.

- Frequency of events in a specific range of a binning variable – such as a specific bin in a frequency distribution histogram (as those in Figure 17; see §4.6.1 re binning)
- Frequency of events beyond a given threshold (threshold exceedance) – as might define extreme events and be based on, for example:
  a. A threshold critical to the system – such as temperatures below freezing or a critical precipitation rate (e.g., Mearns et al. 1984, Beniston and Stephenson 2004)
  b. An upper or lower percentile (e.g., Figure 19a; Climate Extremes Index thresholds per Gleason et al. 2008: Figure 15a)[110]
- Frequency distribution parameters – as those given in box plots
  c. For example, the interquartile range (IQR, Figure 18) determined over short intervals in the record (e.g., by year or decade)

4.6.3 Changes in frequency distribution functions
Figure 19a shows changes in nighttime temperature extremes as a timeseries. Such change can also be evaluated in terms of shifts in its *relative frequency distribution* (Figure 19b).[111] Changes in this distribution can be analyzed using a test for identical distributions, such as the two-sample Kolmogorov-Smirnov test (Conover 1999).[112,113] This nonparametric test is sensitive to differences in shape or position of a frequency distribution – both types of changes are illustrated in Figure 19b.

To apply the Kolmogorov-Smirnov identical distribution test, we use a cumulative form of the relative frequency distribution of a variable. This is as the *cumulative probability distribution function,*[114] where relative frequencies are accumulated (summed) along the x-axis until all observations are accounted for.[115] Figure 20a, b illustrate the relationship between these two distribution functions and how changes in the shape of one is reflected in the other. The Kolmogorov-Smirnov test evaluates if the separation (D in Figure 20c) between the cumulative distribution functions of two sets of observations is significant (Figure 20b, c, d).

Another technique for comparing frequency distributions is quantile-comparison (or quantile-quantile) plots, where the percentiles of two distributions are plotted against each other.[116] These graphics are most useful if accompanied by plotted confidence intervals or statistical tests to evaluate if the distributions are identical.

---

[110] This would be a reduced form of quantile regression.
[111] Relative frequency distribution also referred to as the *probability distribution (or density) function* (PDF).
[112] The 2-sample Kolmogorov-Smirnov test for identical distributions is also known as the Smirnov test (Conover 1999). This test is distinguished from a *one-sample* Kolmogorov-Smirnov test, which is a goodness-of-fit test used to evaluate an observed distribution against a theoretical one (e.g., to test for normality[36]).
[113] An on-line calculator for the 2-sample test is at: http://www.physics.csbsju.edu/stats/KS-test.html (Kirkman 1996).
[114] Also referred to as the *cumulative distribution function* (CDF) or *cumulative probability distribution* (CPD)
[115] A re-expression of the cumulative distribution function is the *probability of exceedance* = [1 - (cumulative distribution function)]. The CDF gives the percentile for a value among all observations, the probability of exceedance gives the chance of observing a value (e.g., a precipitation amount) above a certain level and corresponding return period. This approach is often used in extreme value analysis techniques (§4.6.4) and extended-range weather outlooks: http://www.cpc.noaa.gov/products/predictions/90day/.
[116] Q-Q plots in Helsel and Hirsch (2002: §2.2.5)[7] and Wilks (2006: §4.5.2).

4.6.4 Extreme value analysis

Changes in occurrence of extremes can arise from (1) shifts in location, dispersion, and asymmetry of the frequency distribution (i.e., in the mean, variance, skewness, and other moments) and (2) changes in the structure of the far ends of the distribution (change in tail shape) (Meehl et al. 2000). While testing for overall changes in frequency distribution can rely on methods discussed in §4.6.3, evaluating trends in extreme events requires detection of changes in the tails of frequency distributions – where extremes lie but whose distribution is difficult to estimate by standard statistical methods.

Katz and Brown (1992), Kharin and Zwiers (2000), and Goubanova and Lia (2007) describe and implement an extreme value analysis technique based on a *generalized extreme value distribution* (GEV) which selects from three possible asymptotic distribution models fitted to extremes in the record. Frei and Schär (2001) present a method for detecting trends in extreme event return periods. Coles (2001) and Hosking and Wallis (1997)[117] provide introductions to extremes value analysis in theory and practice.[118,119]

4.7 Spatial pattern analysis. I: Regional connections

Once the temporal dynamics of a site's climate has been explored (§4.3–4.6), we can ask how its behavior fits in with the region or farther afield: Is the station representative of a park and the region? Or does its dynamics contrast with neighboring stations? Tools to assess spatial coherence include geostatistical and cross-spectral techniques (§4.7.1–4.7.3).

These analyses can be complemented by on-line national monitoring and outlook products which provide a near-real time, regional perspective on a park unit's climate. A selection of these is presented in Table 2.

4.7.1 One-point correlation map

A *point-correlation map* displays the spatial distribution of correlation coefficients between the record at a single location with those of other points within a domain (e.g., Figure 21a). These maps are a straightforward means to reveal spatial connections between a site and surrounding region.[120] Interpretation of these maps is facilitated if statistical significance of correlations is also indicated (e.g., by stippling, as on Figure 19d in another application).

---

[117] Hosking and Wallis (1997) available in part on Google Books (see References).

[118] Online resources re extreme value analysis (see Stephenson and Gilleland 2005 for a review of software):

- http://www.cru.uea.ac.uk/projects/mice/html/extremes.html
- http://www.met.rdg.ac.uk/~han/Extremes/extreme1.pdf - a presentation (Stephenson 2002)
- For S-Plus, R, MATLAB, others: http://www.rap.ucar.edu/staff/ericg/softextreme.php; specifically for R: http://www.isse.ucar.edu/extremevalues/evtk.html. See also: http://www.rap.ucar.edu/staff/ericg/extremereading.html
- Statistical Tool for Extreme Climate Analysis (STECA): http://www.cics.uvic.ca/scenarios/index.cgi?Other_Data#steca

[119] See also Makkonen (2008) for additional discussion of this method and related problems.

[120] An online facility for creating correlation maps is http://www.esrl.noaa.gov/psd/data/correlation/. While intended for hemispheric teleconnection correlations (§4.8.2), it can be adapted to creating site point-correlation maps by inserting a site station record as the custom timeseries (http://www.esrl.noaa.gov/psd/data/correlation/custom.html). See help and instruction in the page's left frame.

4.7.2 Spatial variation – Semivariograms

We discussed spatial autocorrelation earlier in the context of infilling missing data (§3.6.4). These techniques can also be used to assess how homogeneous vs. heterogeneous climate is across a domain.

Related to autocorrelation, the semivariogram is the spatial variance as a function of distance.[121] The semivariogram reveals how rapidly or slowly stations become less related to each other the farther apart they are (Figure 21c).[122] The relationship with distance can additionally depend on direction (anisotropy; evident in Figure 21d – see caption discussion).

The semivariogram is usually assessed for a point in time in the record (e.g., a day, week, or month depending on the climate process being evaluated: e.g., February 1996 in Figure 21c). The shape of a domain's semivariogram often changes with season, linked to seasonal climate processes. Controlling for season, the semivariogram can be relatively stable over decades, especially if it is controlled by topography (Fuentes et al. 2006). At the event level, the shape can be tied to synoptic conditions (rainfall type) as well as elevation (Şen and Habib 2000). Mapping the semivariogram at a set distance (e.g., 40km in Figure 21d) can reveal a regional climate's spatial connectivity and anisotropy (Figure 21d, e – see caption). Spatial statistical references, techniques, and issues are discussed further in §3.6.4–3.6.5.

4.7.3 Spectral coherence – Correlation in the frequency domain

Just as spectral analysis (§4.5.2) reveals characteristic oscillatory behavior in a single station's climate record, cross-spectral analysis asks if 2 stations' records co-vary at similar frequencies.[123] Cross-spectra show in which frequency bands station temporal dynamics are highly correlated (spectral coherence, h^2) and if this coordination is in or out of phase (Figure 22). Coherence is the squared correlation coefficient for a given frequency band (von Storch and Zwiers 2001) and so, in a way similar to regression R^2 (§3.4.3.2), can be interpreted as the percent variance shared by the two series in that range of frequencies.

Techniques for cross-spectral analysis include Fourier and wavelet approaches (Ghil et al. 2002, Whitcher et al. 2000).[104] Such analyses are most useful if they test for statistical significance in coherence and phase (Figure 22b, c).

4.8 Spatial patterns. II: Hemispheric teleconnections

As our ability to monitor the global climate system has grown over the last half century, we have become increasingly aware of connections between local climate variability and remote global-scale atmospheric and ocean dynamics. Such 'teleconnections' are fundamental to our understanding the roots of regional interannual climate variability and to exploring the mechanisms by which

---

[121] The semivariogram (γ) for variable z is: $\gamma(h) = \frac{1}{2n(h)} \sum [z(x) - z(x+h)]^2$, where x is any location, h is the separation distance between pairs of data points, and $n(h)$ is the number of pairs that are separated by h; summation is over all such pairs. This equation expresses γ as a function of h (Figure 21c). Note that γ is calculated as *half* the spatial variance, hence '*semi*-variogram.'

[122] Note that semivariance and autocorrelation reflect the same spatial process, but in the opposite manner: as stations become less related with distance, semivariance increases (Figure 21c), while autocorrelation decreases (Figure 11b, d).

[123] Cross-spectral analysis was first introduced in the context of comparing timeseries of two variables (§4.4.1, §4.4.2).

hemispheric processes scale down to the ecology of species and landscapes (cf. Stenseth et al. 2002, Graumlich et al. 2003, Stenseth and Mysterud 2005, Pederson et al. 2006).[124]

In this section, I briefly introduce well-recognized teleconnections (§4.8.1), present analytical approaches (§4.8.2), and discuss key implementation concerns (§4.8.3).

4.8.1 Modes of variation in hemispheric circulation – Multiyear oscillations

Predominant teleconnections for North America have their source in four major interannual to multidecadal oscillations of the climate system:

- El Niño-Southern Oscillation (ENSO)[125,126]
- Pacific Decadal Oscillation (PDO)[127,128]
- Northern Annular Mode (NAM)/North Atlantic Oscillation (NAO)[129]
- Atlantic Multidecadal Oscillation (AMO)[130]

These climate system oscillations have characteristic centers of action typically in places where there is strong coupling between the ocean and atmosphere (e.g., Figure 23a, *top*). These centers are quasistationary in their location, constrained by ocean basin geometry and basin-wide ocean circulation. The centers have characteristic quasiperiodic, multiyear modes of behavior in both ocean and atmospheric measures [e.g., sea surface temperatures (SST), sea level pressure (SLP); Figure 23a, b]. These oscillations are controlled by long-acting, geographically-broad interactions between atmospheric and ocean circulations and run deep vertically in both systems. These dynamics are not fully understood, however.[131] For brief overviews of these oscillation systems, their centers of action, temporal dynamics, and teleconnections, see Stenseth et al. (2003).[132]

---

[124] Comparable analyses for marine ecosystems include: Hare and Mantua (2000; with some terrestrial measures) and Schwing et al. (2009)

[125] General ENSO reference: Trenberth (1997). See also UCAR tutorial webcasts re ENSO under: http://www.nws.noaa.gov/om/csd/pds/PCU2/IC2.4.shtml (free registration, login)

[126] For looking at El Niño-related Pacific Basin dynamics but in the Northern Hemisphere, a counterpart to the Southern Oscillation (SO) is the Northern Oscillation (Schwing et al. 2002). However, indices for the two oscillations show similar dynamics and, in the literature, the SO Index (SOI; Table 4) remains the more common of the two for North American analyses. (cf. http://www.pfeg.noaa.gov/products/PFEL/modeled/indices/NOIx/noix_bkgrnd.html)

[127] General PDO reference: Mantua et al. (1997). Re the PDO and what is called the North Pacific Oscillation (NPO): The term "North Pacific Oscillation" is unfortunately used in the literature to refer to two distinct interannual dynamics in the North Pacific – either as (1) equivalent to the PDO (Gershunov and Barnett 1998) or (2) an oscillation whose spatiotemporal pattern is orthogonal (independent) of the PDO (Minobe and Mantua 1999). The first pattern predominately affects climates across North America (Trenberth and Hurrell 1994, Hurrell 1996, Mantua et al. 1997), the second pattern influences climates of the western Pacific (Linkin and Nigam 2008). Neither of these should be confused with the Northern Oscillation.[126]

[128] Much discussion re the PDO focuses on multidecadal regime shifts at roughly 15-25 and 50-70 year cycles (Minobe 1997, 1999). In addition, the PDO Index (Table 4) has substantial interannual variability (Figure 23a) – index fluctuations in and out of + or – territory can seen as oscillations within a PDO regime. However, whether regime shifts are actually characteristic of the PDO is evaluated using the paleorecord by Gedalof et al. (2002).

[129] The NAO is considered part of NAM dynamics and so are grouped together. The NAM is also referred to as the Arctic Oscillation (AO). General reference for NAO: Hurrell et al. (2003); for NAM/AO: Thompson and Wallace (2000)

[130] See Schlesinger and Ramankutty (1994), McCabe et al. (2004)

[131] In addition, these dynamics are not considered to be genuinely captured by today's global climate models, with implications re our ability to understand how these key modes of climate variability may change under future climates.

[132] See also: Steward (2005): http://oceanworld.tamu.edu/resources/oceanography-book/oceananddrought.html

For land and inland-waters resource science applications, we are interested in these oscillations' teleconnections to the climates of continents adjacent to (or of ocean islands embedded in) corresponding oceans, but removed from the centers of action. These teleconnections are the downstream consequences of location and strength changes in the centers-of-action's (1) warm and cool pools of ocean water and (2) semipermanent high and low pressure systems in the lower troposphere. Consequences are shifts in the position and intensity of the Intertropical Convergence Zone, Mid-latitude Jets, and Subtropical and Polar Highs – and, linked to these, of tropical and mid-latitude storm tracks, summer monsoons, winter advection of warm, moist or cold, dry air masses, and marine layer stability (e.g., Dai et al. 1998, Castro et al. 2001; Figure 24).

Table 3 provides entry links for websites displaying teleconnection patterns for surface air temperature, precipitation, and other station variables in terms of their means and extremes.

4.8.2 Testing for teleconnections – Hemispheric circulation indices

For these oscillations, circulation indices are used to represent their dynamics in a single timeseries. These portray the principal mode of variability at a center of action (e.g., SLP anomalies in the Aleutian Low pressure center, for the PDO-related North Pacific Index, NPI – Trenberth and Hurrell 1994; Table 3) or in the difference between dipoles (e.g., Tahiti-Darwin SLP difference for the Southern Oscillation Index, SOI – Figure 23b, *top*). Table 4 presents indices for oscillations with major teleconnections across North America along with download links. Take care that candidate index datasets for your analysis are not smoothed series because of problems they present in statistical evaluation (§4.5.1: Additional cautions).[133]

To explore teleconnection signals in local and regional climates, common techniques are:

Correlation approaches. These use linear correlation to evaluate the relationship between station records and a circulation index. Techniques include:

1. Simple linear correlation between a station timeseries and the index (§4.4.1).
2. Cross-correlation (§4.4.1) over, for example, weekly, monthly, or seasonal lags based on known or hypothesized mechanics (e.g., Barton and Ramirez 2004, Wright and Calderón 2006).[93]
3. Point-correlation maps (§4.7.1), with the index as the point timeseries (Figure 24a). Teleconnection maps, as on websites in Table 3, are commonly based on this method (e.g., Castro et al. 2001). NOAA provides an on-line facility[123] for generating index-correlation maps with built-in climate fields (demonstrated in Figure 24a).

Cross-spectra/Cross-wavelet. High spectral coherence (§4.7.3) between timeseries for a station and an index may indicate that local variability is dynamically linked to the corresponding hemispheric oscillation. The phase and frequency bands of this coherence may suggest or lend support to hypotheses for such a mechanism. Cross-wavelet analyses can also be applied with these same objectives (§4.7.3).

Data blocking.[134] Dividing an index timeseries into groups (blocks) of years can add power to statistical tests for teleconnections. This is usually done by oscillation phase – positive vs. negative phase years – or, in addition, by eliminating years with low or neutral signal (e.g.,

[133] Also see caveats regarding smoothing in timeseries comparisons (§4.4.2) and spectral analysis (§4.5.2).
[134] Blocking was introduced earlier, in the context of looking for interaction between slow vs. fast-moving processes: §4.4.2: *Process timescales differ, Mode 2*.

Figure 23b: *top*).[135] Blocking by phase focuses an analysis on regime state rather than intensity; blocking out near-neutral conditions focuses on strong episodes most likely to have detectable downstream effects (cf. next section, §4.8.3: *Forcing strength*).

Two-factor blocking, such as by the phases of two indices, allows for an assessment of their interaction (e.g., Figure 23b: *bottom*).[136] Alternatively, the second or yet additional factors can be any variable known or hypothesized to alter teleconnections (such as season, response region). Analysis of teleconnections with multifactor-blocked data can be by individual combinations of the blocks (e.g., the four combinations of PDO and AMO +/– phases in Figure 24b) or with all considered simultaneously, as in multifactor analysis of variance (MANOVA).[137]

Regime shift detection. A reflection of circulation regime shifts in timeseries of landscape climatic, hydrologic, and ecological variables can provide evidence of the impact of hemispheric climate processes on local dynamics (Hare and Mantua 2000; Figure 24c). Techniques for regime shift detection are referenced in §4.5.3. As with local climate data, check for temporal inconsistencies created in the processing of index timeseries you're evaluating (e.g., Table 4: footnote 155).

Principal component analysis. Generally speaking, principal component analysis (PCA) is used to describe the multivariate behavior of a system as a much reduced set of variables. For example, Hare and Mantua (2000) used PCA on an array of environmental variables to distill out principal components strongly related to PDO dynamics. In geophysics, PCA is commonly referred to as *empirical orthogonal function* (EOF) analysis, where it is applied to spatiotemporal data for one parameter (such as temperature anomalies).[138] This technique extracts the most prevailing patterns in space and their corresponding timeseries (e.g., Figure 23a *top* and *bottom*, respectively, for the first component of the PDO).[139] In interpreting EOF's, the first few functions (principal components) tend to explain a sufficient portion of the variance

---

[135] Note that delineation of an oscillation's phases can be an issue: there is no clear or consistent definition of what magnitude or duration of change constitutes a shift to the opposite phase and there is no single index or other definitive measure of these dynamics in terms of geography, variable, or season (hence multiple indices per oscillation in Table 4, Figure 23b; see also points re PDO phases in footnote 136).

[136] An example 2-way blocking by ENSO and PDO phases is given by JISAO (University of Washington): http://www.cses.washington.edu/cig/pnwc/compensopdo.shtml. Some points regarding their presentation:

- PDO phases are presented both in terms of multidecadal regime (columns) and annual state (cell entries).[128]
- Before 1925, there is reduced certainty in climate records used to determine PDO phase. From 1900-1924, the phase is variously considered as negative (in the JISAO table; also Mantua et al. 1997) or positive (e.g., Biondi et al. 2001 based on proxy records: http://www.ncdc.noaa.gov/paleo/pubs/biondi2001/biondi2001.html; Rodionov and Overland 2005: cf. Figure 16b).
- Since 1976/77, shortterm multiyear excursions to negative and back to positive territory have been suggested as possible regime changes (e.g., 1999, 2003, 2006 shifts in the JSIAO table, 1989 shift in Hare and Mantua 2000; see also Rodionov and Overland 2005: Figure 16b).[135] The perspective of additional decades is needed to judge whether these are true regime shifts or reflect year-to-year variation within the positive phase that started in 1977 (see Figure 23a, *bottom*).

[137] See Helsel and Hirsch (2002: §7.2.2)[7]

[138] EOF analysis can be implemented using eigenvalue decomposition (eigenanalysis) or singular value decomposition (SVD); see von Storch and Zwiers (2001: §13.2.9).

[139] EOF analysis terminology (with corresponding PCA terms):

- Empirical orthogonal functions represent the spatial pattern (principal component loadings)
- EOF coefficients express the temporal pattern (principal component scores)

to warrant exploring their scientific meaning; the remaining EOF's are usually only of minor significance.[140]

As a key method, EOF's reveal primary modes in teleconnections (e.g., Wallace and Gutzler 1981). A bivariate technique related to EOF analysis is *canonical correlation analysis* (CCA); this approach extracts spatiotemporal patterns common to two spatially and temporally distributed parameters. von Storch and Zwiers (2001: Chapters 13, 14) present these pattern detection techniques.

4.8.3 Properties of circulation oscillations and their indices – Insights and pitfalls

Circulation indices offer us the opportunity to understand linkages from the global to local. Exploring this scale translation requires insights into the nature of circulation oscillations and their indices. These insights can help design analyses, seek mechanisms in interpretation, and avoid pitfalls.

Key features of oscillations and indices are:

Season dependence. Dynamics at centers of action fluctuate seasonally, and their downstream connections follow suit. As a result, seasonal indices tend to best capture teleconnections. Annual indices often blur the controlling signal, while individual monthly values divide seasons up arbitrarily, weakening detection of a seasonal effect.[141,142] The key teleconnective season (and how it is defined) may depend on the site climate variable evaluated, such as for temperature vs. precipitation.[143] In the extratropics, the strongest signal-to-noise ratio is commonly in winter, giving the most robust teleconnections.[144] However, it can be more fruitful to take the season that corresponds to timing of key local dynamics, such as a summer index for connections with the summer monsoon (e.g., Castro et al. 2001). On the other hand, between-season interactions might alternatively suggest evaluating an index from a different season – for example, if summer moisture conditions are more a function of winter snowpack and its melt regime, then a winter or spring index may be more appropriate than a summer one.

Forcing strength makes for a teleconnection. Strong events are most likely to propagate downstream and result in signals relevant to local dynamics. Weak signals generally get dissipated en route and swamped locally by other sources of variability. As a result, teleconnections are most evident when analyses contrast the strongest periods of opposite phases (e.g., by blocking out near-neutral periods, discussed in §4.8.2: *Data blocking*) (Figure 23b, *top*).

---

[140] Note that in such a spatiotemporal analysis, high percent-variance explained (in the first EOF's) ideally comes from the temporal dynamics of the entire spatial field being explained moderately well. However, note that relatively high variance explained (often in subsequent EOF's) can also come from dynamics of a restricted part of the domain being explained extremely well – this describes how one area is behaving but does little to show how different regions are connected.

[141] This is due to within-seasonal variability either (1) because a seasonal effect may be more concentrated in one month one year but in an adjacent month in the next occurrence or (2) because month boundaries are arbitrary – what's taken to be a month's difference in timing may only be a matter of a few days.

[142] Rather than using a standard scheme to define seasons (e.g., winter = December-January-February), it can be more powerful to delineate seasons based on breaks in system dynamics. For example, Trenberth and Hurrell (1994) define a winter NPI covering November through March (http://www.cgd.ucar.edu/cas/jhurrell/npindex.html).

[143] Plotting monthly teleconnection results can suggest an optimal delineation of key seasons and if that differs for different local variables (e.g., http://cses.washington.edu/cig/pnwc/clvariability.shtml#figure5).

[144] For an example from the Pacific Northwest, see http://cses.washington.edu/cig/pnwc/clvariability.shtml#figure1.

Basin interactions. Circulation oscillations arising in different ocean basins or in different sectors of a basin interact, such as between the PDO and AMO (McCabe et al. 2004) and ENSO and PDO (Gershunov and Barnett 1998), respectively. This leads to conditional teleconnections, where the phase of one oscillation influences the downstream expression of another (e.g., Figure 24b). Two-factor blocking and related multifactor techniques are approaches for evaluating interactions (cf. §4.8.2: *Data blocking*).

Circulation indices are broadly integrative. Oscillation indices track climate dynamics of broad regions of the globe that have strong spatiotemporal coherence across many variables (Figure 23a). Stenseth and Mysterud (2005) layout a conceptual framework for how these hemispheric, seasonal, and multivariate indices present an integrated view of climate that sets up local conditions for a season or longer. These establish prevailing seasonal conditions in ways that can have as much power in explaining ecological dynamics as do analyses of local weather (e.g., Hallett et al. 2004, Forchhammer and Post 2004).[145]

Additional considerations in scale linkages to landscapes and species are:

Local conditionality. How forcing from a given teleconnection plays out across landscapes and regions can be conditional on physiographic features such as aspect, altitude, and latitude. This is especially the case if teleconnections affect where a critical weather threshold, such as a storm's snowline, crosses the domain (Stenseth and Mysterud 2005).

Indirect ecological effects. In evaluating hemispheric linkages to ecological dynamics (while skipping over local climate), keep in mind that some consequences of circulation teleconnections may be indirect, possibly with strong temporal lags and spatial offsets. These may arise from population and trophic dynamics and from biogeographic linkages (e.g., for regional or hemispheric migrants[146]) (Forchhammer and Post 2004).

Important caveats and common pitfalls in teleconnection analyses include:[147]

Responses nonlinear. Within a given oscillation's phase, teleconnections are not expressed the same way in each occurrence. The relationship between an oscillation's phase and its local teleconnections may in fact change sign between lower forcings vs. higher ones. This is because oscillation dynamics often shift the position of an atmospheric circulation system (e.g., storm track latitudes) into a region at first and, subsequently under a more intense teleconnection, push the circulation farther along but out of the region. Such non-monotonic, complex dynamics are of course not adequately explored with linear correlation methods. Careful period blocking can help reveal these dynamics.

Responses nonstationary. Teleconnections also appear to change with time. Some nonstationarity can be attributed to behavior conditional on the phase of other oscillation systems, as noted earlier. Aside from such conditionality, keep in mind that key teleconnections

[145] *N.B.* – Descriptions of circulation patterns in Forchhammer and Post (2004) are not entirely accurate; Stenseth et al. (2003) provide a better review for ecological audiences. Forchhammer and Post (2004) do, however, present key insights from three case studies on teleconnections and ecological dynamics.

[146] For example, when migrants have distant seasonal ranges strongly affected by a teleconnection seen there, but not in the local study domain.

[147] Stenseth et al. (2003) also review benefits and drawbacks of teleconnection analyses.

found may not have held throughout the historical period nor persist in the future (e.g., Gedalof et al. 2002).

Prediction. Much attention has been given to the prospect of predicting local climate and ecological dynamics based on teleconnections. While teleconnections with local dynamics may be statistically significant, the percent variance explained may put into question the utility of relationships for prediction.[148] Low explained variance comes from various sources:

- Nonlinearity, nonstationarity, and indirect effects of teleconnections that reduce their detection.
- Centers of action for oscillations are much removed and their signal is altered in transit by unaccounted-for downstream climate processes – including stochastic and chaotic behavior inherent to the climate system.
- Hemispheric circulation dynamics are only part of the local story – Other, independent local factors also control site climate and ecological behavior.

4.9 Interpretation of results – The good, the bad, and the ugly

4.9.1Statistical interpretation

On completing analyses, comes interpretation of descriptive and statistical results and their scientific review (next section, §4.9.2) (Figure 1c). Recall your research questions and corresponding stated hypotheses – in doing so, confirm that your analyses are addressing what you intended. For statistical tests, review results and their significance level, and formally state outcomes (e.g., rejecting null hypotheses or not).

Scientific interpretation of statistical results

Caution is needed in interpreting both significant and nonsignificant statistical results:

Significant results. Statistically-significant results lend support to your hypothesized dynamics, but not to validation. In tests comparing variables, keep in mind, as noted earlier, the proposition that "correlation does not imply causation" (§4.4.2: *Interpretation of correlation*). There may be other underlying mechanisms that give rise to the relationships you see. In such bivariate as well as spatial and temporal relationships, significant patterns may turn out be nonstationary – that is, only present under conditions set by some undetected overriding process (e.g., a seemingly persistent circulation regime). Given a longer timeframe, relationships seen today may shift or disappear.[149] The role of undetected, unevaluated factors can interfere with seeking mechanisms in the interpretation of results.

Weak results. Nonsignificant results also need to be evaluated with care. Weak results (tendencies consistent with a hypothesis, but not backed by statistical significance) should not be reported as apparent support for your hypothesis. Rather they support re-evaluation of:

- Available data – Are additional, appropriate data sources available that would lengthen a station's record, or allow evaluation across several locations? The advantage of this is that the power of statistical tests increases with number of independent observations.

---

[148] The distinction here is between significance level (e.g., p<0.05) and % variance explained, e.g., R^2 for regression analysis (§3.4.3.2). It is not unusual for teleconnection analysis to yield highly significant regressions, say p<0.01, but with it low to modest % variance explained, commonly with R^2's<0.30.

[149] See Schumm (1991: Chapter 3) for pitfalls in interpreting results in space and time.

- Data processing – Review if dataset development methods adequately corrected problems or, instead, obscured patterns being testing for.
- Analysis methodology – Some techniques are more powerful than others in detecting patterns. Review analysis options (data transformations, alternate tests, etc.) and check that their assumptions are followed.

Clearly nonsignificant results. From a statistical perspective, clearly nonsignificant results obviously do not support your hypothesis. From a scientific perspective, this can be the end of story or lead to alternate hypotheses. However, another proposition that "correlation is needed to prove causation" is not necessarily the case. We may just not be looking at the question from the right perspective or with the right tools: system complexity can mask causation, making it recalcitrant to standard approaches. Climate processes and affected biotic components (1) are highly interactive with positive and negative feedbacks, (2) act across a range of time and space scales, and (3) often entail non-monotonic or threshold responses. Numerical simulation models are tools employed to understand dynamics of such highly connected systems.

5.0 Synopsis

The objectives of this report are two-fold. First, to layout a methodology for developing quality climate datasets appropriate for resource management science. Second, to introduce the array of analysis techniques available for answering the many questions we often ask regarding climate dynamics and their interaction with a region's ecology and other landscape processes.

Overarching guidelines for creating datasets are:

(1) *Hypotheses dictate the requirements* of datasets, selection of data clean-up methodologies, and corresponding analyses. Keeping hypotheses in mind throughout the process will help ensure a successful outcome: a credible dataset and valid results.

(2) *No dataset is perfect.* There's a trade off between working with a clean highly-processed data set and one that is as unadulterated as possible.

   a. *No raw observational dataset is free of collection and archive errors*. Decisions on what types of errors to be concerned about and to correct should be based on goals and analysis requirements. Spending effort on ridding a dataset of all problems may not be necessarily for intended uses, and may create a set that is not appropriate for addressing some questions. Keep in mind that some adjustment techniques may interfere with your intended analysis.

   b. *No cleaned-up dataset is free of assumptions* about what its planned or perceived use is. Keep track of decisions made along the way and limitations arising as a consequence of these decisions and processing. Test generated data for unintended emergent features, and that these aren't going to create spurious results in the analysis stage.

(3) *Nonetheless, techniques for improving dataset utility can detect and correct* data errors, biases, and artificial record inhomogeneities and infill missing observations.

(4) *Document achieved improvements, intended uses, and caveats* for yourself and other users to understand what the dataset is good vs. inappropriate for.

Strategies for data analysis are:

(5) *Analyses must include statistical significance testing* to have results that can be relied upon for understanding your system and for decision-making. This is in addition to analyses being well designed and properly interpreted. Descriptive and graphic techniques are valuable for exploring data, but ultimately these must lead to hypothesis generation and testing.

(6) *Statistical techniques make certain assumptions about input data.* For a given analysis, some methods have more relaxed requirements than others regarding statistical distribution, independence, and missing values, for example. Keep track of the properties of your data with respect to analysis requirements – violating these will most likely give overestimated significance to your results.

(7) *Analytical tools can evaluate temporal and spatial patterns at different scales*, including trends, oscillations, regime shifts, daily events, regional correlations, and hemispheric teleconnections.

(8) *Keep in mind that results contain uncertainties not assessed statistically.* Uncertainties external to an analysis will mean we're likely to assign too much confidence to statistical results. Sources of uncertainty include unresolved data issues, analysis limitations, and simplifying assumptions re natural systems that we use to guide our studies.

Lastly:

(9) *Studies are most likely to be successful if driven by a sense for underlying mechanisms* that connect ecosystems and species with climate. This will provide more confidence in interpretation of analysis results beyond just their statistical basis, adding insights into the dynamics of park landscapes.

Tables

Table 1. An example of a plausibility limits table that can be used in basic validity checks (after Burroughs 2008). Limits are set to catch values not physically plausible; these are tailored for a given site (§3.3.1). This is separate from screening for outliers which would employ tighter limits (§3.3.2).

Item	Valid Range
Year	1954 - present
Month	1 - 12
Day	1 - last day in corresponding month
Observation Hour	0 - 23
Temperature	-50 to +50 °C
Precipitation	0 to 100 mm/day
Dewpoint Depression	0 to 50 °C
Wind Speed	0 - 100 m/s
Wind Direction	0 - 360°
etc.	

Table 2. On-line resources for station metadata (indicated by an asterisk*; §3.4.1) and monitoring and outlook products providing near-real time, regional context to park climates (§4.7).

Source	Web Entry Point
*NOAA Regional Climate Centers	http://www.wrcc.dri.edu/rcc.html - links to all NOAA regional centers ▪ e.g., for Western U.S. – http://www.wrcc.dri.edu/CLIMATEDATA.html
*State Climatologists	http://www.stateclimate.org/ ▪ e.g., for Wyoming http://www.wrds.uwyo.edu/sco/climate_office.html
NOAA Climate Prediction Center	http://www.cpc.ncep.noaa.gov/ – Climate monitoring: U.S., Pacific Islands, Global: ▪ http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/ (maps & data) ▪ http://www.cpc.ncep.noaa.gov/products/precip/CWlink/ (including near-real time and historical maps)
NOAA National Climate Data Center	http://www.ncdc.noaa.gov/ ▪ http://www.ncdc.noaa.gov/climate-monitoring/ – NCDC climate monitoring links ▪ http://lwf.ncdc.noaa.gov/oa/climate/research/cag3/cag3.html – *US Climate at a Glance* (data) ▪ http://www.ncdc.noaa.gov/sotc/ – *State of the Climate:* US, Global (reports)
National Integrated Drought Information System	http://www.drought.gov/ ▪ National Drought Mitigation Center – http://drought.unl.edu/
Natural Resource Conservation Service	http://www.wcc.nrcs.usda.gov/snowcourse/ – Snow course maps
Western Water Assessment	http://wwa.colorado.edu/IWCS/index.html - Intermountain West climate summary ▪ http://wwa.colorado.edu/forecasts_and_outlooks/forecasts.html - Links to other US climate-related websites

Table 3. Web entry-point resources for major teleconnection patterns for United States and territories. Sites give descriptions and maps for teleconnection patterns for annual and monthly surface climate variables; some sites also include teleconnections for extremes. Oscillation abbreviations are given in the text §4.8.1. Some patterns are illustrated in Figure 24 (for NPI, ENSO, PDO, and AMO). Data sources for oscillation indices are given in Table 4. A broader summary of Northern Hemisphere teleconnections is provided at: http://www.cpc.noaa.gov/data/teledoc/telecontents.shtml

Teleconnection Pattern*†	Web Resources
ENSO teleconnections	▪ http://www.cru.uea.ac.uk/cru/info/enso/ ▪ http://www.esrl.noaa.gov/psd/enso// – o http://www.esrl.noaa.gov/psd/enso//enso.climate.html o http://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/composites/ U.S. by climate region and state – o http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/ENSO_connections.shtml U.S.: for climate extremes – o http://www.esrl.noaa.gov/psd/enso/climaterisks/ http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/lanina/us_impacts/ustp_impacts.shtml ▪ http://www.cses.washington.edu/cig/pnwc/aboutenso.shtml o See *PDO teleconnections* for 'combined PDO & ENSO effects' ▪ http://jisao.washington.edu/analyses0500/#enso
PDO teleconnections	▪ http://www.atmos.washington.edu/~mantua/REPORTS/PDO/PDO_cs.htm; also Mantua et al. (1997) ▪ http://www.cses.washington.edu/cig/pnwc/aboutpdo.shtml o Combined PDO & ENSO effects – ▪ http://www.cses.washington.edu/cig/pnwc/clvariability.shtml ▪ http://www.cses.washington.edu/cig/pnwc/compensopdo.shtml ▪ http://jisao.washington.edu/analyses0500/#pdo ▪ http://www.beringclimate.noaa.gov/data/BCinclude.php?filename=in_PDO ▪ See *AMO teleconnections* for 'combined AMO & PDO effects'
& the related North Pacific Index (NPI)–Pacific/North American (PNA) teleconnections	▪ NPI: see Trenberth and Hurrell (1994), Hurrell (1996) ▪ http://jisao.washington.edu/data/pna/ ▪ http://jisao.washington.edu/analyses0500/#pna
NAM/AO & NAO teleconnections	▪ http://jisao.washington.edu/analyses0500/#ao ▪ http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml ▪ http://nsidc.org/arcticmet/patterns/arctic_oscillation.html ▪ http://www.ldeo.columbia.edu/res/pi/NAO/ ▪ http://www.cru.uea.ac.uk/cru/info/nao/
AMO teleconnections	▪ http://oceanworld.tamu.edu/resources/oceanography-book/oceananddrought.html o Includes combined AMO & PDO effects; also McCabe et al. (2004) ▪ http://www.aoml.noaa.gov/phod/amo_faq.php
Other Major Teleconnection Patterns – East Atlantic (EA) West Pacific (WP) East Pacific-North Pacific (EP-NP) Tropical/Northern Hemisphere (TNH) Pacific Transition (PT)	▪ http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml

* For creating custom teleconnection correlation maps, see: http://www.esrl.noaa.gov/psd/data/correlation/.[123]
† For a glossary of terms related to circulation oscillations and teleconnections, see http://www.ucar.edu/news/backgrounders/patterns.shtml

Table 4. Hemispheric circulation oscillations indices and their online data sources for oscillations with significant teleconnections affecting the United States and territories. Related oscillations are listed together. The analysis variable that underlies each index is given in [brackets]. Several of the indices are plotted in Figure 23 (PDO, SOI, and MEI). Links describing their US teleconnection patterns are listed in Table 3. A complementary compilation of index sources is at: http://www.esrl.noaa.gov/psd/data/climateindices/list/. (Link to text §4.8.1 and 4.8.2)

Oscillation	Circulation Indices [Analysis Variable]*	Sources for Index Timeseries**
El Niño / Southern Oscillation (ENSO)	Niño Region 3.4 SST Index / Oceanic Niño Index (ONI)[150] [SST]	http://www.cgd.ucar.edu/cas/catalog/climind/Nino_3_3.4_indices.html http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml ▪ http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml
	Southern Oscillation Index (SOI) [SLP]	http://www.cgd.ucar.edu/cas/catalog/climind/soi.html http://www.cru.uea.ac.uk/cru/data/soi.htm
	Multivariate ENSO Index (MEI) [multivariate]*	http://www.esrl.noaa.gov/psd/people/klaus.wolter/MEI/mei.html ▪ http://www.esrl.noaa.gov/psd/people/klaus.wolter/MEI/table.html
Pacific Decadal Oscillation (PDO)[131]	PDO Index [SST]	http://jisao.washington.edu/pdo/ ▪ http://jisao.washington.edu/pdo/PDO.latest [152]
	North Pacific Index (NPI)[151] [SLP]	http://www.cgd.ucar.edu/cas/jhurrell/npindex.html
North Atlantic Oscillation (NAO)[132]	NAO Index [SLP]	http://www.cgd.ucar.edu/cas/jhurrell/indices.html [154] http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml http://www.cru.uea.ac.uk/~timo/projpages/nao_update.htm
Northern Annular Mode (NAM) / Arctic Oscillation (AO)[153]	NAM Index [SLP]	http://www.cgd.ucar.edu/cas/jhurrell/indices.info.html#nam
	AO Index [1000mb ht]	http://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml
Atlantic Multidecadal Oscillation (AMO)	AMO Index [SST]	http://www.esrl.noaa.gov/psd/data/timeseries/AMO/

* SST = sea surface temperatures, SLP = sea level pressure, ht = pressure surface height. MEI is based on SLP, SST, surface wind, surface air temperature, and cloudiness anomalies.
** Where more than one source is listed, it is generally because they use slightly different means to determine the index. Note that some sites present both monthly and seasonal series for some indices. I've attempted to list sites that keep index series up-to-date, but no guarantee.

[150] Trenberth (1997) explains the difference between these similar indices based on Niño Region 3.4 SST's: http://www.cgd.ucar.edu/cas/catalog/climind/. SST's in this region of the equatorial Pacific is a strong indicator of ENSO events, especially with respect to global teleconnections: http://www.ucar.edu/news/backgrounders/patterns.shtml#mno
[151] The North Pacific Index (NPI) reflects the PDO – not the second NPO pattern discussed in footnote 127 (Minobe and Mantua 1999). Note that the NPI is negatively correlated with the PDO index.
[152] SST source data for this PDO Index changed in 1982 and 2002 – see data source notes (and graphic link at end of file).
[153] See discussion of Arctic/Antarctic Annular Modes at http://ao.atmos.colostate.edu/introduction.html and their indices at http://ao.atmos.colostate.edu/Data/index.html
[154] Hurrell lists an array of NAO indices, for a overview see: http://www.cgd.ucar.edu/cas/jhurrell/naointro.html

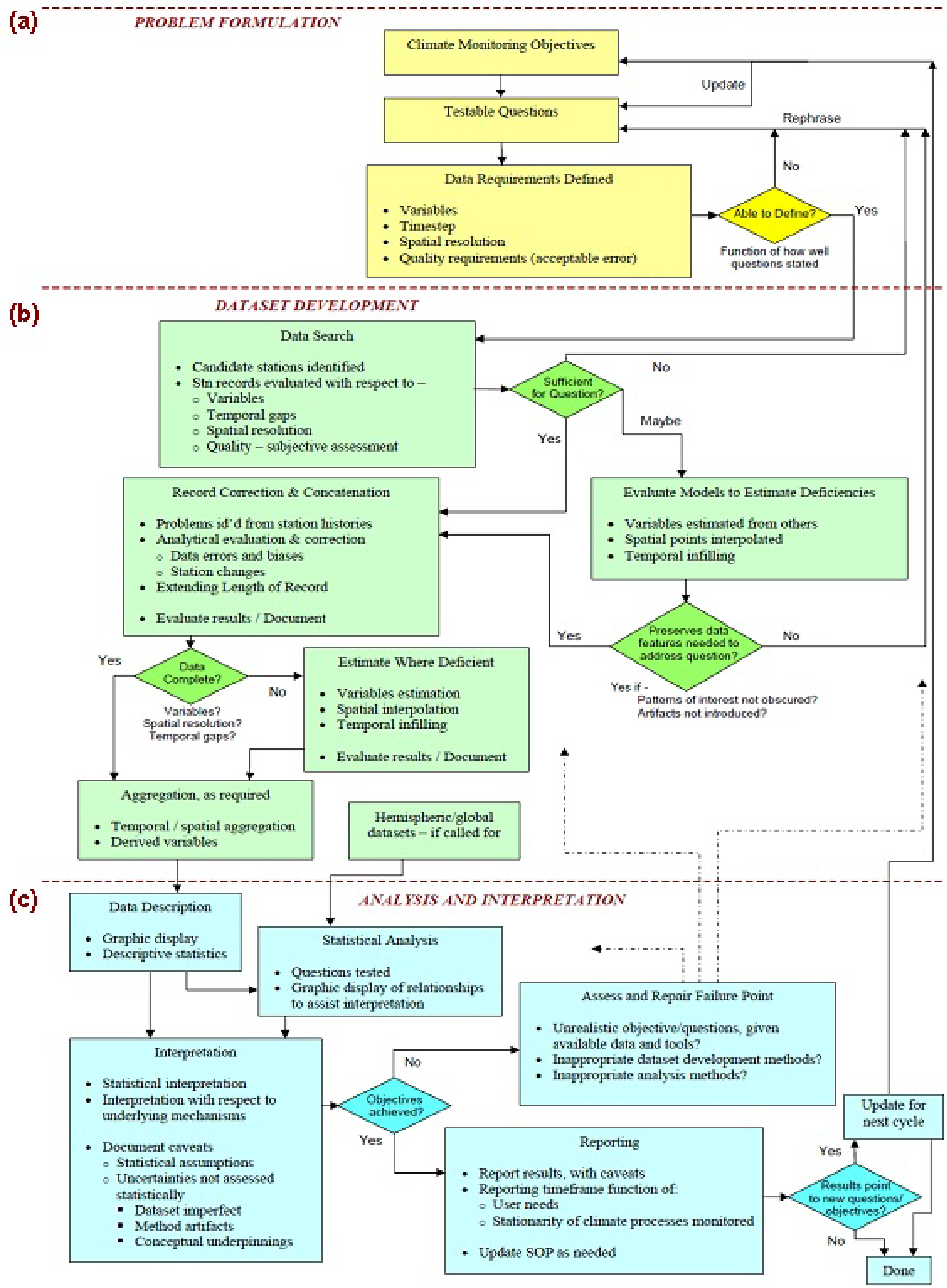


Figure 1. Workflow diagram laying out the overall process for developing and analyzing climate datasets. Three key stages are: (a) problem formulation (covered in section §2.0), (b) dataset development (§3.0), and (c) analysis and results interpretation (§4.0). Specific objectives and questions will dictate specific procedures for each stage. Not all boxes are covered in this report, such as variable estimation ('Evaluate Models' box), reporting, and program updating processes. (SOP = standard operating procedure document.) (Link to text §1.2)

Copyrighted figure removed

Figure 2. The relationship between characteristic temporal and spatial scales for (a) various geophysical processes in general,[155] (b) atmospheric processes in particular (from von Storch and Zwiers 2001), and (c,d) terrestrial ecological processes.[156] The blue oval in (a) roughly gives an alternate scheme for oceans from von Storch and Zwiers's (2001); the green oval in (b) generally represents that for the terrestrial biosphere in (c). Panel (d) gives a specific example for the behavior of large wading birds [the domain of this figure is indicated by the brown box in (c)]. Note that temporal and spatial axes are log10, and those in (b) are switched from those in (a,c,d). (in a-d, color annotations added)

Discussion. Regardless of the geophysical or ecological system, these figures show a positive log-log relationship between characteristic temporal and spatial scales for biogeophysical processes: larger processes take longer to operate. How this relationship lays out is, however, system dependent, with little or no overlap in the time-space relationship for climate and ecological processes (b). This means there is not a simple translation of variation in one system to the other at a given scale, but rather that there must be scale interactions. Differences among time-space relationships are illustrated in detail for ecological external forcings vs. internal processes by Urban et al. (1987; see also Delcourt et al. 1983). (Link to text §2.2)

Copyrighted figure removed

Figure 3. Types of errors often present in data from electronic sensors and other automated instruments (from Wade 1987).

Discussion: When errors carry the data outside of an expected dynamic range, they can be detected by straightforward screening techniques (e.g., plausibility tests, §3.3.1). Detection and correction of more subtle errors (with physically reasonable values), require more involved techniques (§3.3.2–3.3.3).

[155] Image from Water Cycle Study Group, 2001, Predictability of Variations in Global and Regional Water Cycles, Ch. 3, in: *A Plan for a New Science Initiative on the Global Water Cycle*. USGCRP. (http://www.usgcrp.gov/usgcrp/Library/watercycle/wcsgreport2001/wcsg2001chapter3.htm)

[156] Image from Sheehan, P., 1995, Assessments of Ecological Impacts on a Regional Scale, Ch. 14 in: *SCOPE 53 - Methods to Assess the Effects of Chemicals On Ecosystems* (http://www.icsu-scope.org/downloadpubs/scope53/chapter14.html)

Copyrighted figure removed

Figure 4. Illustrations of record outliers and different detection methods. (a) Ice coverage for a lake in Ontario (from Drews 2003). The circled outlier was manually detected. While physically plausible, the outlier was rejected when a manual review of observations found that the record for the suspect year was incomplete. (b) Tropospheric temperature record over Norway in 1984. The outliers (spikes) passed physical plausibility tests but were detected and rejected using a temporal consistency test (from Burroughs 2008). (c) Daily mean surface air temperature for Walnut Creek, CA. Automated screening used day-dependent outlier detection limits (curves) based on record absolute daily maxima and minima – this identified a late fall outlier (indicated by the solid arrow). However, the method can be overly sensitive as it does not allow any leeway for new valid extremes. Follow-on visual inspection showed that the outlier was temporally consistent with a week-long cold snap (dashed arrows show the trace of dailies leading down to and then up from the outlier) and so was not rejected (from Meek and Hatfield 1994; brown annotations added). (d) The 1932 monthly temperature record for Linyi, China (dots) was automatically screened first with climatological limits based on 2.5 and 5 standard deviations (SD) from the longterm mean seasonal cycle (black curve; gray-shaded contours are 1 SD increments from the average).[157] Depending on which SD threshold was exceeded, outliers were further evaluated for spatial consistency against observations from nearby stations and for temporal consistency (Hansen et al. 1999). Finally, flagged outliers were plotted for visual checks – if the datapoint was not clearly faulty, it was retained. In this case, the September 1932 mean was identified as outlier (> 2.5 SD), but was confirmed as reasonable value by other stations' records and was retained. (Link to text §3.3.2)

Copyrighted figure removed

Figure 5. Time-of-observation bias. Difference in March mean monthly temperature between daily minimum/maximum thermometer observations taken in the early morning vs. end of the afternoon (0700h vs. 1700h local time). (from Karl et al. 1986)

Discussion: Temperature bias due to time of observation is strongest for continental regions and in spring (as shown here) and late fall. This geographic and seasonal pattern of high biases is related to the prevalence in the interior and in spring and fall of strong warm and cold fronts, which are linked to this bias (see text §3.3.3). Morning-based records give monthly means consistently colder than afternoon ones by as much as 2ºC or greater. Over the last 50 years, U.S. reporting times have steadily shifted from late afternoon to early morning, contributing an artificial cooling tendency to uncorrected regional trends (Vose et al. 2003).

[157] Image from: Herring, D. 2007. *Earth's Temperature Tracker*. Earth Observatory, NASA. http://earthobservatory.nasa.gov/Study/GISSTemperature/giss_temperature3.html (brown annotations added)

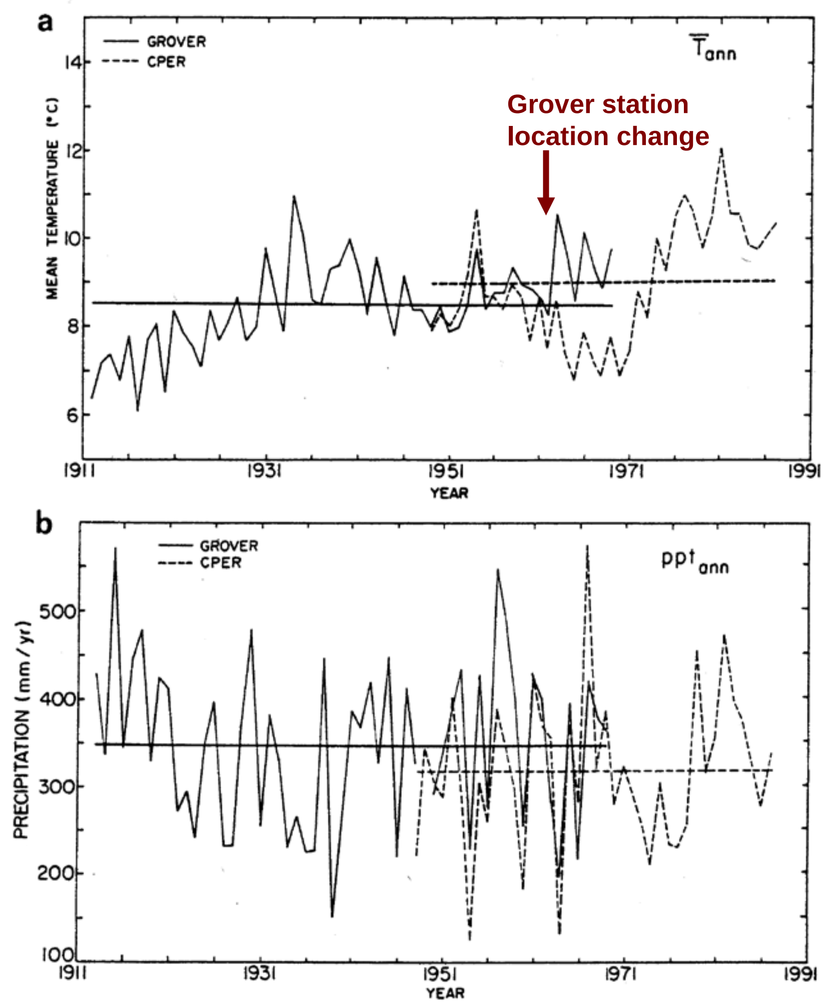


Figure 6. Comparing records of nearby stations, illustrating issues with combining records. Water-year (Oct-Sept) (a) mean temperature and (b) precipitation for two stations in the Colorado shortgrass steppe: Grover, CO and the Central Plains Experimental Range (CPER) LTER (Long-Term Ecological Research Program) site. (Kittel 1990; in a, brown annotation added)

Discussion: The objective was to create a combined record for exploring longterm climate dynamics of the region. Three issues must be addressed: (1) station differences in means, (2) station differences in variance structure, and (3) degree of correlation in interannual variability. For the period of overlap, the two sites track each other well until 1962 (correlation r = +0.71 for temperature, +0.77 for precipitation, p<0.01). This correlation breaks down after 1962 (arrow) especially for temperature (entire overlap period 1949-69: r = +0.14, n.s.). A check of the metadata for Grover revealed that the station was moved a significant distance in 1962. A truncated overlap period (prior to the Grover station change, 1949-1962) was then used to evaluate station differences in annual means and standard deviations, which were not grossly different. The long overlap permitted spectral analysis, which also revealed that variance structure was similar at sub- and multidecadal scales. These results confirmed that the annual records could be combined (concatenated) by adjusting records based on differences in means and ratios of standard deviations. (Link to text §3.5).

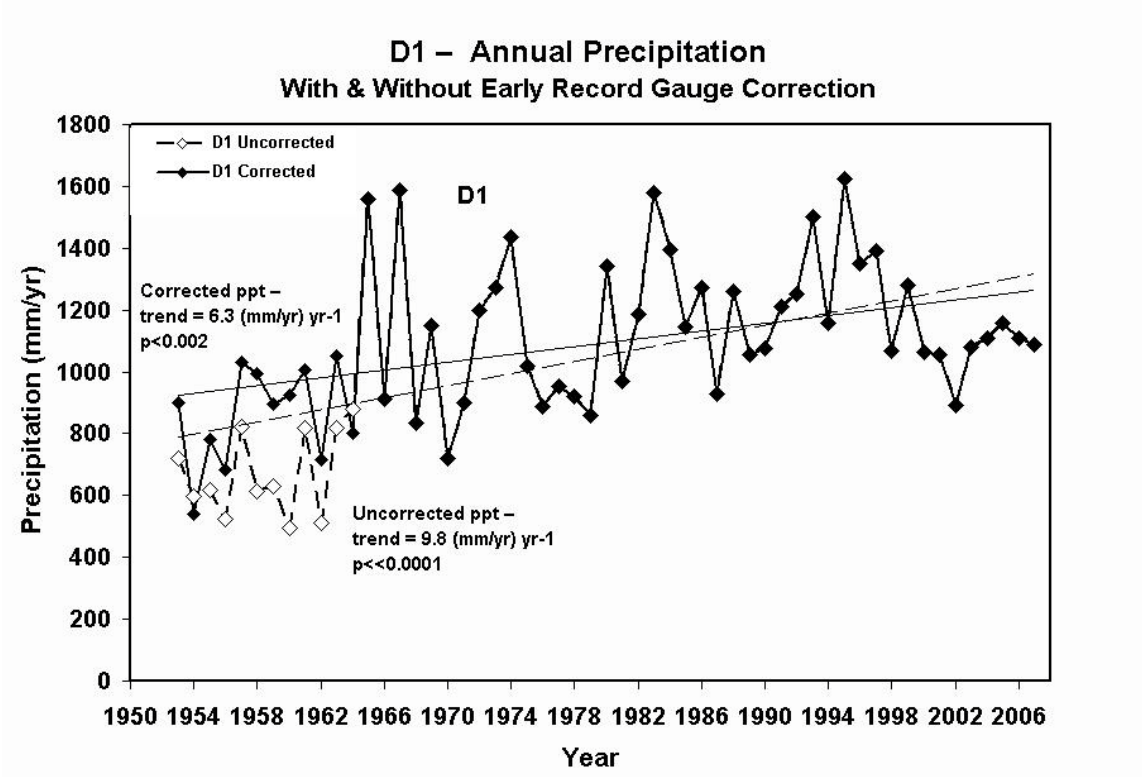


Figure 7. Accounting for a precipitation gauge change at a high elevation station in the Colorado Front Range. A unshielded Belford gauge was replaced at the end of 1964 with a recording bucket gauge along with an Alter-type shield and a Wyoming fence to prevent wind-caused undercatch. A 2-year overlap in instruments permitted the development of a correction factor, applied to the original data (*dashed line*) and reflected in the solid line (through 1964). (Kittel et al., in preparation)

Discussion: The correction had a marked effect on the longterm trend analysis: with correction, the trend decreased from +9.8 mm yr^{-1} yr^{-1} ($p<10^{-5}$) to +6.3 mm yr^{-1} yr^{-1} ($p<0.002$). (Link to text §3.4)

Copyrighted figure removed

Figure 8. Step changes in (a) maximum and (b) minimum temperature records in the mid-1980's due to a change from liquid-in-glass thermometers to the electronic sensor Maximum/Minimum Temperature System (MMTS) (from Quayle et al. 1991). Plotted values are the average timeseries across stations in the conterminous U.S., co-registered to the time of the change over.

Discussion: Opposing step changes in T_{min} and T_{max} resulted in a severely reduced diurnal temperature range after the switch. Such instrument artifacts severely interfere with climate trend assessments. Quayle et al. (1991) presents bias correction factors to adjust records for this change. (Link to text §3.4.5)

Copyrighted figure removed

Figure 9. Urban heat island effect: Evolution of differences in annual minimum temperature between stations in cities (1980 populations ≥100,000) relative to paired rural stations (populations <2,000) (N= number of pairs). (from Karl et al. 1988).

Discussion: The magnitude of the urban heating trend is substantial (~0.5ºC/80y), especially when compared to non-urban influenced global trends. The urban effect is real, but is often removed from regional datasets designed to evaluate trends arising from other sources (e.g., the USHCN dataset[18,60]). (Link to text §3.4.7)

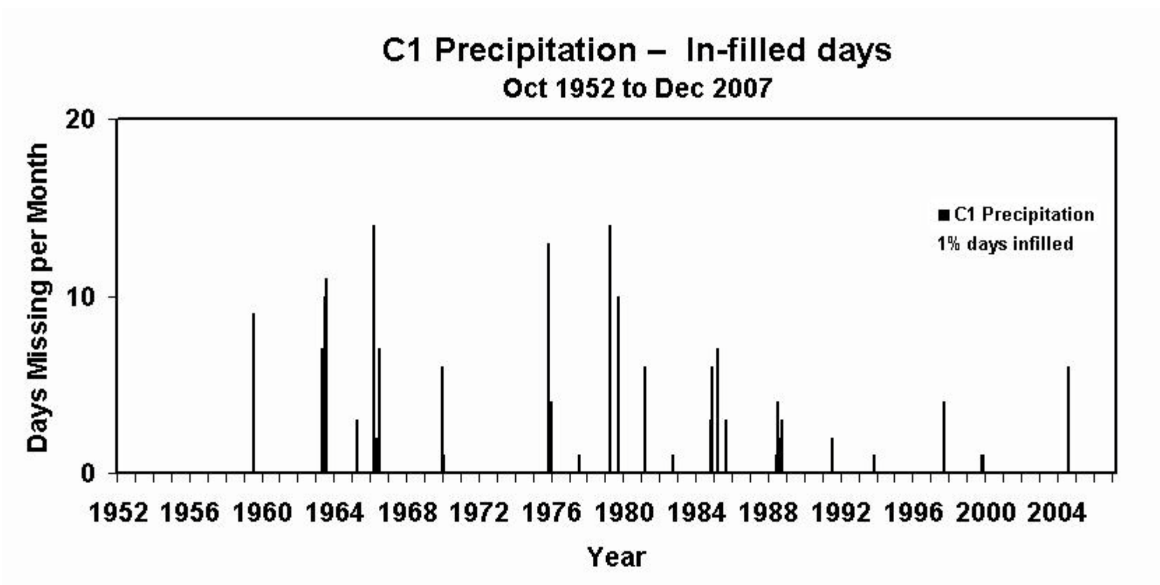


Figure 10. Number of days missing per month in the precipitation record of a high elevation station in the Colorado Front Range (Niwot Ridge). (Kittel et al., in preparation)

Discussion: The plot shows that the record is nearly complete with missing values well distributed – in particular, not concentrated at either end of the record which otherwise would raise the question of temporal biasing (§3.3.1). Missing values were infilled using adjacent instrument and nearby station records (§3.6).

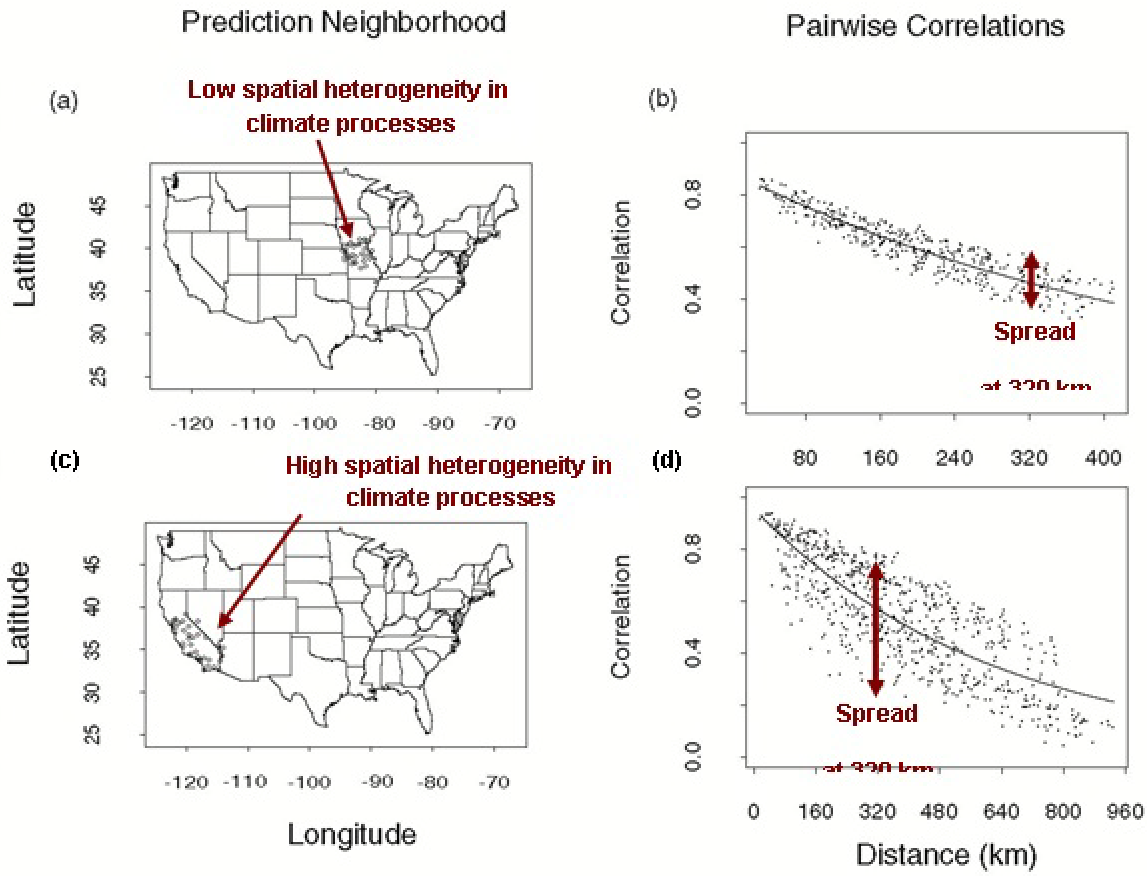


Figure 11. Spatial autocorrelation structure in monthly precipitation anomalies is dependent on heterogeneity in factors controlling climate, demonstrated here for two climate regions: (a, b) in the Midwest where autocorrelation structure was strong and (c, d) in southern California where it was weak. (a, c) Maps show a neighborhood ('local window') of 20 stations used to develop site correlograms shown in (b, d). In (b, d), note that *x*-axis distance scales differ between these two figures (vertical double-headed arrows are placed at *x*=320 km to facilitate comparison). Site correlograms show correlations among all station pairs as a function of distance (dots), along with the best fit line from an exponential model. (from Kittel et al. 2004; brown annotations added). (Link to text: §3.6.4)

Discussion: The double arrows show that the spread of correlations at a distance of 320 km was greater for a more climatically heterogeneous California than for the more uniform Midwest.

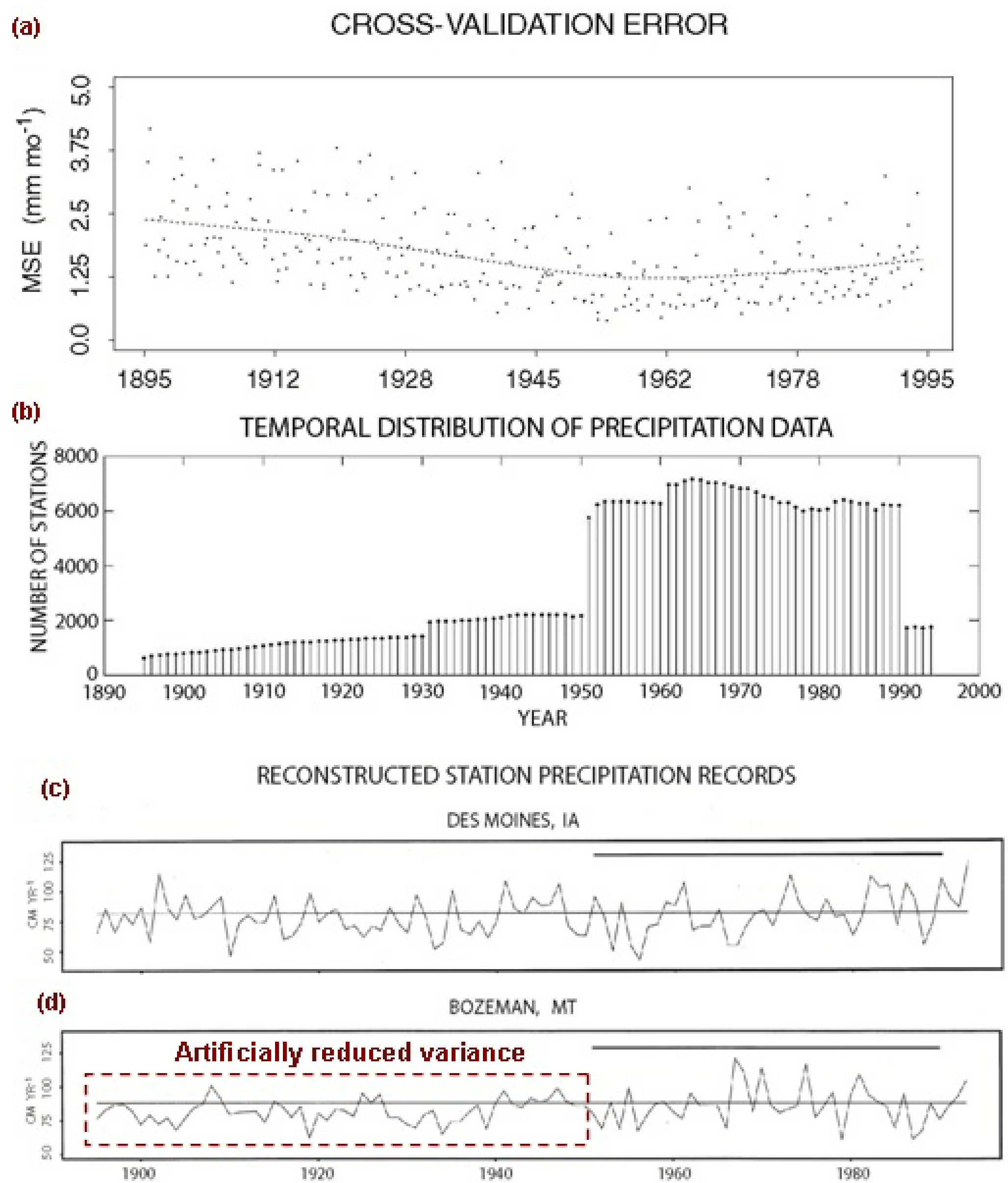


Figure 12. Time and station-density dependence in errors for the spatial (kriging) model in Figure 11. (a) Time dependence in cross-validation errors for precipitation, expressed as mean squared error (MSE) for square-root transformed precipitation, is higher during periods of reduced station density in (b). Cross-validation is where the model predicts a subset of observations withheld from the original analysis. The decrease in station number after 1990 was due to delayed updating of some national datasets. (c,d) Visualization of the effects of station density on interannual variability of infilled station precipitation records created using the kriging model, for different climatic regions: (c) Midwest (Des Moines) and (d) the Mountain West (Bozeman). The heavy horizontal bar in the upper right spans the period when observed data were available for the station. (from Kittel et al. 2004; in d, brown annotation added)

Discussion: Note in (a) that the model is relatively robust to station density change: average precipitation errors roughly doubled from recent decades back to the early part of the record (MSE = 1.3 to 2.5 mm/mo), corresponding to a nearly 10-fold decrease in station numbers across the domain (in b). In (d), interannual variability was artificially reduced during the early part of reconstructed records for areas where station densities were depleted, as around Bozeman. This is a consequence of the model generating overly smoothed spatial fields when there were too few station observations to adequately capture regional variability. As station density drops, the process reached farther away from a site to find predictor stations and blended unrelated anomaly patterns from adjacent regions to make a point's prediction. Poorly-related anomalies tended to counter each other, diminishing overall variance in reconstructed timeseries. This is not an issue with infilled series for regions where stations densities were sufficiently high early in the record, for example, in the Midwest (in c). (Link to text §3.6.5).

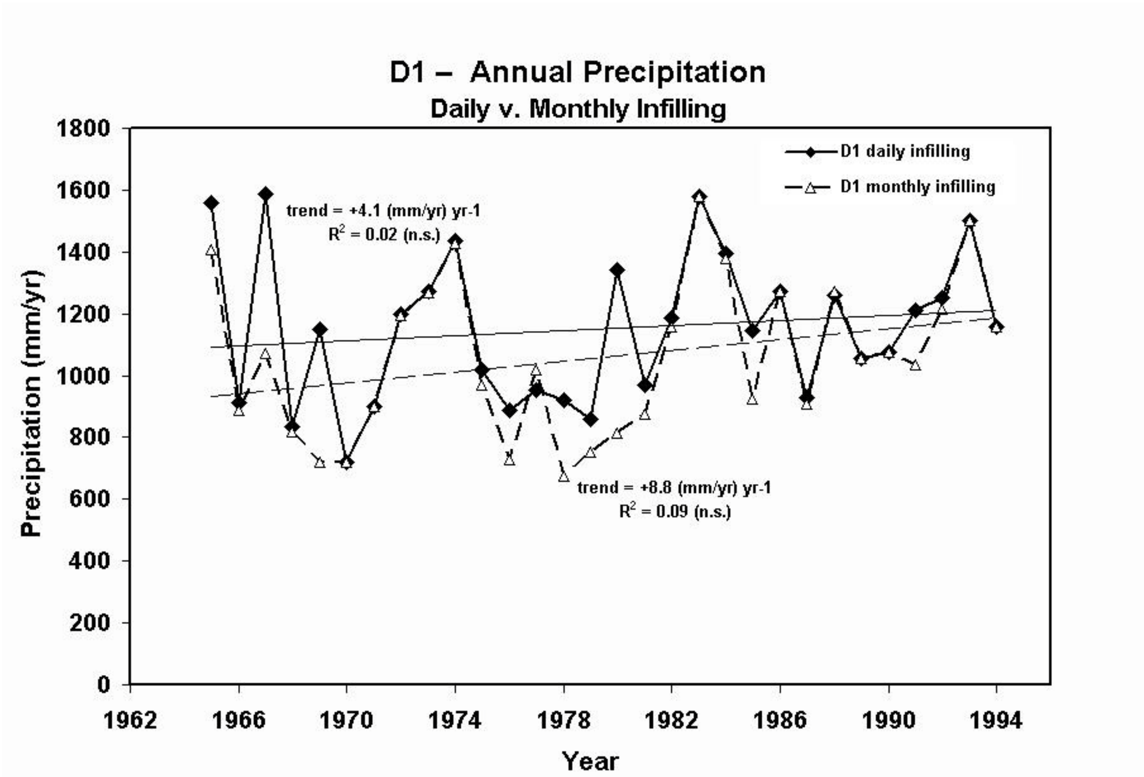


Figure 13. The effects of applying a missing-value method at different temporal resolutions on the annual precipitation record at a high elevation station in the Colorado Front Range (Figure 7, Figure 10) for 1965-1999. The long-standing technique was to infill monthly precipitation for months with missing days using monthly correlations with nearby stations (*dashed curve*). Later, the same spatial method was repeated but applied at a daily timestep based on daily correlations (*solid curve*). (Kittel et al., in preparation)

Discussion: While infilling method did not change the significance of longterm trends (both nonsignificant, n.s.), the daily method produced a markedly different series. Differences in infilled data were accumulated over the year and were due to the daily model being able to select different stations for the regression on a daily basis, rather than just one to represent a month. The daily-method timeseries was taken to be more true to the station record as it preserved more of the observed data. (Link to text §3.6.5)

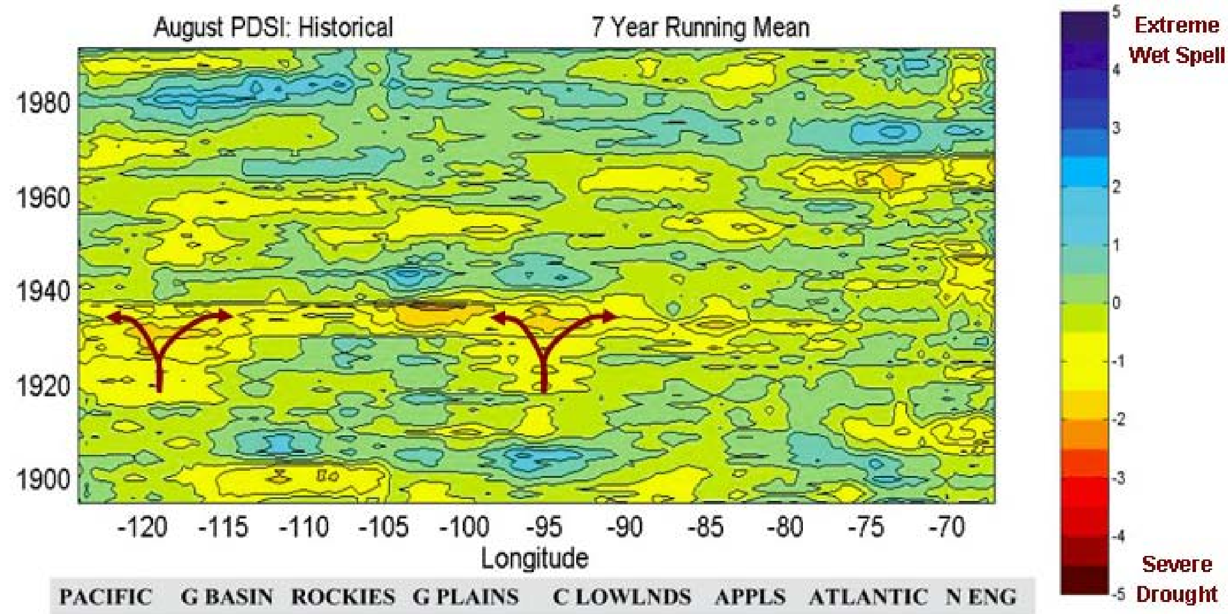


Figure 14. A longitude-time section plot[73] of the Palmer Drought Severity Index (PDSI) for August for the conterminous United States (from Kittel et al. 2004; brown annotations added). Vertical axis is time (year), horizontal axis is longitude (°W). Mapped values are latitudinal (north-south) averages for a given year and longitude, and smoothed using a 7-year running average. Longitude labels: Pacific = Pacific states, G Basin = Great Basin, Rockies = Rocky Mountains, G Plains = Great Plains, C Lowlnds = Central Lowlnds, Appls = Appalachians, Atlantic = Atlantic Coastal Plain, N Eng = New England.

Discussion: The longitude-time section plot shows how droughts (yellow-deep orange) in the far western and central US developed in the 1920's, intensified and merged in the 1930's, and then rapidly dissipated. (Links to text: section plots §4.1.2, PDSI §4.2.2)

Copyrighted figure removed

Figure 15. Temporal smoothing revealing decadal and longer-period oscillatory behavior. (a) Timeseries of the pervasiveness of cold-season climate extremes across the conterminous United States, as represented by the aggregate Climate Extremes Index (CEI, from Gleason et al. 2008; §4.6.2).[158] Bars are yearly values, smoothed curve is a 5-year centered moving average, and black horizontal line is the period average. The value of CEI is the percent of the three dimensional time-space domain (season × region) experiencing extreme climate. The occurrence of extreme climate is based on any of five thermal and moisture measures exceeding their upper or lower 10%-tiles. (b) Weights used in IPCC WG1's centered low-pass 13-year moving filter (figure created from data in Trenberth et al. 2007). The 13 weights have numerators {1-6-19-42-71-96-106-96-71-42-19-6-1} and a denominator of 576. (c) 1850-2005 tropical Atlantic sea surface temperature annual anomalies (bars) and smoothed curve (thick line) using weights in (b) designed to remove less than decadal scale fluctuations (from Trenberth et al. 2007). At the end points, where 13 points are not longer available to the filter, the smoothed line is constrained by minimizing its slope (Mann 2004; see text §4.5.1).

Discussion: In (a), the smoothed timeline suggests that the cold-season extremes index varies with a period of roughly 20 years. In (c), the low-pass filter curve suggests several scales of quasi-periodic behavior with decadal and longer periods. In both cases (a, c), the next step would be to evaluate the original series using spectral techniques (§4.5.2). (Link to text §4.5.1)

Copyrighted figure removed

Figure 16. Multiyear temporal pattern analysis. (a) Spectral analysis of the Southern Oscillation Index (SOI; §4.8, Figure 23b) (from Ghil et al. 2002; brown annotations added). Spectral power is plotted on the *y*-axis. The bottom *x*-axis is frequency (1/month), converted to period (months) along the top *x*-axis. The uppermost smoothed line is the 99% confidence level, and below that, the 95% confidence level. (b) Regime shift detection in the Pacific Decadal Oscillation (PDO; §4.8, Figure 23a) for winter.[159] *Top* vs. *bottom panels* show a sensitivity to detection model parameters.

Discussion: At interannual or lower frequencies [left end of the spectrum in (a)], three peaks exceed the 99% level. Those marked 0.015 and 0.034 cycle/month correspond to dominant modes of the SOI, with periods spanning 2-6 years. The third, left-most peak is the longterm trend. (Link to text §4.5.2) In (b: *top*), two major regime shifts are detected in North Pacific climate dynamics over the last century: in 1946 and more recently in 1977 (marked by vertical arrows). Smaller, less persistent shifts were detected with relaxed cutoff-length thresholds in the regime shift analysis, including a warm-to-cold shift in 1989 (b: *bottom*). These two recent shifts (1977, 1989) are consistent with shifts in an array of abiotic and biotic measures across the northeastern Pacific (Hare and Mantua 2000; see also §4.8.1, §4.8.2). (Link to text §4.5.3)

[158] http://www.ncdc.noaa.gov/oa/climate/research/cei/cei.html

[159] Image from: http://www.beringclimate.noaa.gov/data/BCinclude.php?filename=in_PDO. Model parameters: *p* target significance level, *l* cutoff length (years), *h* Huber weight parameter, and AR1 autoregressive parameter are noted in website links to specific images and described in Rodionov and Overland (2005). Brown annotations added.

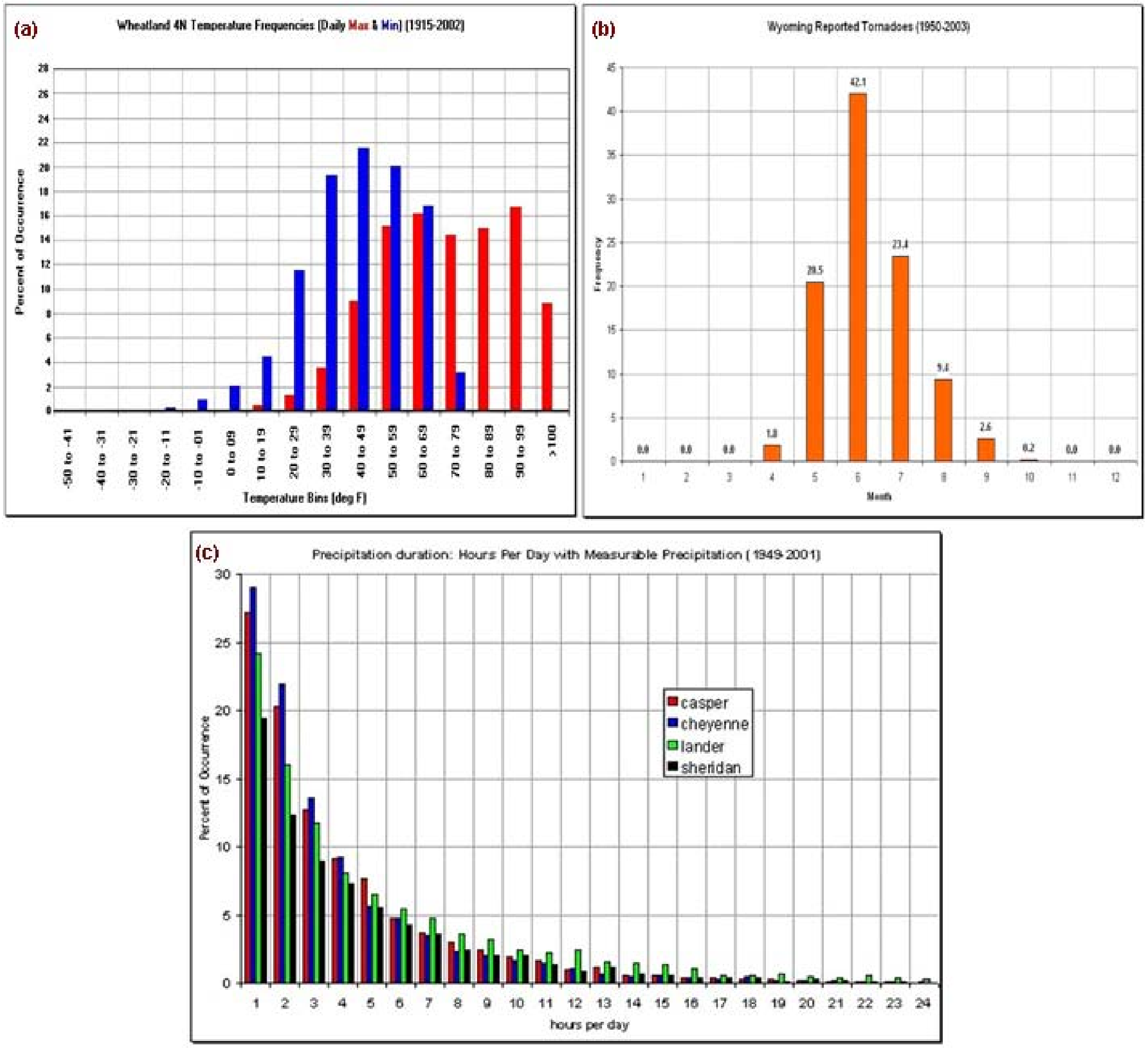


Figure 17. Evaluating the frequency of events. Frequency distribution of (a) daily T_{min} and T_{max} for Wheatland, WY by 10ºF (~6ºC) bins, (b) tornados in Wyoming by month of the year, and (c) precipitation duration by hours per day bins for four stations in Wyoming.[160]

Discussion: The occurrence of these three kinds of weather events is evaluated by binning by event magnitude (a, c) or timing (b). In (a) minimum temperatures (blue bars) at Wheatland are slightly negatively skewed, with a stretched out distribution of the most extreme cold events. Maximum temperatures (red bars) are more broadly distributed, but note that the bins are truncated at the high end (>100ºF), so the distribution of extreme maxima cannot be evaluated. Here, there is a suggestion of bimodal structure, which might reflect two separate (perhaps seasonal) processes, warranting further analysis. In (b), the seasonal distribution of tornados is slightly positively skewed, with more occurrences loaded up at the beginning of the warm season. This distribution is consistent with our understanding that strong gradients in

[160] Images from Wyoming Climate Atlas, http://www.wrds.uwyo.edu/sco/climateatlas/title_page.html, 17 Oct 2008 update.

temperature and/or moisture – most often occurring in late spring through early summer – are precursors for tornado development. For (c), as is characteristic for dry climates, precipitation events are highly concentrated on the short-duration end and are primarily less than one hour, yet with some full-day (or perhaps longer) events. Cheyenne appears to have more very short events, Lander more longer events – this difficult could be tested with a Kolmogorov-Smirnov identical distribution test (§4.6.3). (Link to text §4.6.1).

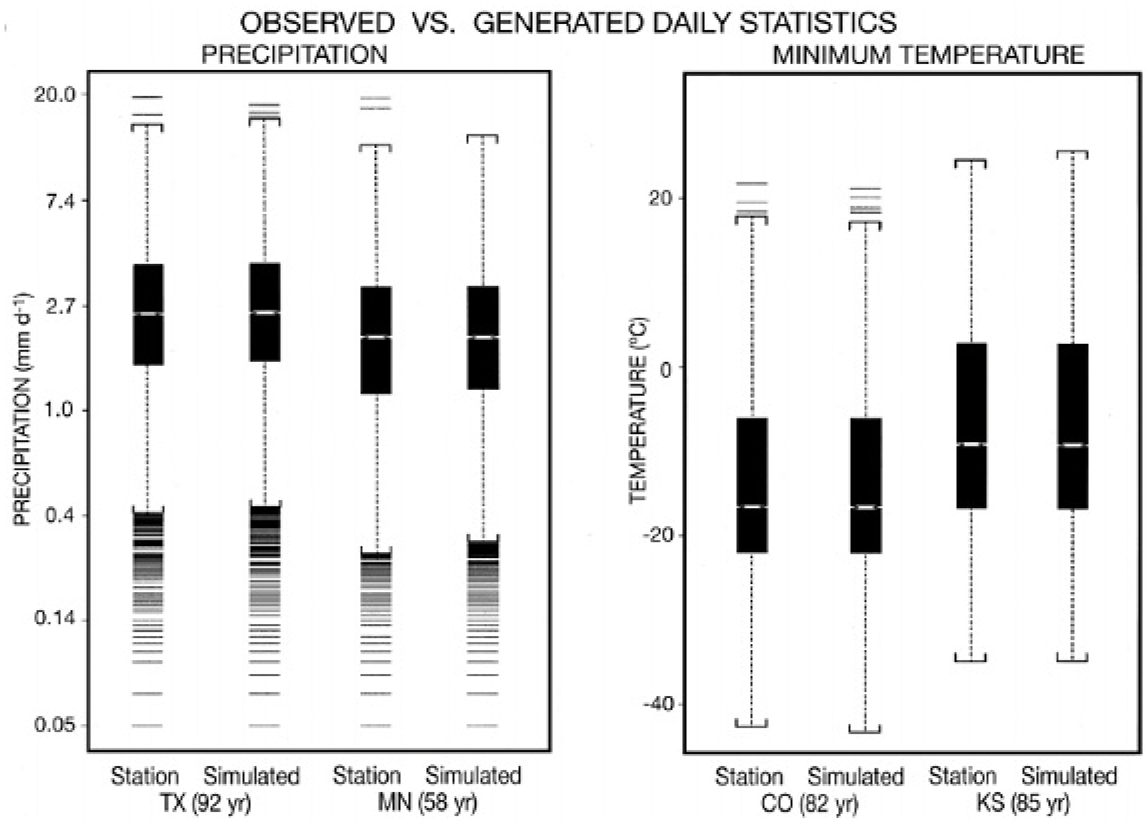


Figure 18. Structure of daily weather events as illustrated by notch plots. Plots are for January (*left panel*) precipitation and (*right*) minimum temperature from daily station records and daily weather model simulations for selected points (from Kittel et al. 2004; brown annotation added). The daily weather model was WGEN (§3.6.7).[70] Points are in central Texas (TX), southern Minnesota (MN), western Colorado (CO), and northwestern Kansas (KS). Record length (years) is indicated in parentheses after each location. Notch plots show median (center of notch), standard deviation (white area = 2 SD), interquartile range (IQR; dark boxes around median, spanning 25 to 75% of values), data range (brackets), and outliers (bars beyond the brackets; defined as values > +1.5 IQR or < –1.5 IQR). Precipitation is for wet days (only non-zero values are included) and is plotted in natural-log space (though note, *y*-axis tick labels are in original precipitation units, mm/day). Box plots present the same distribution information but show the median without the standard deviation 'notch.'

Discussion: This graphic analysis shows that daily precipitation event size distribution is similar for both Texas and Minnesota stations with the median size and distribution set slightly higher at the Texas site. The distributions are symmetrical in natural-log space, meaning in linear space that the distributions have long tails on the high end. Both sites are characterized by many small («0.4 mm/day) outlier events. The daily minimum temperatures distribution for the Colorado and Kansas stations are similar, both asymmetrical with a broader distribution above the median than below. Relative to the Kansas site, the Colorado station has a distribution overall shifted towards lower minimum temperatures, a less broad IQR, and a few high outliers over the 82-year record. (Link to text §4.6) The frequency distribution of WGEN-simulated values strongly follows observed (§3.6.7).

Copyrighted figure removed

Figure 19. Historical changes in temperature and precipitation event frequency. (a, b) Global change in the occurrence of high extreme daily T_{min} ('warm nights'), defined as T_{min}'s in the top 10% of observations: (a) days with warm nights by year and (b) the frequency distribution of the percent-of-year with warm nights based on 202 stations, evaluated for 3 successive multidecadal periods (from Trenberth et al. 2007, based on Alexander et al. 2006; brown annotations added). Smoothed line in (a) is a 21-term filter.[161] (c, d) Frequency of precipitation over the conterminous United States (for all but the smallest events, ≥ 0.1 mm/hour) (from Trenberth 1998): (c) January precipitation frequency (percentage of hours with precipitation) and (d) historical trend in winter precipitation frequency (change in precipitation hours per decade) over the period 1963–1994. In (c), areas with less than 8% of hours are stippled and greater than 16% are hatched (contour interval=2%). In (d), negative trends are shown with dashed contour lines (in the Southeast and Pacific Northwest) and zero or positive change with solid contours (contour interval=4 hours/decade; brown -0- contour labels added). Areas with trends that are statistically significant (at 5% level) are stippled.

Discussion: In (a) – on a global basis, the number of days with warm nighttime temperature extremes increased from the early part of the 20th century to the beginning of the present one. The resulting greater portion of the year experiencing warm nights is expressed in (b) in terms of both a positive shift (horizontal arrow) in and a broadening of the frequency distribution of the % of year with warm nights. The next step would be to evaluate changes in the frequency timeseries (a) using trend analysis (§4.3), while changes in the frequency distribution (b) could be analyzed using identical-distribution tests (§4.6.3). Winter U.S. precipitation frequency [shown for January in (c)] is greatest in the Southeast and Pacific Northwest and low throughout the continental interior. (d) shows statistically significant late 20th century changes in this pattern, with precipitation frequency reduced most strongly in the northern U.S. Rockies and enhanced broadly across the Southern Great Plains and into the Southwest. (Link to text §4.6).

[161] http://hadobs.metoffice.com/hadcrut3/smoothing.html

Copyrighted figure removed

Figure 20. (a) Cumulative probability function (dashed line; right *x*-axis scale) for a binomial relative frequency distribution function (solid line; left axis). (b) Comparison of two theoretical binomial curves with narrow (dashed lines) vs. broad (solid lines) dispersion, plotted as both relative frequency (simple lines) and cumulative (lines with symbols) functions. The cumulative probability curves are significantly different (p=0.03, Kolmogorov-Smirnov identical distribution test, 'KS-test' – evaluated with an online calculator[116]). (c) Comparison of empirical cumulative functions for two sets of observations with similar means (t-test, nonsignificant at p>>0.05) but with strongly differing variances, resulting in statistically different distributions (KS-test, p=0.023). D is the maximum vertical deviation between the two curves and is the Kolmogorov-Smirnov test statistic (from Kirkman 1996; brown annotation added).[162] (d) Comparison of Beaufort force shipboard wind records in the English Channel for successive multidecadal periods. The frequency distribution of observations pre-1900 (open circles) was significantly different from subsequent records through 1939 (open squares and closed circles) (KS-test, p<0.01) (from Peterson and Hasse 1987; brown annotation added).

Discussion: The two cumulative distributions in (c) have similar means (at x=0). However, just as in the theoretical example (b), the dispersion of observations in (c) about the mean is much stronger for one set of observations (dashed curve, which starts earlier and climbs more slowly) than the other (solid curve). The significance of this difference is confirmed with the KS-test. In (d), the KS-test allowed Peterson and Hasse (1987) to identify a significant distribution shift in the Beaufort force record after 1899 – they suggest this maybe an artifact of a shift from sailing to steam ships at the end of the 19th century, perhaps related to concomitant changes in navigation and observer habits. (Link to text §4.6.3)

Copyrighted figure removed

Figure 21. Spatial analyses showing regional connections. (a) Point-covariance map for surface ozone of the eastern U.S. between the grid point indicated by the red dot (in northeast Ohio) and the rest of the domain (from Matsuo 2005). Covariance units are [(ppb O_3)^2]. (b-e) Spatial autocorrelation model of monthly precipitation across the Central and Southern Rocky Mountains (Colorado is the state in the center of all maps). (b) Stations used in the analysis. (c) Semivariogram (mm/mo)[163] vs. spatial separation (km) at a given time, February 1996 – showing increasing differences (spatial variance) in precipitation between stations as inter-station distance increases. (d) Map of July 1996 precipitation semivariograms across the region at a set distance, 40km [marked with a dotted line in (c)]. Areas with lower semivariogram values have stronger spatial connectivity. (e) Topographic relief[164] for the domain in (b, c, d). (b-d from Fuentes et al. 2006; brown annotations added).

Discussion. Point-covariance map (a) shows the spatial extent of sites that have a strong correlation in time with the focus site's ozone dynamics. The text (§4.7.1) discusses this type of presentation in terms of point-correlation maps, which impart the same information – recall that

[162] Image from Kirkman (1996): http://www.physics.csbsju.edu/stats/KS-test.html

[163] Note re units: Units for semivariance are the square of the variable's units. In this example, however, semivariance units are 'mm' rather than 'square mm' because precipitation was first square-root transformed to better meet analysis requirements.

[164] Image from Jones et al., 1996, *Nature* 381:37-41. http://cires.colorado.edu/people/jones.craig/GSA/slide1l_big.JPG

covariances and correlations are directly related.[165] (c) and (d) show how monthly precipitation at any place in the domain is related to that of the surrounding region: in (c) as a function of distance between stations – spatial variance increases as spacing increases, and in (d) as a function of location within the domain – spatial variance is generally higher in regions with highest topographic heterogeneity (in e). We also see in (d) that the semivariogram at 40km is anisotropic, i.e., that spatial correlation depends on compass direction. (Link to text §4.7.2)

Copyrighted figure removed

Figure 22. Cross-spectral Fourier analysis of two oceanic regional paleoclimate proxies, ^{18}O and ^{13}C isotopic variables.[166] (a) Power spectra for the two individual timeseries, (b) their spectral coherence (h^2), and (c) phase. Significance levels (5 and 20%) for coherence peaks are shown in (b). In (c), confidence intervals (C.I.; vertical 'I' bars) for the phase of major coherence peaks in (b) test if the phase is significantly different from 0º. Phase is presented in terms of degrees, with cycles defined as in phase at 0º and of opposite phase at 180º. In x-axis frequency and period units, 'ky'=10^3 years. Analysis bandwidth is shown in the upper right ('bw'=0.01/ky).

Discussion: The two climate proxy records share significant coherent (b) and in-phase (c) spectral power at periods of ~40 and ~100ky. Coherence values for these peaks ($h^2 \geq 0.80$) indicate that 80% or more of variance in these frequencies is shared (§4.7.3). In c, note that their C.I. bars extend to 0º. The periods of these peaks in ocean regional climate proxies suggest climate links with Earth orbital dynamics as they closely correspond to Milankovitch orbital parameter cycles for tilt (41ky) and eccentricity (100ky; Imbrie et al. 1992, 1993). (Link to text §4.7.3).

[165] Correlation(x,y) = covariance(x,y)/[SD(x)*SD(y)]
[166] Figure modified from: McDuff, R.E., and G.R. Heath. 2001. Phase Relationships of Proxy Variables. Oceanography 540. Marine Geological Processes. Course web page. University of Washington, School of Oceanography. http://www2.ocean.washington.edu/oc540/lec01-26/ Figure based on Specmap data from the National Geophysical Data Center, NGDC. Copyright (©) 1994-2001 Russell E. McDuff and G. Ross Heath. All Rights Reserved. Used with permission.

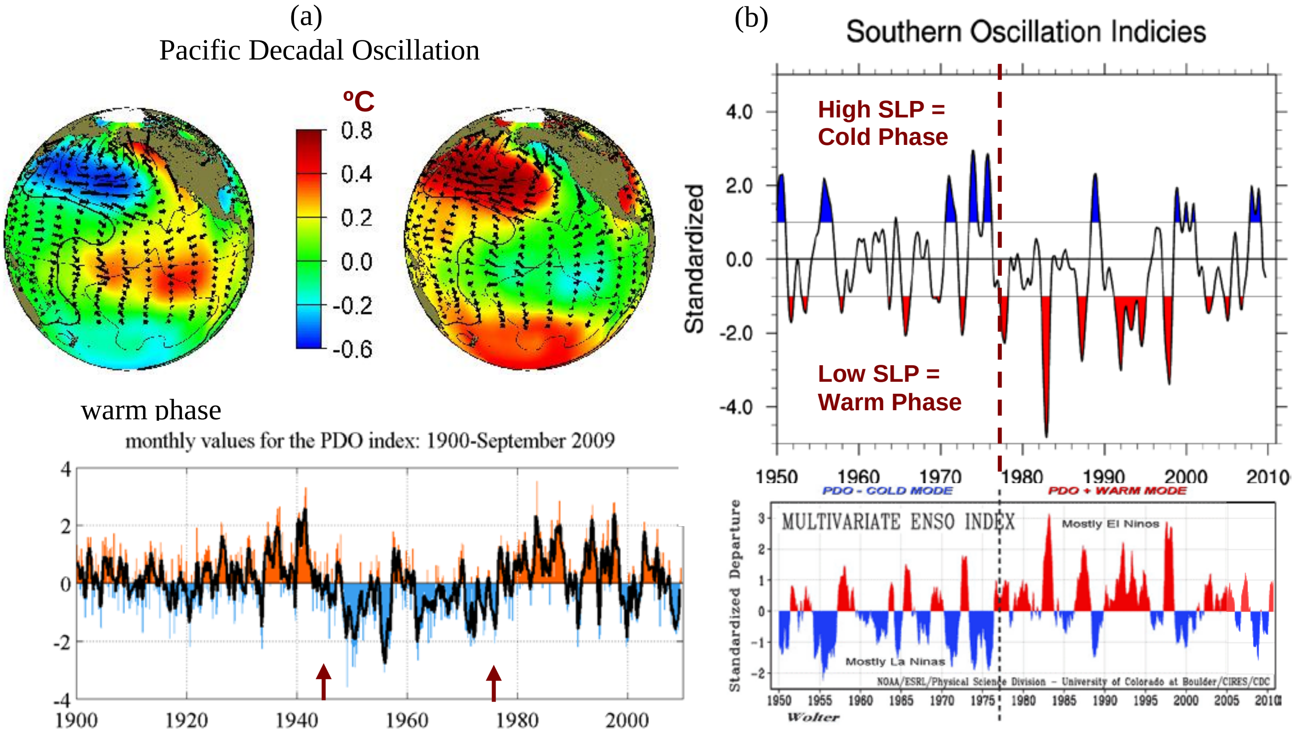
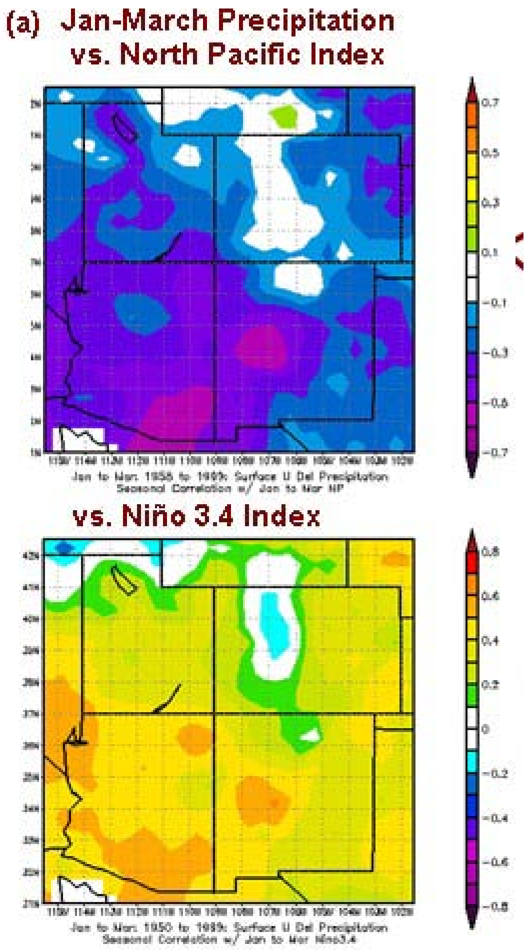


Figure 23. Hemispheric climate oscillations' spatial and temporal signatures: (a) The Pacific Decadal Oscillation (PDO) characteristic warm (*top left panel*) and cool (*right*) phase conditions across the Pacific: wintertime anomaly patterns in sea surface temperature (SST; color zones, ºC), sea level pressure (SLP, contours) and surface windstress (arrows). (*Bottom*) PDO Index monthly timeseries (standardized SST anomalies for the first principal component, §4.8.2).[167] (b) El Niño-Southern Oscillation (ENSO) timeseries since 1950 for two indices: Southern Oscillation Index (SOI, based on SLP anomalies; Trenberth 1984) and Multivariate ENSO Index (MEI, based on SLP, SST, surface wind, surface air temperature, and cloudiness anomalies; Wolter and Timlin 1993).[168] Plots are presented with the *x*-axes' year scales lined up. The plots in (b) illustrate two different approaches to data blocking: *(top panel)* High magnitude SOI events are selected using a threshold, *(bottom)* MEI is blocked by PDO cold vs. warm regime, which shifted after 1976 (vertical dashed line). For timeseries in both (a) and (b), blue = cool phase, red/orange = warm phase [note in (b) that SOI phases are of opposite sign than El Niño/La Niña (warm/cold) temperature anomalies]. (In a-b, brown annotations added)

Discussion: PDO phases are termed 'warm' and 'cool' (or 'cold') based on SST anomalies in central tropical to northeast extratropical Pacific (off of North American west coast) (a, *top*). Wind anomalies reverse directions in phase with temperature shifts. The PDO Index timeseries (a, *bottom*) shows the multidecadal persistence of these phases, with regime shifts following 1946 and 1976 (arrows along *x*-axis; see Figure 16b). In (b), note that the two ENSO indices capture the main events, though their relative structures vary. Both indices show the prevalence of cold ENSO events during the PDO's cold regime and warm ENSO in the PDO's warm regime – a pattern that suggests constructive interference between the two dynamics (Biondi et al. 2001). (Link to text §4.8.1)

---

[167] Images from: http://jisao.washington.edu/pdo/ (timeseries through September 2009, periodically updated)
[168] SOI image from: http://www.cgd.ucar.edu/cas/catalog/climind/soi.html. MEI image from: http://www.intellicast.com/Community/Content.aspx?ref=rss&a=126 through November 2006 – to this image is appended updated series through October 2009 from http://www.esrl.noaa.gov/psd/people/klaus.wolter/MEI/mei.html.



Copyrighted figure removed

Figure 24. Pacific Ocean Basin teleconnections to conterminous U.S. and Alaskan climates. (a) Teleconnection correlation maps for January-March precipitation for the U.S. Four Corners region with *(top)* the North Pacific Index (NPI) and *(bottom)* Niño 3.4 Index (for 1958-1999 and 1950-1999, respectively). Warm colors = positive correlation, cold colors = negative. Plots generated on-line: http://www.esrl.noaa.gov/psd/data/correlation/. (b) Teleconnections for conterminous U.S. drought frequency as a function of the interaction between phases (+/–) of the Pacific Decadal Oscillation (PDO) and Atlantic Multidecadal Oscillation (AMO): blue areas = fewer droughts than normal, red = more (McCabe et al. 2004; image from Steward 2005, brown annotations added). (c) Mean annual temperature departure for Alaska (1949-2008) showing a dramatic shift in the late 1970's corresponding to the 1976/77 PDO regime shift (Figure 23a).[169]

Discussion: The negative correlation field in (a, *top*) indicates that a *negative* NPI in January-March corresponds to *high* precipitation into the American Southwest and Colorado River Basin. The underlying mechanism is that a negative NPI means a strong, deep Aleutian Low – this forces the mid-latitude jet stream and embedded winter storms to track farther south coming into North America, bringing moisture directly to the Southwest. In (b), these dynamics are similarly reflected in PDO teleconnections – *negative* NPI generally correspond to *warm* phases of the PDO ('+PDO') and so giving *wetter* conditions in the Southwest (b, *top*) than under –PDO (b, *bottom*). In (a, *bottom*), the *warm* SST phase of ENSO (positive Niño 3.4 Index) is *positively*

[169] From Alaska Climate Research Center (http://climate.gi.alaska.edu/ClimTrends/Change/TempChange.html; brown annotations added). Used with permission.

related to Southwest precipitation in January-March. Warm SST's in the eastern Pacific are linked to greater rainfall in the eastern tropical Pacific and a weakened Subtropical High off the coast of northern Mexico – these lead to more precipitation in northern Mexico and the American Southwest. In (b), the +AMO phase brings drought across far more of the conterminous U.S. than under –AMO (*right* vs. *left panels*). However, which regions are affected by this drought is strongly determined by PDO phase. During the –PDO phase, when the Aleutian Low is weak, the mid-latitude jet stream is more zonal (following latitude lines) and brings moisture directly into the Pacific Northwest – keeping the +AMO drought out of this region (*top* vs. *bottom right panels*). This interaction could be evaluated statistically for a given region using multifactor techniques (§4.8.2). (Links to text §4.8, Table 3, Table 4).

Appendix E. Literature Cited

Alexander, L. V., et al. [23 coauthors]. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res*. 111, D05109, doi:10.1029/2005JD006290.

Allen, T. F. H. and T. B. Starr. 1982. *Hierarchy: Perspectives for Ecological Complexity.* Chicago: The University of Chicago Press. 310 p.

Alley, W. M. 1984. The Palmer Drought Severity Index: limitations and assumptions. *Journal of Climate and Applied Meteorology* 23:1100-1109.

Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2003. *Hierarchical Modeling and Analysis for Spatial Data.* Monographs on Statistics and Applied Probability, Routledge. 472 p.

Baron, J. S. 2006. Hindcasting nitrogen deposition to determine an ecological critical load. *Ecological Applications* 16:433-439

Barton, S. B. and J. A. Ramirez. 2004. Effects of El Niño Southern Oscillation and Pacific Interdecadal Oscillation on water supply in the Columbia River Basin. *J. Water Resources Planning and Management* 130:281-289.

Beniston, M. and D. Stephenson. 2004. Extreme climatic events and their evolution under changing climatic conditions. *Global and Planetary Change* 44:1-9.

Billings, W. D. and L. C. Bliss. 1959. An alpine snowbank environment and its effects on vegetation, plant development, and productivity. *Ecology* 40:388-397.

Biondi, F., A. Gershunov, and D. R. Cayan. 2001. North Pacific decadal climate variability since 1661. *J. Climate* 14:5–10.

Biondi, F., T. J. Kozubowski, and A. K. Panorska. 2002. Stochastic modeling of regime shifts. *Climate Research* 23:23–30.

Box, G. E. P. and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. *J. Amer. Stat. Assoc.* 70:70–79.

Braswell, B. H., D. S. Schimel, E. Linder, and B. Moore III. 1997. The Response of Global Terrestrial Ecosystems to Interannual Temperature Variability. Science 278:870-873.

Burnaby, T. P. 1953. A suggested alternative to the correlation coefficient for testing the significance of agreement between pairs of time series, and its application to geologic data. *Nature* 172:210–211.

Burroughs, J. 2008. *Integrated Global Radiosonde Archive - Quality Control.* National Climate Data Center, NOAA. Available from:

http://www.ncdc.noaa.gov/oa/climate/igra/index.php?name=quality and http://www.ncdc.noaa.gov/oa/climate/igra/caption.php?fig=quality-fig3.

Cade, B. S. and B. R. Noon. 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* 1:412-420.

Castro, C. L., T. B. McKee, and R. A. Pielke, Sr. 2001. The relationship of the North American monsoon to tropical and North Pacific sea surface temperatures as revealed by observational analyses. *J. Climate* 14:4449-4473.

Clark, I. and W.V. Harper. 2000. *Practical Geostatistics 2000.* Ecosse North America. 442 p.

Chang, Y. S., D. Jeon, H. Lee, H. S. An, J. W. Seo, and Y. H. Youn. 2004. Interannual variability and lagged correlation during strong El Niño events in the Pacific Ocean. *Climate Research* 27:51-58.

Chatfield, C. 1995. *Problem Solving: A Statistician's Guide*, 2nd ed. Chapman & Hall. 350 p. Available in part from: http://books.google.com/books?id=EA3jBSe0c3wC

Chen, M., W. Shi, P. Xie, V.B.S. Silva, V.E. Kousky, R.W. Higgins, and J.E. Janowiak. 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res*. 113:D04110.

Christy, J.R., W.B. Norris, and R.T. McNider. 2009. Surface temperature variations in East Africa and possible causes. *J. Climate* 22:3342–3356

CLIMAS. 2002. Climate Divisions: to use or not to use? *END InSight*, November 2002, p. 3-4. CLIMAS (Climate Assessment for the Southwest), University of Arizona. Available from: http://www.climas.arizona.edu/forecasts/articles/climdiv_nov2002.pdf

Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values.* Springer. 225 p.

Conover, W. J. 1999. *Practical Nonparametric Statistics*, 3rd ed. Wiley. 584 p.

Crawley, M. J. 2002. *Statistical Computing: An Introduction to Data Analysis using S-Plus.* Wiley. 772 p.

Crawley, M. J. 2007. *The R Book.* Wiley. 950 p.

Cressie, N. A. C. 1993a. *Statistics for Spatial Data.* Wiley, New York. 928 p.

Cressie. N. A. C. 1993b. Aggregation in geostatistical problems. In: Soares, A. (ed) *Geostatistics Troia '92*, Vol 1, pp. 25–36. Kluwer Academic Publishers, Dordrecht.

Dai, A., K. E. Trenberth, and T. R. Karl. 1998. Global variations in droughts and wet spells: 1900-1995. *Geophysical Research Letters* 25(17):3367-3370.

Daly, C., W.P. Gibson, G.H. Taylor, G.L. Johnson, and P. Pasteris. 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research* 22:99-113. Available from: http://www.prism.oregonstate.edu/.

Daly, C., W. P. Gibson, G. H. Taylor, M. K. Doggett, and J. I. Smith. 2007. Observer bias in daily precipitation measurements at United States Cooperative Network Stations. Bull. Amer. Meteor. Soc. 88:899-912.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G.H. Taylor, J. Curtis, and P.P. Pasteris. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28:2031-2064.

Delcourt, H. R., P. A. Delcourt, and T. Webb III. 1983. Dynamic plant ecology: the spectrum of vegetation change in space and time. *Quat. Sci. Rev.* 1:153-175.

Díaz, S. C., C. A. Salinas-Zavala, and S. Hernandez-Vazquez. 2008. Variability of rainfall from tropical cyclones in northwestern Mexico and its relation to SOI and PDO. *Atmósfera* 21: 213 – 223.

Di Luzio, M., G. L. Johnson, C. Daly, J. K. Eischeid, and J. G. Arnold. 2008. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States. J. Appl. Meteor. Climatol. 47:475–497.

Doesken, N. J. 2005. The National Weather Service MMTS (Maximum-Minimum Temperature System) -- 20 years after. *15th Conference on Applied Climatology and 13th Symposium on Meteorological Observations and Instrumentation.* Paper JP1.26. America Meteorological Society, Boston. Available from: http://ams.confex.com/ams/pdfpapers/91613.pdf.

Draper, N. R., and H. Smith. 1998. *Applied Linear Regression*, 3rd ed. Wiley. 736 p.

Drews, C. 2003. *Detecting Climate Change in Canadian Ice Data.* Research Paper, USDA Graduate School. http://www.highestlake.com/canadice.html.

Easterling, D.R., T.R. Karl, E.H. Mason, P.Y. Hughes, D.P. Bowman, R.C. Daniels, and T.A. Boden. 1996. *United States Historical Climatology Network (US HCN) Monthly Temperature and Precipitation Data.* ORNL/CDIAC-87, NDP-019/R3. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory. Available from: http://www.ncdc.noaa.gov/oa/climate/research/ushcn/ushcn.html. Data access: http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn.html.

Easterling, D.R., B. Horton, P.D. Jones, T.C. Peterson, T.R. Karl, D.E. Parker, M.J. Salinger, V. Razuvayev, N. Plummer, P. Jamason, and C.K. Folland. 1997. Maximum and minimum temperature trends for the globe. *Science* 277:364-367.

Easterling, D.R., T.R. Karl, J.H. Lawrimore, and S.A. Del Greco. 1999. *United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data for 1871-1997.* ORNL/CDIAC-118, NDP-070. Carbon Dioxide Information Analysis Center, Oak Ridge

National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee. Available from: http://www.ncdc.noaa.gov/oa/climate/research/ushcn/ushcn.html. Data access: http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn.html.

Easterling, D.R., J.L. Evans, P.Ya. Groisman, T.R. Karl, K.E. Kunkel, and P. Ambenje. 2000. Observed variability and trends in extreme climate events: a brief review. *Bulletin of the American Meteorological Society* 81:417–425.

Eischeid, J.K., C.B. Baker, T.R. Karl, and H.F. Diaz. 1995. The quality control of long-term climatological data using objective data analysis. *J. Appl. Meteor.* 34:2787–2795.

Endoe, J. 2009. *The USHCN Version 2 Serial Monthly Dataset.* National Climate Data Center, NOAA. Available from: http://www.ncdc.noaa.gov/oa/climate/research/ushcn/.

Forchhammer, M.C., and E. Post. 2004. Using large-scale climate indices in climate change ecology studies. *Population Ecology* 46:1-12.

Frei, C., and C. Schär. 2001. Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. *Journal of Climate* 14:1568-1584.

Fuentes, M., T.G.F. Kittel, and D. Nychka. 2006. Sensitivity of ecological models to their climate drivers: Statistical ensembles for forcing. *Ecological Applications* 16:99–116.

Gandin, L.S. 1988. Complex quality control of meteorological observations. *Monthly Weather Review* 116:1137-1156.

Garson, G.D. 2009. *Statnotes: Topics in Multivariate Analysis.* North Carolina State University http://www2.chass.ncsu.edu/garson/pa765/statnote.htm.

Gedalof, Z. and D..J. Smith. 2001. Interdecadal climate variability and regime-scale shifts in Pacific North America. *Geophys. Res. Lett.* 28:1515–1518.

Gedalof, Z., N..J. Mantua, and D..L. Peterson. 2002. A multi-century perspective of variability in the Pacific Decadal Oscillation: new insights from tree rings and coral. *Geophys. Res. Lett*. 29(24): 2204.

Gershunov, A. and T. Barnett. 1998. Interdecadal modulation of ENSO teleconnections. *Bulletin of the American Meteorological Society* 79(12):2715-2725.

Ghil, M., M.R. Allen, M.D. Dettinger, K. Ide, D. Kondrashov, M.E. Mann, A.W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. 2002. Advanced spectral methods for climatic time series. *Rev. Geophys.* 40(1):1003.

Gleason, K.L., J.H. Lawrimore, D.H. Levinson, T.R. Karl, and D.J. Karoly. 2008. A revised U.S. Climate Extremes Index. J. Climate 21:2124-2137. Available from: http://journals.ametsoc.org/doi/full/10.1175/2007JCLI1883.1.

Goubanova, K., and L. Lia. 2007. Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations. *Global and Planetary Change* 57:27-42.

Graumlich L.J., M.F.J. Pisaric, L.A. Waggoner, J.S. Littell, and J.C. King. 2003. Upper Yellowstone River flow and teleconnections with Pacific Basin climate variability during the past three centuries. *Climatic Change* 59:245-262.

Gray, S. T., J. L. Betancourt, C. L. Fastie, and S. T. Jackson. 2003. Patterns and sources of multidecadal oscillations in drought-sensitive tree-ring records from the central and southern Rocky Mountains. *Geophys. Res. Lett.* 30(6):1316.

Guttman, N. B. 1999. Accepting the Standardized Precipitation Index: a calculation algorithm. *J. Amer. Water Resour. Assoc.* 35(2):311-322.

Guttman, N. and R. Quayle. 1996. A historical perspective of U.S. climate divisions. *Bulletin of the American Meteorological Society* 77(2):293–304.

Haas, T.C. 1995. Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *J Am Stat Assoc* 90:1189-1199.

Haimberger, L. 2007. Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate* 20:1377–1403.

Hale, R..C., K..P. Gallo, T..W. Owen, and T..R. Loveland. 2006. Land use/land cover change effects on temperature trends at U.S. Climate Normals stations. *Geophys. Res. Lett.* 33, L11703.

Hallett, T..B., T. Coulson, J..G. Pilkington, T..H. Clutton-Brock, J..M. Pemberton, and B. Grenfell. 2004. Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature* 430:71–75.

Hamed, K..H., and A..R. Rao. 1998. A modified Mann-Kendall test for autocorrelated data. *J. Hydrology* 204:182-196.

Hansen, J., R. Ruedy, J. Glascoe, and M. Sato. 1999. GISS analysis of surface temperature change. *J. Geophys. Res.* 104(D24):30997-31022.

Hao, L., and D.Q. Naiman. 2007. *Quantile Regression.* Quantitative Applications in the Social Sciences, v. 149. Sage Publications. 136 pp. In part on Google Books: http://books.google.com/books?id=nKVBXePWtSoC.

Harcum, J. B., J. C. Loftis, and R. C. Ward. 1992. Selecting trend tests for water quality series with serial correlation and missing values. *Water Resources Bulletin* 28:469-478.

Hare, S. R. and N. J. Mantua. 2000. Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Prog. Oceanogr.* 47(2-4):103-146.

Heim, R. R. 2002. A review of twentieth-century drought indices used in the United States. *Bull. Am. Meteorol. Soc.* 83:1149−1165.

Helsel, D..R., and R..M. Hirsch. 2002. *Statistical Methods in Water Resources.* U.S. Geological Survey, Techniques of Water-Resources Investigations Book 4, Chapter A3. http://pubs.usgs.gov/twri/twri4a3/index.html.

Hijmans, R..J., S..E. Cameron, J..L. Parra, P..G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965–1978. doi: 10.1002/joc.1276.

Hirsch, R.M., and J.R. Slack. 1984. A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research* 20(6):727–732.

Hirsch, R. M., J. R. Slack, and R. A. Smith. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research* 18:107-121.

Holling, C..S. 1992. Cross-scale morphology, geometry and dynamics of ecosystems. *Ecol. Monogr.* 62(4):447-502.

Hosking, J..R..M. and J. R. Wallis. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, Cambridge, UK. 224 p. Available in part on Google Books: http://books.google.com/books?id=OfB57wi7VcUC.

Hu, Q., and G.D. Willson. 2000. Effects of temperature anomalies on the Palmer Drought Severity Index in the central United States. *Int. J. Climatol.* 20:1899-1911.

Hubbard, K.G., N.B. Guttman, J.S. You, and Z. Chen. 2007. An improved QC process for temperature in the daily cooperative weather observations. *Journal of Atmospheric and Oceanic Technology* 24:206-213.

Hurrell, J. W. 1996. Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophys. Res. Lett.* 23:665–668.

Hurrell, J.W., Y. Kushnir, G. Ottersen and M. Visbeck. 2003. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact.* Geophysical Monograph Series. American Geophysical Union, Washington, DC.

Hutchinson, M.F. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems* 9:385–403.

Hutchinson, M. F. 2004. *Anusplin Version 4.3.* Centre for Resource and Environmental Studies. The Australian National University: Canberra, Australia.

Imbrie, J., E. Boyle, S. Clemens, A. Duffy, W. Howard, G. Kukla, J . Kutzbach, D. Martinson, A. Mcintyre, A. Mix, B. Molfino, J. Morley, L. Peterson, N. Pisias, W. Prell, M. Raytoo, N. Shackleton, and J. Toggweiler. 1992. On the structure and origin of major glaciation cycles. 1. Linear responses to Milankovitch forcing. *Paleoceanography* 7:701-738.

Imbrie, J., A. Berger, E. Bogle, S. Clemens, A. Duffy, W. Howard, G. Kukla, J. Kutzbach, D. Martinson, A. McIntyre, A. Mir, B. Molfino, J. Morley, L. Peterson, N. Pisias, W. Prell, M. Raymo, N. Shackleton, and J. Toggweiler. 1993. On the structure and origin of major glaciation cycles. 2. The 100,000 year cycle. *Paleoceanography* 8:699-735.

Karl, T.R., and R.W. Knight. 1985. *Atlas of Monthly Palmer Hydrological Drought Indices (1931–1983) for the Contiguous United States*. Historical Climatology Series 3-7, National Climatic Data Center, Asheville, North Carolina, USA.

Karl, T.R., and C.N. Williams, Jr. 1987. An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Climate Appl. Meteor.* 26:1744-1763.

Karl, T.R., C.N. Williams, Jr., P.J. Young, and W.M. Wendland. 1986. A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States. *J. Climate Appl. Meteor.* 25:145-160.

Karl, T.R., H.F. Diaz, and G. Kukla. 1988. Urbanization: its detection and effect in the United States climate record. *J. Climate* 1:1099-1123.

Katz, R.W., and B.G. Brown. 1992. Extreme events in a changing climate: Variability is more important than averages. *Climatic Change* 21:289–302.

Kestin T., D.J. Karoly, J.L. Yang, and N.A. Rayner. 1998. Time-frequency variability of ENSO and stochastic simulations. *J. Climate* 11:2258-2272.

Kharin, V.V., and F.W. Zwiers. 2000. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *J. Climate* 13:3760–3788.

Kimball, S.L., B.D. Bennett, and F.B. Salisbury. 1973. The growth and development of montane species at near-freezing temperatures. *Ecology* 54:168-173.

Kirkman, T.W. 1996. *Statistics to Use.* College of Saint Benedict and Saint John's University. http://www.physics.csbsju.edu/stats/ .

Kittel, T.G.F. 1990. Climatic variability in the shortgrass steppe. Pp. 67-75, in: D. Greenland and L.W. Swift, Jr. (eds.). *Climate Variability and Ecosystem Response*. U.S. Forest Service, Southeastern Region. Gen. Tech. Rpt. SE-65.

Kittel, T.G.F., N.A. Rosenbloom, J.A. Royle, C. Daly, W.P. Gibson, H.H. Fisher, P. Thornton, D. Yates, S. Aulenbach, C. Kaufman, R. McKeown, D. Bachelet, D.S. Schimel, and VEMAP2 Participants. 2004. The VEMAP Phase 2 bioclimatic database. I: A gridded historical (20th century) climate dataset for modeling ecosystem dynamics across the conterminous United States. *Climate Research* 27:151-170.

Kittel, T.G.F., T. Ackerman, M. Hartman, M.W. Williams, M.V. Losleben, and K. Chowanski. *In Prep.* Elevational and seasonal dependence in recent climate change across an alpine and subalpine landscape in the Colorado Front Range, USA.

Koenker, R. 2005. *Quantile Regression*. Cambridge University Press. 368 pp. In part on Google Books: http://books.google.com/books?id=Xi_dTAeAmGcC.

Koenker, R., and F. Schorfheide. 1994. Quantile spline models for global temperature change. *Climatic Change* 28: 395-404.

Kohler, M.A. 1949. Double-mass analysis for testing the consistency of records and for making adjustments. *Bull. Amer. Meteor. Soc.* 30:188–189.

Kondrashov, D., and M. Ghil. 2006. Spatio-temporal filling of missing points in geophysical data sets. *Nonlin. Processes Geophys.* 13:151-159.

Kondrashov, D., and M. Ghil. 2007. Reply to T. Schneider's comment on "Spatio-temporal filling of missing points in geophysical data sets." *Nonlin. Processes Geophys.* 14:3-4.

Kulkarni, A. and H. von Stroch. 1995. Monte Carlo experiments on the effect of serial correlation on the Mann-Kendall test of trend. *Meteorol. Z.* 4(2):82-85.

Kvålseth, T.O. 1985. Cautionary note about R^2. *The American Statistician* 39(4):279-285.

Labat, D. 2005. Recent advances in wavelet analyses: Part 1. A review of concepts. *Journal of Hydrology* 314:275-288.

Lane, D.L. 2007. *HyperStat Online Statistics Textbook.* http://davidmlane.com/hyperstat/index.html.

Leffler, R., and K. Redmond. 2004. *PCU6 – Unit No. 2 Factors Affecting the Accuracy and Continuity of Climate Observations.* National Weather Service, NOAA. 40 p. http://www.weather.gov/om/csd/pds/PCU6/IC6_2/tutorial1/PCU6-Unit2.pdf (see corresponding tutorial at: http://www.weather.gov/om/csd/pds/PCU6/IC6_2/tutorial1/Factors.htm).

Legates, D.R. 1995. Global and terrestrial precipitation: A comparative assessment of existing climatologies. *International Journal of Climatology* 15:237-258.

Lettenmaier, D.P., E.F. Wood, and J.R. Wallis. 1994. Hydro-climatological trends in the continental United States, 1948-88. *J. Clim.* 7:586-607.

Linkin, M.E., and S. Nigam. 2008. The North Pacific Oscillation–West Pacific Teleconnection Pattern: mature-phase structure and winter impacts. *Journal of Climate* 21(9):1979-1997.

Makkonen, L. 2008. Problems in the extreme value analysis. *Structural Safety* 30:405–419.

Malmgren, B.A., A. Winter, and D. Chen. 1998. El Niño–Southern Oscillation and North Atlantic Oscillation control of climate in Puerto Rico. J. Climate 11:2713–2717.

Mann, M.E. 2004. On smoothing potentially non-stationary climate time series. *Geophys. Res. Lett*. 31, L07214, doi:10.1029/2004GL019569.

Mantua, N.J., S.R. Hare, Y. Zhang, J.M. Wallace, and R.C. Francis. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of American Meteorological Society* 76:1069–1079.

Marchand, P.J. 1996. *Life in the Cold.* 3rd ed. University Press of New England, Hanover, NH. 304 p.

Marlon, J. R., P. J. Bartlein, M.K. Walsh, S.P. Harrison, K.J. Brown, M.E. Edwards, P.E. Higuera, M.J. Power, R.S. Anderson, C. Briles, A. Brunelle, C. Carcaillet, M. Daniels, F.S. Hu, M. Lavoie, C. Long, T. Minckley, P. J.H. Richard, A.C. Scott, D.S. Shafer, W. Tinner, C.E. Umbanhowar, Jr, and C. Whitlock. 2009. Wildfire responses to abrupt climate change in North America. *Proc. National Academy of Science* 106: 2519-2524. Supporting information: www.pnas.org/cgi/content/full/0808212106/DCSupplemental.

Martinez-Yrizar, A., and J. Sarukhan. 1990. Litterfall patterns in a tropical deciduous forest in Mexico over a five-year period. *Journal of Tropical Ecology* 6(4):433-444.

Matsuo, T. 2005. Nonstationary covariance modeling using wavelets. Institute for Mathematics Applied to Geosciences (IMAGe), National Center for Atmospheric Research, Boulder CO. http://www.image.ucar.edu/GSP/Projects/ResearchNuggets.shtml#WaveletCovariance.

McCabe, G.J., M.A. Palecki, and J.L. Betancourt. 2004. Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences* 101(12):4136-4141.

McDonald, J.H. 2009. *Handbook of Biological Statistics.* 2nd ed. Sparky House Publishing, Baltimore, Maryland. http://udel.edu/~mcdonald/statintro.html and http://www.lulu.com/content/3862228.

McKee, T.B., N. J. Doesken, and J. Kliest. 1993. The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference of Applied Climatology, 17-22 January, Anaheim, CA.* American Meteorological Society, Boston, MA. p. 179-184.

Mearns, L.O., R.W. Katz, and S.H. Schneider. 1984. Extreme high temperature events: Changes in their probabilities with changes in mean temperature. *J. Climate Appl. Meteor.* 23:1601–1613.

Meehl, G.A., et al. [16 coauthors]. 2000. An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bull. Amer. Meteor. Soc.* 81:413–416.

Meek, D.W., and J.L. Hatfield. 1994. Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology* 69:85-109.

Menne, M.J., and C.N. Williams, Jr. 2005. Detection of undocumented change points using multiple test statistics and composite reference series. *J. Climate* 18:4271-4286.

Minobe, S. 1997. A 50-70 year climatic oscillation over the North Pacific and North America. *Geophysical Research Letters* 24:683-686.

Minobe, S. 1999. Resonance in bidecadal and pentadecadal climate oscillations over the North Pacific: Role in climatic regime shifts. *Geophys. Res. Lett.* 26:855-858.

Minobe, S., and N. Mantua. 1999. Interdecadal modulation of interannual atmospheric and oceanic variability over the North Pacific. *Progress in Oceanography* 43:163–192.

Miranda, P.M.A., and A.R. Tomé. 2009. Spatial structure of the evolution of surface temperature (1951–2004). *Climatic Change* 93:269-284.

Muggeo, V.M.R. 2009. *Segmented: Segmented Relationships in Regression Models*. R Package. http://cran.es.r-project.org/web/packages/segmented/segmented.pdf.

PACN. 2008. *Climate Monitoring Protocol for the Pacific Island Network (PACN), Standard Operating Procedure (SOP) # 24: Data Analysis - Instructions for Graphs and Calculations.* Version 1.0 (June 13, 2008). Pacific Island Network, National Park Service, USDI. 11 p. Unpublished document.

Palmer, W.C. 1965. *Meteorological Drought.* Research Paper No. 45, US Department of Commerce Weather Bureau, Washington, DC.

Parlange, M.B., and R.W. Katz. 2000. An extended version of the Richardson model for simulating daily weather variables. *Journal of Applied Meteorology* 39:610-622.

Pederson, G., S. Gray, D. Fagre, and L. Graumlich. 2006. Long-duration drought variability and impacts on ecosystem services: a case study from Glacier National Park, Montana. *Earth Interactions* 10(1):1-28.

Peterson, E.W., and L. Hasse. 1987. Did the Beaufort scale or the wind climate change. *J. Physical Oceanogr*. 17:1071-1074.

Peterson, T.C., and D.R. Easterling. 1994. Creation of homogeneous composite climatological reference series. *Int. J. Climatol*. 14:671–679.

Peterson, T.C., et al. 1998a. Homogeneity adjustments of *in situ* atmospheric climate data: A review. *Int. J. Climatol.* 18:1493-1517.

Peterson, T.C., R. Vose, R. Schmoyer, and V. Razuvaëv. 1998b. Global historical climatology network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.* 18:1169-1179.

Quayle, R.G., D.R. Easterling, T.R. Karl, and P.Y. Hughes. 1991. Effects of recent thermometer changes in the cooperative station network. *Bull. Am. Meteorol. Soc.* 72:1718-1724.

Redmond, K., G. McCurdy, and G. Kelly. 2008. *NPS Climate Data and Monitoring Options.* Report to the National Park Service. NPS Evaluation Report 20080518 / WRCC Report 08-01. Western Regional Climate Center, Reno, NV. 45 p.

Rehfeldt, G.E., N.L. Crookston, M.V. Warwell, and J.S. Evans. 2006. Empirical analyses of plant-climate relationships for the western United States. *International Journal of Plant Sciences* 167:1123-1150.

Richardson, C.W. 1981. Stochastic simulation of daily precipitation, temperature and solar radiation. *Water Resources Research* 17:182-190.

Rodionov, S.N., and J.E. Overland. 2005. Application of a sequential regime shift detection method to the Bering Sea ecosystem. *ICES J. Mar. Sci*. 62:328-332.

Running S.W., R.R. Nemani, and R.D. Hungerford. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Can. J. For. Res.* 17:472-483.

Schlesinger, M.E., and N. Ramankutty. 1994. An oscillation in the global climate system of period 65-70 years. *Nature* 367:723–726.

Schreuder, H.T., R. Ernst, and H. Ramirez-Maldonado. 2004. *Statistical Techniques for Sampling and Monitoring Natural Resources.* Gen. Tech. Rep. RMRS-GTR-126. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 111 p. Available online: http://www.fs.fed.us/rm/pubs/rmrs_gtr126.html.

Schumm, S.A. 1991. *To Interpret the Earth: Ten Ways to be Wrong.* Cambridge University Press. 133 p.

Schwing, F.B., T. Murphree, and P.M. Green. 2002. The Northern Oscillation Index (NOI): a new climate index for the northeast Pacific. *Progress in Oceanography* 53:115-139.

Schwing, F.B., R. Mendelssohn, S.J. Bograd, J.E. Overland, M. Wang, and S. Ito. 2009. Climate change, teleconnection patterns, and regional processes forcing marine populations in the Pacific. *J. Mar. Syst.* doi:10.1016/j.jmarsys.2008.11.027

Scott, A. and C. Wild. 1991. Transformations and R^2. *The American Statistician* 45:127-129.

Şen, Z., and Z. Habib. 2000. Spatial precipitation assessment with elevation by using point cumulative semivariogram technique. *Water Resources Management* 14:311-325.

Sokal, R.R., and F.J. Rohlf. 1994. *Biometry*. 3rd ed. W.H. Freeman. 880 p.

Steinskog, D..J., D..B. Tjøstheim, and N..G. Kvamstø. 2007. A cautionary note on the use of the Kolmogorov–Smirnov Test for normality. Monthly Weather Review 135:1151–1157.

Stenseth, N. C. and A. Mysterud. 2005. Weather packages: finding the right scale and composition of climate in ecology. *Journal of Animal Ecology* 74:1195–1198.

Stenseth, N. C., A. Mysterud, G. Ottersen, J.W. Hurrell, K.-S. Chan, and M. Lima. 2002. Ecological effects of climate fluctuations. *Science* 297:1292–1296.

Stenseth, N. C., G. Ottersen, J. W. Hurrell, A. Mysterud, M. Lima, , K. S. Chan, N. G. Yoccoz, and B. Ådlandsvik. 2003. Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *Proceedings of the Royal Society of London, Series B* 270:2087–2096.

Stephenson, A., and E. Gilleland. 2005. Software for the analysis of extreme events: the current state and future directions. Extremes 8:87-109.

Stephenson, D. 2002. Estimation and attribution of changes in extreme weather and climate events. In: *IPCC Workshop on Extreme Weather and Climate Events*, 11-13 June 2002, Beijing. Presentation. http://www.met.rdg.ac.uk/~han/Extremes/extreme1.pdf.

Steward, R. R. 2005. The oceanic influence on North American drought. Chapter in: *Our Ocean Planet – Oceanography in the 21st Century*. Texas A&M University. On-line textbook: http://oceanworld.tamu.edu/resources/oceanography-book/oceananddrought.html.

Stohlgren, T.J., T.N. Chase, R.A. Pielke, Sr., T.G.F. Kittel, and J. Baron. 1998. Evidence that local land use practices influence regional climate, vegetation, and stream flow patterns in adjacent natural areas. *Global Change Biology* 4:495-504.

Thompson, D.W.J., and J.M. Wallace. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate* 13:1000-1016.

Thornton, P.E., H. Hasenauer, and M.A. White. 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application of complex terrain in Austria. *Agricultural and Forest Meteorology* 104:255-271.

Toews, M.W., P.H. Whitfield, and D.M. Allen. 2007. Seasonal statistics: The 'seas' package for R. *Computers and Geosciences* 33:944-951.

Tomé, A.R., and P.M.A. Miranda. 2004. Piecewise linear fitting and trend changing points of climate parameters. *Geophys. Res. Lett.* 31:L02207.

Torrence, C. and G. P. Compo. 1998. A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.* 79:61–78

Trenberth, K.E. 1984. Signal versus noise in the Southern Oscillation. *Monthly Weather Review* 112:326-332.

Trenberth, K.E. 1997. The definition of El Niño. *Bull. Am. Meteorol. Soc.* 78:2771–2777.

Trenberth, K.E. 1998. Atmospheric moisture residence times and cycling: Implications for rainfall rates with climate change. *Climatic Change* 39:667-694.

Trenberth, K.E., and J.W. Hurrell. 1994. Decadal atmosphere-ocean variations in the Pacific. *Climate Dyn.* 9:303-319.

Trenberth, K.E., A. Dai, R.M. Rasmussen, and D.B. Parsons. 2003. The changing character of precipitation. *Bull. Am. Meteorol. Soc.* 84:1205–1217.

Trenberth, K.E., P.D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J.A. Renwick, M. Rusticucci, B. Soden and P. Zhai. 2007. Observations: surface and atmospheric climate change. In: Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.). *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA.

Urban, D.L., R.V. O'Neill, and H.H. Shugart. 1987. Landscape ecology. *Bioscience* 37:119-127.

Verzani, J. 2004. *Using R for Introductory Statistics.* Chapman & Hall/CRC. 432 p. In part on Google Books: http://books.google.com/books?id=jwolc192c5kC.

Verosub, K.L., and J. Lippman. 2008. Global impacts of the 1600 eruption of Peru's Huaynaputina Volcano. *Eos* 89:141-142.

von Storch, H., and F.W. Zwiers. 2001. *Statistical Analysis in Climate Research.* Cambridge University. 484 p. In part on Google Books: http://books.google.com.br/books?id=5QgAfL1N6koC.

Vose, R.S., C.N. Williams Jr., T.C. Peterson, T.R. Karl, and D.R. Easterling. 2003. An evaluation of the time of observation bias adjustment in the US Historical Climatology Network. *Geophysical Research Letters* 30(20):2046, pp. CLM3-1 to 3-4.

Wade, C.G. 1987. A quality control program for surface mesometeorological data. *J. Atmos. Oceanic Technol.* 4:435–453.

Wallace, J.M., and D.S. Gutzler. 1981. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.* 109:784–812.

Wessa, P. 2009. *Free Statistics Software (v1.1.23-r1).* Office for Research Development and Education, Resa Corporation R&D: http://www.wessa.net/.

Whitcher B., P. Guttorp, and D.B. Percival. 2000. Wavelet analysis of covariance with application to atmospheric time series. *Journal of Geophysical Research-Atmospheres* 105(D11):14941-14962.

Wilks, D. 2006. *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press / Elsevier. Available from: http://web.unbc.ca/~ytang/text_book.pdf.

Willett, J. B., and J. D. Singer. 1988. Another cautionary note about R^2: its use in weighted least-squares regression analysis. *The American Statistician* 42:236-238.

Willmott, C. J., C. M. Rowe, and W. D. Philpot. 1985. Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer* 12:5-16.

Wolter, K., and D. Allured. 2007. New climate divisions for monitoring and predicting climate in the U.S. *Intermountain West Climate Summary* 3(5):2-6 (June 2007). http://wwa.colorado.edu/IWCS/archive/IWCS_2007_Jun.pdf.

Wolter, K. and M.S. Timlin. 1993. Monitoring ENSO in COADS with a seasonally adjusted principal component index. *Proc. of the 17th Climate Diagnostics Workshop*, Norman, OK, NOAA/N MC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, pp. 52-57.

Wright, S.J., and O. Calderón. 2006. Seasonal, El Niño and longer term changes in flower and seed production in a moist tropical forest. *Ecology Letters* 9:35–44.

Yandell, B.S. 1997. *Practical Data Analysis for Designed Experiments.* CRC Press. 437 p. In part on Google Books: http://books.google.com/books?id=K-e05Mrab0oC.

Yang, D., D. Kane, Z. Zhang, D. Legates, and B. Goodison. 2005. Bias-corrections of long-term (1973-2004) daily precipitation data over the northern regions. *Geophysical Research Letters* 32:L19501.

Yiou, P., E. Baert, and M.F. Loutre. 1996. Spectral analysis of climate data. *Surveys in Geophysics* 17:619-663.

You, J.S., K.G. Hubbard, S. Nadarajah, and K. Kunkel. 2007. Performance of quality assurance procedures on daily precipitation. *Journal of Atmospheric and Oceanic Technology* 24:821-834.

Sierra Nevada Network Climate Reporting Protocol

SOP 1: Station Selection

Version 1.0

This SOP is part of the Sierra Nevada Network Climate Reporting Protocol, but is designed to be printed and viewed as a separate document. This document has been adopted from 'Rocky Mountain Climate Protocol SOP Station Selection' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Introduction

This SOP describes the process and criteria used to select stations from which we will acquire, analyze, and report data in annual status reports and periodic trend reports. First, we describe general considerations for station selection and their application within SIEN. We follow with selection criteria specific to each relevant program (e.g., NWS-COOP stations) and the results of their application by SIEN. Appendix B contains a list of all weather stations in the SIEN as well as the final list of stations selected for use by this protocol. On occasion it is necessary to relax the criteria in order to achieve adequate station coverage for a region or time frame, and these exceptions are noted. As new stations are added in the future by other organizations, we will follow this SOP to determine if and when data from the new station will be used by this protocol. After applying objective criteria listed in this SOP, SIEN will also seek expert opinion to confirm that selected stations will contribute to meeting the goals of the protocol.

General Approach to Station Selection

Below we list station selection criteria appropriate for various types of analyses. The strictest criteria apply to stations supplying data for temporal analyses. These criteria originate from Kittel et al. (2010) and discussions with the Western Regional Climate Center (Edwards and Redmond *in prep*). Criteria were developed to assure adequate temporal extent, high data quality, and maximize spatial coverage to the extent possible. Even under the best of circumstances, obtaining, analyzing and summarizing climatic records can be difficult. Stations will be selected based on their ability to contribute to one or both of the following objectives:

1. To provide an extensive *spatial* picture of climate across the network and its conterminous climate domain. This is a key part of describing the year in review in the climate status reports.
2. To provide for a *temporal* analysis of interannual variability, regime shifts, and long-term trends in key daily, monthly, seasonal, and annual variables. This is primarily restricted to trend reports although status reports will compare annual results to the 30-year average.

Stations for Temporal Analysis

The strictest selection criteria identify daily and monthly stations suitable for temporal analyses. These criteria are largely based on Gray (2008). Stations in this analysis group should meet the following requirements:

- A continuous record having gaps no more than one year (per Gray 2008).

- A record of at least 25-40 years depending on the variable and time step; at least 50 years are needed for regime shift analysis. The longer period for regime shift analyses is to limit analyses to records most likely to capture multi-decadal shifts. However, 50 years is an arbitrary cutoff and can be relaxed. When selecting stations to represent a given climate zone, stations with longer records will be given preference over others. For monthly temperature and precipitation (COOP), records will ideally cover the 1971-2000 period, the most recent period for calculating ‘30-year climate normals.’ This period is the baseline for statistical comparisons.

- Relatively high quality – that is, records with manageable data issues (biases, errors, inhomogeneities, missing values). See Gray (2008; Data Issues Report, p. 1-7) for specifics.

- Providing representation for the climate zone or for a park. Stations with problematic records but in critical locations will warrant a high level of effort to create a ‘clean’ data series.

- Finally, as the “best of the best,” stations will be narrowed down to key stations for temporal analysis. This is to make quality control and correction tasks manageable during the 5-year Trends Report cycle.

We have determined that six of the seven long-term COOP stations within the SIEN boundaries meet these criteria; this is a manageable number for annual status. These six stations will be included in both annual status and periodic trends reports, and the long-term trends analyses may incorporate additional NWS-COOP stations near the parks’ boundaries that are relevant to conditions within the parks and have records of high quality and adequate length.

Spatial Coverage Stations
Stations to provide greater geographic coverage include the ‘temporal analysis stations’ *plus* additional stations that:

- Are currently reporting monthly parameter values or provide data to compute monthly values (i.e., daily values not required).

- Contribute significantly to spatial coverage within the network – that is, are not redundant in their information about the year.

- Have sufficient quality during the previous calendar and water-years, requiring minimal quality correction or whose problems are judged by a climatologist as acceptable for the purpose of spatial analyses.

We have selected the DEPO-AWS for inclusion in reports based on the above criteria. Although DEPO is within the same climate zone as the other SIEN parks, the geographic location causes the climate at DEPO to be unique compared to nearby stations. For example, a rain shadow can form at DEPO despite the park’s location west of the Sierra crest. Nearby stations with long records, such as the stations on Mammoth Mountain, may not experience the same climatic conditions as those at DEPO. Therefore, although the period of record at the DEPO station is short (2006-present) and nearby stations may have longer periods of record, SIEN will rely on records from the DEPO-AWS station to report annual climate status that is relevant to DEPO resource management.

Additionally, we adopt one of the long-term NWS-COOP stations, Tuolumne RS, for inclusion under the above criteria. The station has gone through periods of personnel changes and was classified as an official station in 2008, although the records extend back to 1978. The station contributes significantly to the spatial coverage of the SIEN desired for reporting. However, portions of the record may be suspect and it may not be possible compare yearly values to an historical “normal”.

Review of Station Selection

Annual review of stations will be necessary as some stations are dropped and others added to climate monitoring networks or as data quality issues are encountered. The SIEN Physical Scientist will examine the data completeness of all stations selected for reporting each year and determine whether each station has adequate data quality for inclusion in each annual status report. During the first years of implementing the protocol, selection may need refining. Previously unidentified data problems will arise and will need to be resolved with a workable solution or by removing a station from the selection pool.

Station Selection Guidelines by Program

The specific selection criteria below originate from Gray (2008) and are designed to first identify a network of stations that provide good spatial coverage within each park or network and then screen them to produce a set of stations with the highest quality data. Limiting the number of stations reduces the workload involved to acquire, process, review, analyze, and report on station data. The intention is to obtain a high-quality, representative data set, but not to overburden report writers with an unmanageable number of stations.

Station Selection Criteria Used for NWS-COOP Stations

COOP weather stations are generally accepted to be the best of the nationwide weather station networks in terms of data completeness and quality and metadata documentation. In the absence of a formal data and metadata review for each station, we assume that the data quality is good unless otherwise noted.

Representation of Intraregional Variability

First we considered stations that span the variability within and near the SIEN parks by mapping all COOP stations and listing each station’s history and elevation.

Data Quality and Availability

Next we reviewed the length, quality and completeness of potential station records and applied the following requirements:

1. Keep only those stations with > 30 years of continuous record (i.e., no gaps > 1 year)
2. Keep only those stations that have operated continuously between 1971 and 2000. (Criteria 1 and 2 reflect the need to calculate 30-year averages or so-called “climatic normals” for these stations).
   SIEN note: This requirement was not met for one COOP station, Tuolumne Meadows RS. However, the station has a relatively long and complete period of record and the station provides coverage in an area that is otherwise not addressed spatially and for which there is high management interest. Therefore, the station will be included in the database and annual status reports.
3. Station records must include daily min/max temperature and precipitation.
4. Station records must not contain more than 5% missing data since January 1, 1990.
5. Stations may not have known siting issues or recurring instrumentation problems
6. Stations must provide year-round observations.

Expert Input

While objective criteria such as data completeness and length-of-record provided the basis for an initial screening of candidate sites, expert input played a critical role in the final selections. Experts were consulted to review potential sites based on the following criteria:

1. Siting. Based on site visits, plus local knowledge and/or photographs, the location of each station should be reviewed to determine if the station is reasonably well-sited (e.g., away from heat sources) and maintained.

2. Station history. Station histories from National Climatic Data Center (http://mi3.ncdc.noaa.gov/) and “Cooperative Network Station Reports (B-44)” should be reviewed for station moves and potential siting and/or equipment problems. Expert knowledge of station histories could also be considered.

3. Representation of intra-regional variability. Expert input on the distribution of sites and their ability to capture regional variability should also be considered.

Station Selection Criteria for California Water Resources Snow Courses

While snow course data throughout the U.S. are commonly served through the Natural Resources Conservation Service (NRCS), the snow courses in California are primarily operated by the California Department of Water Resources (CDWR) and served through CDWR’s data system – California Data Exchange Center (CDEC). The Inventory and Monitoring Division (IMD) has the ability to acquire data from NRCS and store it in the Climate Database, however they do not have the ability to access data from CDEC. Therefore, it falls to SIEN to acquire any snow course data for analysis and reporting. Currently, there are 98 snow courses (some also have automated snow pillows) and five snow pillows (not associated with a snow course) in the major SIEN watersheds (43 within the park boundaries. See Table B.2 in the SIEN Climate Protocol Appendix B for a list of these stations. SIEN will download snowcourse data for all stations in the SIEN watersheds from CDEC, then process, format, and upload it to the IMD Climate Database following instructions provided in IMD SOP: Submitting Data for Upload (Frakes 2012).

Representation of Intraregional Variability

Variability in snow depth between individual snowcourses because of elevation, aspect and microsite differences preclude a meaningful comparison between snowcourses. Instead, for annual status reports, SIEN will rely on the watershed averages, which incorporate all stations within a watershed, provided through the CDEC website. Thus, all snow courses within the SIEN watersheds are included in reporting.

Data Quality and Availability

Most snow courses in the SIEN have records of at least 50 years. The snow course measurements are done manually by trained individuals and the data can be assumed to be of high quality.

Station Selection Criteria for USGS Streamgage Stations

All streamgages reported by the USGS through their website, the National Water Information Service (NWIS), can be considered of high quality with extensive QA/QC of the data and maintenance of equipment. There are five stations reported through NWIS within the major SIEN watersheds, three with records of greater than 90 years and two that have short records.

Those with long records will be incorporated in annual status and periodic trend reports. For the purposes of annual status reports, we do not consider a long record critical for inclusion. Although it will not be possible to compare hydrologic statistics, such as peak discharge, to historic averages at those stations with short records, the annual hydrographs at the stations demonstrate timing of snowmelt and other climatic events within the year.

Representation of Intra-regional Variability

By including all USGS streamgages, regardless of the length of record, we are able to report on yearly conditions in four of the seven major SIEN watersheds.

Data Quality and Availability

Gages with > 30 years of continuous record are an appropriate target for period trends reports. Generally speaking, data completeness is far less of an issue with streamgages than with NWS-COOP stations. Automated and telemetered gages reduce the chances for missing data. In addition, USGS estimates missing data using a robust, well-documented procedure, Gages with less than 30 years of data remain valuable for status reporting.

Literature Cited

Edwards, L. M., and K. T. Redmond. 2011. Climate assessment for Sierra Nevada Network Parks. NPS/SIEN/NR-2011/482. National Park Service, Fort Collins, Colorado. Available from: http://science.nature.nps.gov/im/units/sien/monitoring/Reports/SIENClimateAssessment_EdwardsRedmond_Final_20111230.pdf (accessed 26 March 2012).

Frakes, B. 2012. Submitting data for upload to the I&M Enterprise climate database. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166928 (accessed 30 March 2012).

Gray, S. T. 2008. Framework for linking climate, resource inventories, and ecosystem monitoring. Natural Resource Technical Report NPS/SIEN/NRTR—2008/110. National Park Service, Fort Collins, Colorado.

Kittel, T., S. Ostermann-Kelm, B. Frakes, M. Tercek, S. Gray, and 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, Montana, 7–8 April 2009. Unpublished final report (18 February 2010). National Park Service, Greater Yellowstone Network, Bozeman, Montana.

Rocky Mountain Climate Working Group. 2011. Standard operating procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sierra Nevada Network Climate Reporting Protocol

SOP 2: Quality Control

Version 1.0

This SOP is part of the Sierra Nevada Network Climate Reporting Protocol, but is designed to be printed and viewed as a separate document. This document has been adopted from 'Rocky Mountain Climate Protocol SOP Quality Control' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Figures

Page

Tables

Page

Data Quality Control for Status and Trend Reporting

Data quality control requirements and activities for annual status reports are limited in scope pending the more comprehensive and time-consuming expert review required for data used in Climate Variability and Trend reports. Data reported annually are first downloaded from the climate database, then reviewed for obvious errors and problems according to the standards provided in Table SOP 2.1 and Figure SOP 2.1, which are from Kittel et al. (2010). To prepare data for semi-annual variability and trend reporting a climatologist is required to perform the more rigorous quality control activities provided in Kittel et al. (2010), which is further described in “SIEN SOP 6: Trend Analysis and Reporting” (Tables SOP 2.2 and SOP 2.3, Figure SOP 2.2).

Table SOP 2.1. Quality control for annual status reports[1].

Input	Event Structure	Temporal or Spatial	Data Errors, Biases, Outliers, and Multiday Observations	Inhomogeneities		Completeness		Network Dataset Product
				Known	Unknown	Daily Data	Monthly Data	
COOP	Daily	Temporal	As for monthly COOP dataset, *except*: Multiday observations omitted	As for COOP seasonal	As for COOP seasonal	Missing values *not* infilled	*N/A*	Daily dataset (*Provisional*)
	Seasonal Inter-annual variability	Spatial Temporal	Numerical & visual checks Consult forms and observer Manual removal Multiday obs crossing month boundary parsed (e.g., ppt[1]) or omitted	Accept source QC Known but not corrected by source - only correct if easily done If not, document for next *Trends* report	Accept source QC Attach caveats to results	Missing values – remove month if value > threshold: - missing T >5d - missing ppt >3d - other vars: 15% missing	Accept source QC, document missing months Attach caveats to results	Monthly-Seasonal-Annual Dataset (*Provisional*)
Streamflow	Daily Seasonal Inter-annual variability	Spatial Temporal	Numerical & visual checks Manual removal	*As for COOP*	*As for COOP*	Accept source QC	*N/A*	Unregulated Stations *(Provisional)*
PRISM	Seasonal	*N/A*	Accept source QC	Accept source QC	Accept source QC	*N/A*	Accept source QC	Monthly-Seasonal-Annual Dataset (*Provisional*)

[1]Abbreviations include: ppt = precipitation. T = temperature. QC = quality control. N/A = not applicable

Table SOP 2.2. Quality control for variability and Trends report, temporally representative stations only[1]. See Figures SOP 2.1 and SOP 2.2 for more detailed instructions.

Input	Event Structure	Data Errors, Biases, Outliers, and Multiday Observations	Inhomogeneities		Completeness		Network Dataset Product
			Known	Unknown	Daily Data	Monthly Data	
COOP	Daily	**Apply high-quality QC** Numerical & visual checks Consult forms and observer Manual removal Multiday observations omitted	**Apply high-quality QC**	**Apply high-quality QC**	Missing values not infilled	*N/A*	Daily dataset (*Final*)
	Inter-annual variability	As for Daily COOP, *except*: Multiday observations parsed (ppt) or omitted, then infilled	**Apply high-quality QC**	**Apply high-quality QC**	**Apply high-quality QC** Missing values infilled	If only monthlies available – Accept source QC, document missing months Attach caveats to results	Monthly-Seasonal-Annual Dataset (*Final*)
SNOTEL	Inter-annual variability	Numerical & visual checks Manual removal	*As for COOP*	*As for COOP*	Accept source QC	*N/A*	*Final*
Streamflow	Daily	**Apply high-quality QC** Numerical & visual checks Manual removal Multiday observations omitted Check data gone from provisional to official Consult observer	**Apply high-quality QC**	**Apply high-quality QC**	Missing values not infilled	*N/A*	Daily dataset *(Final)*
	Inter-annual variability						Monthly dataset *(Final)*

[1]Abbreviations include: ppt = precipitation. T = temperature. QC = quality control. N/A = not applicable

Table SOP 2.3. Quality control for decadal region shifts, long term trends, and teleconnections, the latter two of which are treated in the same manner concerning inter-annual variability [1].

All datasets	Decadal Regime Shift	Regime-shift subset of Monthly-Seasonal-Annual Dataset (*Final*)	Correct only with great care – will affect shift detection	Ignore possibility of artificial inhomogeneities not supported by station histories or other data. Attach caveats to results	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability
Long-term trends	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability
Indices	*N/A*	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability	As for inter-annual variability

[1]Abbreviations include: ppt = precipitation. T = temperature. QC = quality control. N/A = not applicable

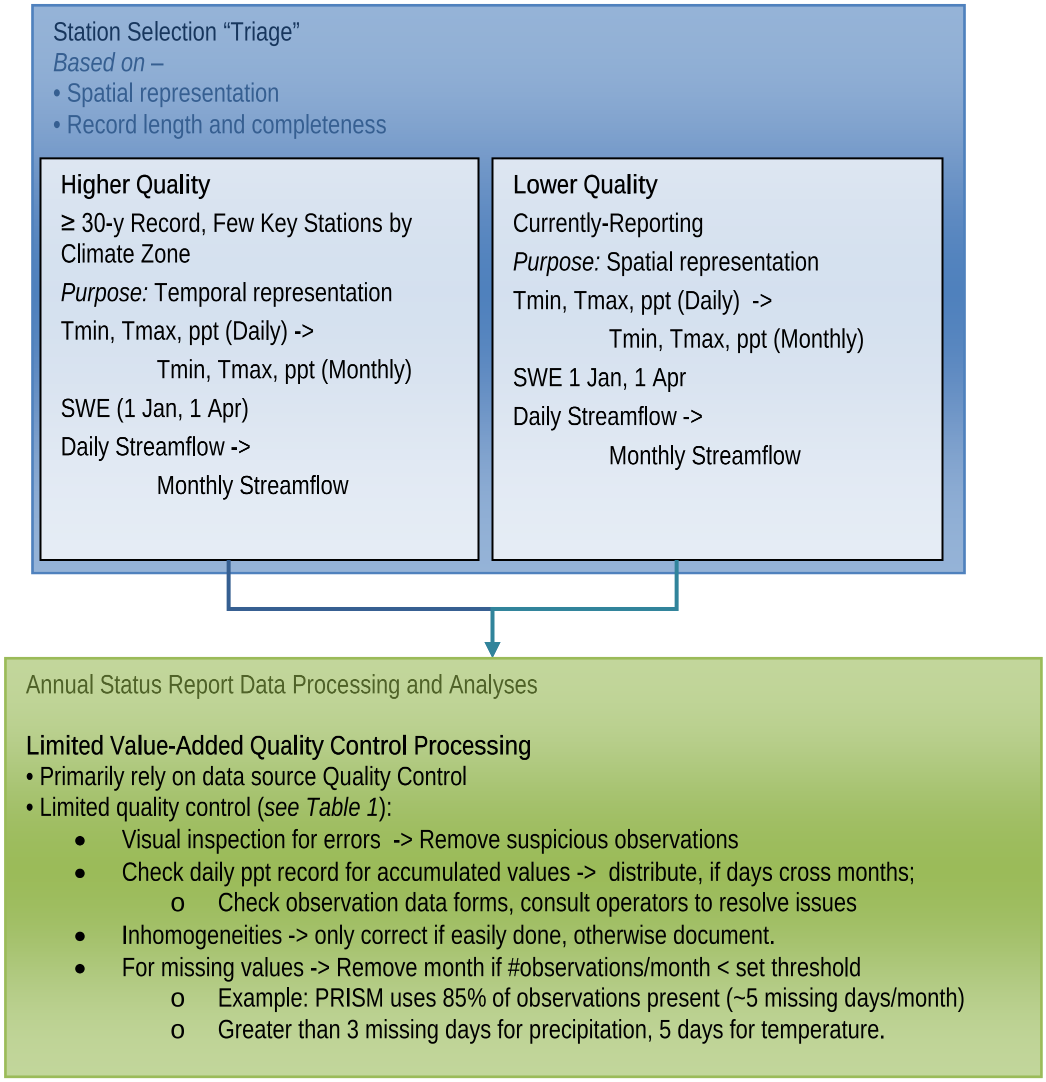


Figure SOP 2.1. Process framework for annual climate status report using limited-value quality control. Refer to Kittel et al (2010) for supporting details.

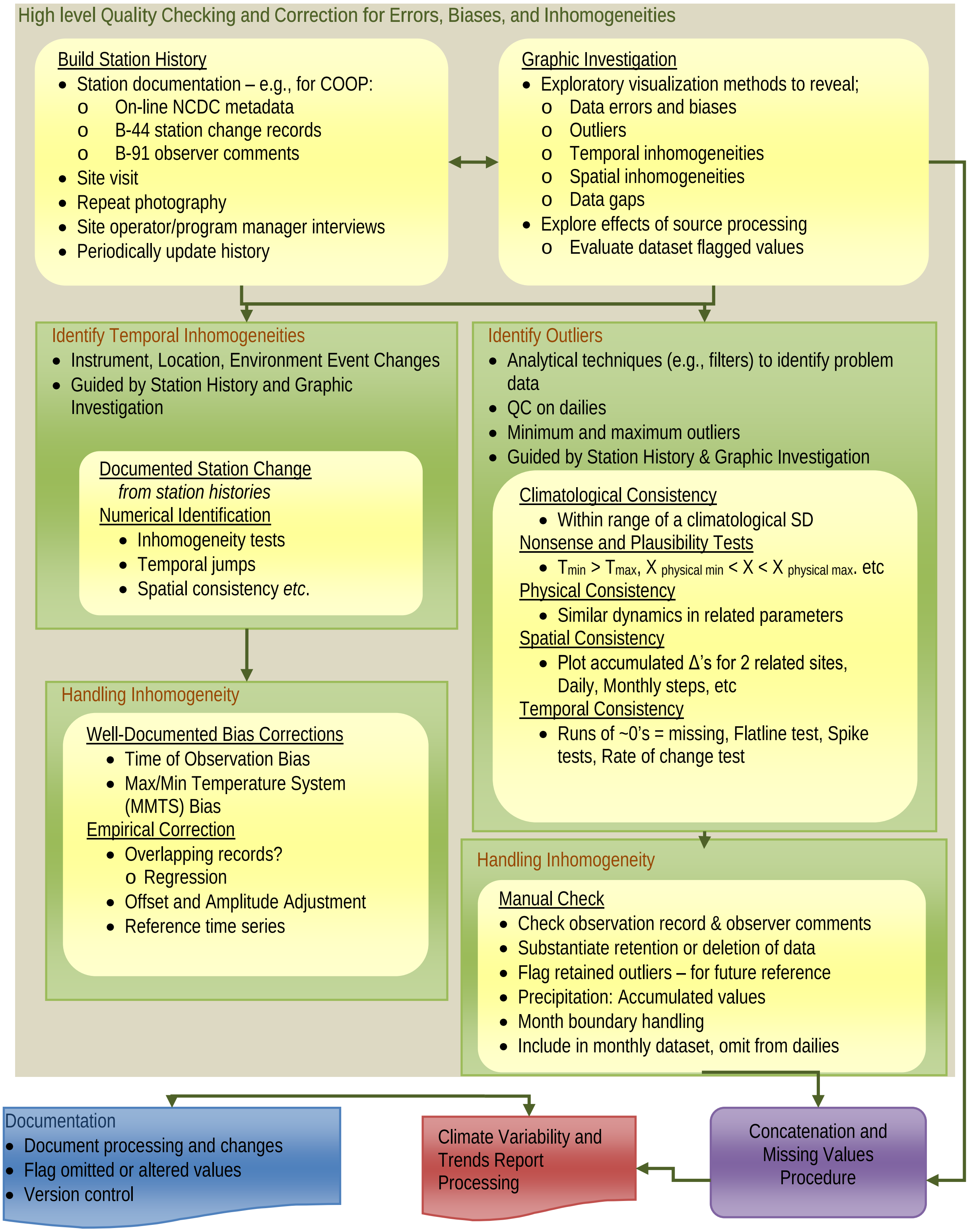


Figure SOP 2.2. Processing framework for high-level quality control of the climate datasets. Refer to Kittel et al (2010) for supporting details.

Literature Cited

Kittel, T., S. Ostermann-Kelm, B. Frakes, M. Tercek, S. Gray, and 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, Montana, 7–8 April 2009. Unpublished final report (18 February 2010). National Park Service, Greater Yellowstone Network, Bozeman, Montana.

Rocky Mountain Climate Working Group. 2011. Standard operating procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sierra Nevada Network Climate Reporting Protocol

SOP 3: Tabular Data Acquisition and Processing

Version 1.0

This document has been adopted from 'Rocky Mountain Climate Protocol SOP Data Acquisition' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Figures

Page

Tables

Page

Introduction

The Sierra Nevada Climate protocol relies entirely on data acquired from established weather and climate monitoring systems operated by the federal and academic entities listed in Table SOP 3.1. This SOP describes the process for acquiring and managing these data, some of which will be addressed by the NPS Inventory and Monitoring Division (IMD), and some of which will be addressed by SIEN. Tabular climate data from National Weather Service Cooperative Observer (NWS-COOP) stations will be acquired and managed within a central database run by IMD. This SOP describes the steps to ensure that IMD incorporates climate data for SIEN parks into the database. Procedures for downloading data directly from provider websites are listed which can be used in cases where the IMD cannot acquire the data for the network. SIEN is responsible for the management and analyses of gridded datasets (PRISM), and SOP 4 describes where and how to download raster data.

Listed here are the processes that will be undertaken during the first years of protocol implementation with the assumption that the IMD database is in operation. We also include the processes that would be undertaken to acquire NWS-COOP data if the IMD database were not in operation. Additionally, we include instructions for acquiring supplemental data that may be useful for periodic trend reports.

Table SOP 3.1. Sources of climate data to be obtained for the SIEN climate program. IMD will acquire data from climate monitoring programs/systems in bold text. SIEN will acquire the data or data products from the other programs and either store locally or upload to the Climate Database.

Agency	Acronym	Data	Web Address
National Oceanic and Atmospheric Administration, National Climatic Data Center	**NOAA-NCDC**	**COOP (temp and precip)**	**http://www.ncdc.noaa.gov/oa/climate/stationlocator.html** **http://www7.ncdc.noaa.gov/IPS/coop/coop.html**
California Department of Water Resources	CDWR	Snow Course	http://cdec.water.ca.gov/snow/current/snow/
U.S. Geologic Survey	USGS	Streamgage	http://waterdata.usgs.gov/nwis
Scripps Institution of Oceanography	Scripps	Temperature and Precipitation	Not available online
PRISM Climate Group	PRISM	PRISM	http://prism.oregonstate.edu/
Drought Indices and maps			http://drought.unl.edu/dm/monitor.html and http://lwf.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html

Acquiring Common Data by Request to the NPS Natural Resource Program Center

Climate data from the NWS-COOP network are acquired and managed by IMD based on a request from SIEN using the instructions found in "IMD SOP: Requesting Essential Parameter Values" (Frakes 2012a). IMD has the capability to import data from additional programs, but SIEN will only request NWS-COOP data at this time. NWS-COOP data is available through the NOAA National Climatic Data Center (NCDC) under their Global Historical Climatology Network (GHCN). Although the data is still being collected through the NWS-COOP program, the NCDC places the data under the GHCN and the "SourceName" and "SourceCode" in the Climate Database will be GHCN rather than COOP.

It is SIEN's responsibility to provide IMD with a list of all stations for which SIEN would like to have data uploaded to the database. The request must be sent by the end of April each year. If there are no changes to the previous year's request, IMD will assume the request is the same and it is not necessary to send a request. Store each new request at: J:\sien\monitoring_projects\weather_climate\data\SQL_serverDatabase.

Requesting Data

(1) Make a list of all GHCN (NWS-COOP) stations that are relevant to the network. This list can be derived from the interactive map on the NCDC website at: http://gis.ncdc.noaa.gov/map/cdo/. SIEN has determined that only stations within the parks will be uploaded to the database at this time. However, the climatologist preparing the first periodic trend report may choose to request stations outside but near the SIEN parks. The stations in Table SOP 3.2 have been requested.

Table SOP 3.2. List of CDB codes for the GHCN (NWS-COOP) stations for SIEN's annual data request to IMD.

Park	COOP_CDB_code	GHCN_CB_code	Database COOP StationCode	Database GHCN StationCode	Database StationName
SEQU	CDB_S_12459	CDB_S_75970	040343	USC00040343	ASH MTN
KICA	CDB_S_30768	CDB_S_75978	043551	USC00043551	GRANT GROVE
YOSE	CDB_S_8373	CDB_S_75979	043939	USC00043939	HETCH HETCHY
SEQU	CDB_S_7504	CDB_S_75982	045026	USC00045026	LODGEPOLE
YOSE	CDB_S_5639	CDB_S_75986	048380	USC00048380	SO ENTR YOSEMITE NP
YOSE	CDB_S_5164	NA	049063	NA	TUOLUMNE MEADOWS RS
YOSE	CDB_S_30200	CDB_S_75988	049855	USC00049855	YOSEMITE PARK HQ

(2) Download current copy of the IMD SOP: Requesting Essential Parameter Values (Frakes 2012a) and the "Wx_RequestForm.xls" from IRMA at https://irma.nps.gov/App/Reference/Profile/2166927.

Create a copy of the spreadsheet and name it Wx_RequestForm_[UnitCode]_[FirstInitialLastName].xlsx. For example, if Andi Heard were making the request for SIEN, the submitted file would be Wx_RequestForm _SIEN_AHeard.xlsx

The first four fields (CDB_Code, ParkUnits, YourName, YourEmail) must be completed in order for the request to be accepted. Do not enter values in the other fields (columns G through J) – these are for NRPC processing

Park Units refers to what units the station should be affiliated with. If there is more than one unit, they should be comma-delimited.

(3) Send request form to IMD by April 1st (as of 2011 send list to: brent_frakes@nps.gov).

(4) For all subsequent years, check the list from the previous year and submit any additions or changes in station selection by submitting the full list with changes.

Acquiring Data from the Climate Database

IMD will fill data requests near the end of April each year. The IMD SOP: Connecting to the I&M Enterprise Climate Database for the Purpose of Data Retrieval, Summary and/or Analysis (Frakes and Kingston 2011) details the steps that SIEN uses to connect to the database and utilize the queries and views to obtain data. Both tabular daily data and summary statistics such as the annual average temperature, average total precipitation, and calculated "normals" for specified reference periods can be obtained from the database. Always check IRMA for the most recent version of Frakes and Kingston (2011) and read through it.

In order to work in the Climate Database, it is important to be clear about a few things. Each year when IMD requests data from the data providers, such as NCDC, they request the entire record (not only the current year) and replace the previous record. Also, it is important to understand some of the terms used in the database. A data version is defined as a time series of observations for a given parameter at a given station. For example, precipitation observed at Ash Mountain (COOP Station 040343) from 1931-01-01 to 1946-01-30 is considered a version and has a unique version ID of 10357. This version ID will change each year, even for the same station and parameter.

The first step towards acquiring data is to query the database to determine the version numbers for each parameter of interest and record those version numbers, which will be used to perform other queries. Queries are also referred to as "table-valued functions". SIEN maintains a "database cheatsheet" with joined and modified queries from the IMD SOP (J:\sien\monitoring_projects\climate\data\tabular\ClimateDatabase).

1) Open SQL Server Management Studio Express
2) Log into the Climate database with the password provided in the Accessing Climate Data SOP. Check often for the most recent version of the Accessing Daa SOP on NRInfo: https://nrinfo.nps.gov/Reference.mvc/Profile?code=2167699
   Server name is: INP2300FCSGUMO6\GIS_Working_Data
   Password: cl1m@te
3) Open Databases/D_Climate_DB.
4) Click on "New Query" in upper left corner.

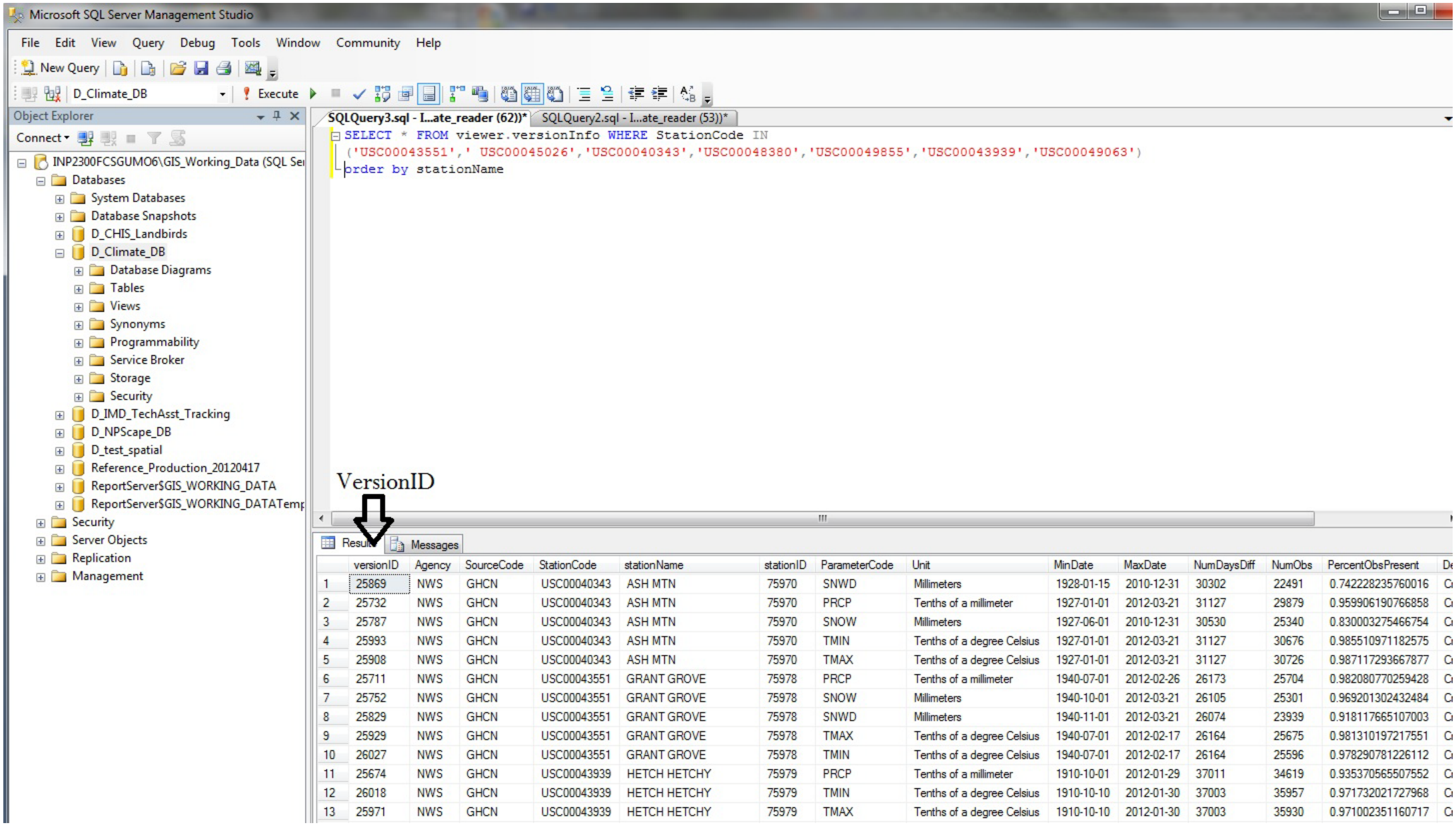


Figure SOP 3.1. An example query to obtain the versionIDs and period of record for the SIEN stations.

5) To check the period of record available in the database and retrieve the versionID for each station and parameter of interest, enter the query to look for version info from the cheatsheet into the window and click “Execute” (Figure SOP 3.1).
6) Start a new Excel spreadsheet to record the information obtained from the database as you go through the steps. Copy and paste the version numbers into the spreadsheet.
7) Perform the QA/QC steps as outlined in the next section, then return to this list.
8) Submit any corrected data to IMD to be uploaded to the database. Ask IMD to contact you when they’ve uploaded the corrected data.
9) Use queries from Frakes and Kingston (2011) and the SIEN cheatsheet to obtain the monthly and annual min, mean, and max temperature and precipitation. Also obtain the total monthly snowfall from each station. Record these in the working spreadsheet.
10) Use SigmaPlot or Excel to create figures for the annual status report. Refer to SOP 5 for the suggested content of status reports.

The database queries aren’t included here because they are frequently being modified with assistance from IMD.

Performing QA/QC on Data from the Climate Database

After the versions have been recorded, the next step is to check for missing records for the current year. It is should be noted that IMD performs some basic QA/QC when the records are acquired from the provider. IMD will scan for missing records and values that are abnormal. However, it is likely that not all errors will be caught by IMD. Further, SIEN may need to make some phone calls and obtain copies of the original data forms to correct some data. For example, for some SIEN stations, the data is recorded from the COOP station by an individual from interpretation who will relay the data to someone in the dispatch office who will type the data into the NWS-COOP webpage. The dispatch staff may not get the data entered into the website even though the data was actually collected.

First, run the query to identify missing records for each version. Record the number of missing days and the dates that are missing on the working spreadsheet. Next, use the “saved views” identified in Frakes and Kingston (2011) to look for instances of observation values that are outside the expected range or minimum daily temperatures that are higher than the max.

Now, obtain the B-91 forms from the COOP website: http://www7.ncdc.noaa.gov/IPS/coop/coop.html?foreign=false&_page=0&jsessionid=BA2FE892A96CA766C581D36CE57F1B58&state=CA&_target1=Next+%3E. The website will allow for immediate viewing of a pdf version of the observations that were reported by the observers in the parks. As previously noted, the data may be documented on paper, then relayed to an individual who will type the data into the NWS webpage. The B-91 form pdf obtained from the COOP website is the version of the data that was entered, it is not a replica of the paper copy of the data that was recorded from the station.

If it is possible to correct or fill any data, first note those changes on the working spreadsheet, the follow the instructions in the most recent version of “IMD SOP Submitting Data for Upload” (Frakes 2012b) to resubmit the corrected dataset to the database. The corrected dataset will be given a new version number.

Acquiring and Uploading Data to the NPS Climate Database

The NPS Climate Database accommodates data that has been locally acquired such as the data from the DEPO-AWS station. SIEN is responsible for acquiring and correctly formatting these data in order to submit them for upload. SIEN will compile and upload data from the DEPO-AWS or other relevant sources.

The Tuolumne Meadows COOP station data is not being made available through NCDC. Thus, we will acquire the data through NWS and submit it for upload. Each year, we will build on the previous year's submittal, thus submitting the entire period of record each year, which will be given a new version number in the database.

To submit improved and other data sources to the NPS Enterprise Climate Database, follow the instructions in the most recent version of "IMD SOP Submitting Data for Upload" (Frakes 2012b). This is normally completed and submitted by April 1st of each year.

Acquiring Snow Course Data from the California Data Exchange Center (CDEC)

SIEN will acquire tabular data for all snow courses in the major SIEN watersheds as well as summary statistics for each snow course and major watershed.

The tabular data can be downloaded at http://cdec.water.ca.gov/cgi-progs/snowQuery (Figure SOP 3.2).

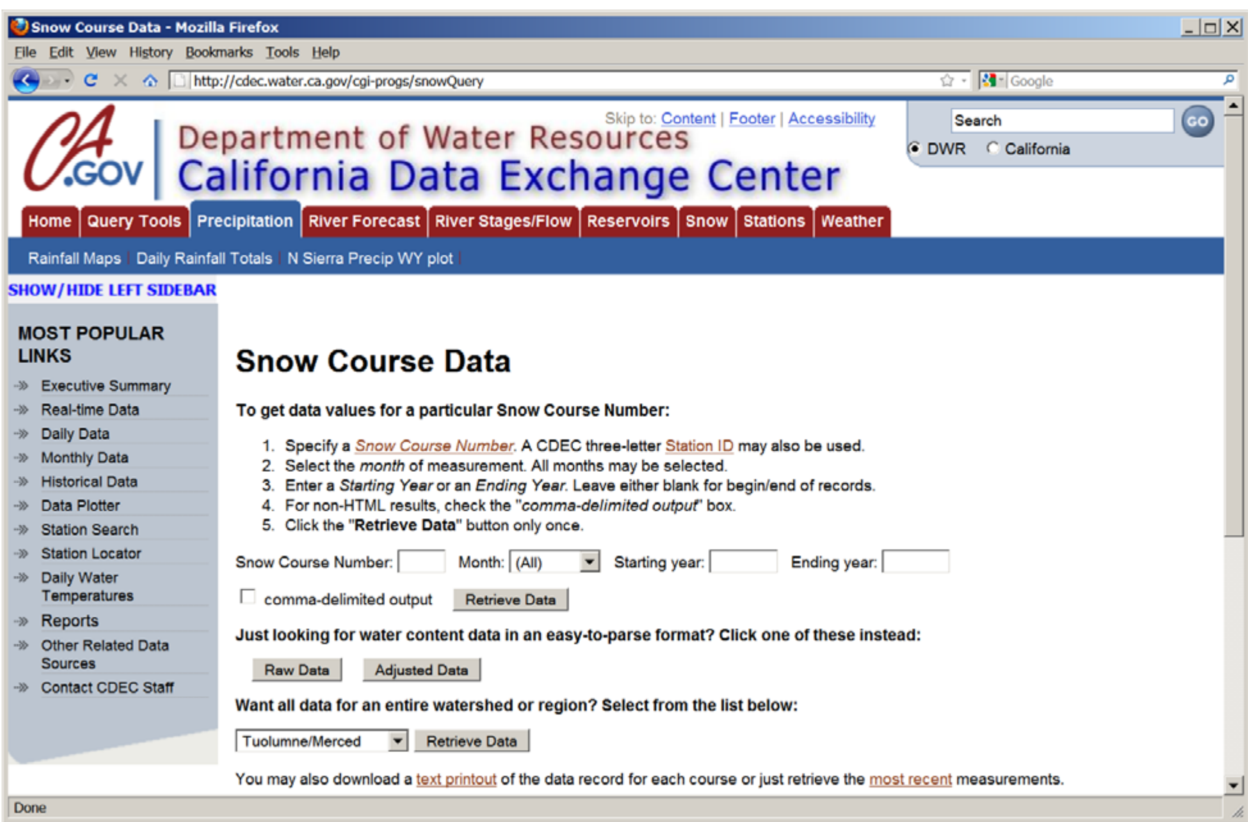
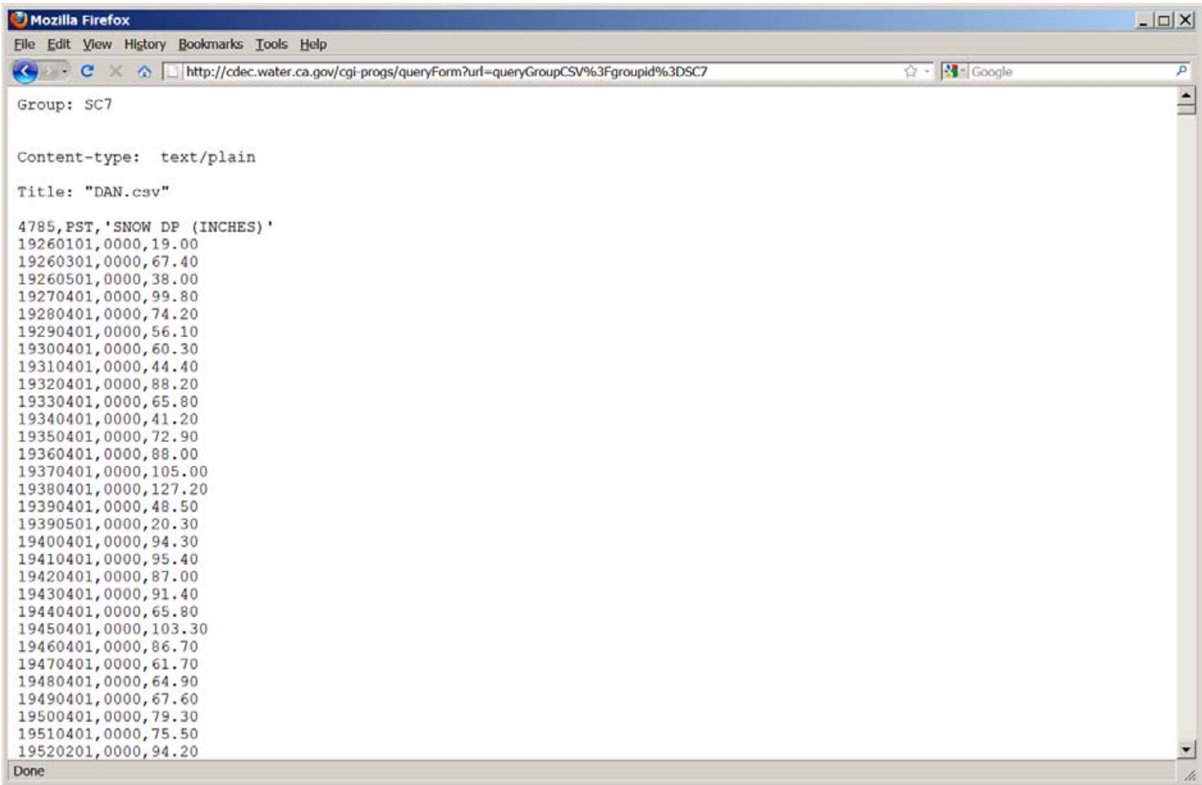


Figure SOP 3.2. The CDEC Snow Course data query page.

Query snow course data by watershed/region for the Tuolumne/Merced, San Joaquin, Kings, and Kaweah/Kern/Tule. These data include all snow courses in the watershed/region and span the entire period of operation for each snow course. These data are updated monthly from December through June. Note that not all snow courses are surveyed every month. The query returns data displayed as ASCII text (Figure SOP 3.3).




```
Group: SC7

Content-type:  text/plain

Title: "DAN.csv"

4785,PST,'SNOW DP (INCHES)'
19260101,0000,19.00
19260301,0000,67.40
19260501,0000,38.00
19270401,0000,99.80
19280401,0000,74.20
19290401,0000,56.10
19300401,0000,60.30
19310401,0000,44.40
19320401,0000,88.20
19330401,0000,65.80
19340401,0000,41.20
19350401,0000,72.90
19360401,0000,88.00
19370401,0000,105.00
19380401,0000,127.20
19390401,0000,48.50
19390501,0000,20.30
19400401,0000,94.30
19410401,0000,95.40
19420401,0000,87.00
19430401,0000,91.40
19440401,0000,65.80
19450401,0000,103.30
19460401,0000,86.70
19470401,0000,61.70
19480401,0000,64.90
19490401,0000,67.60
19500401,0000,79.30
19510401,0000,75.50
19520201,0000,94.20
```

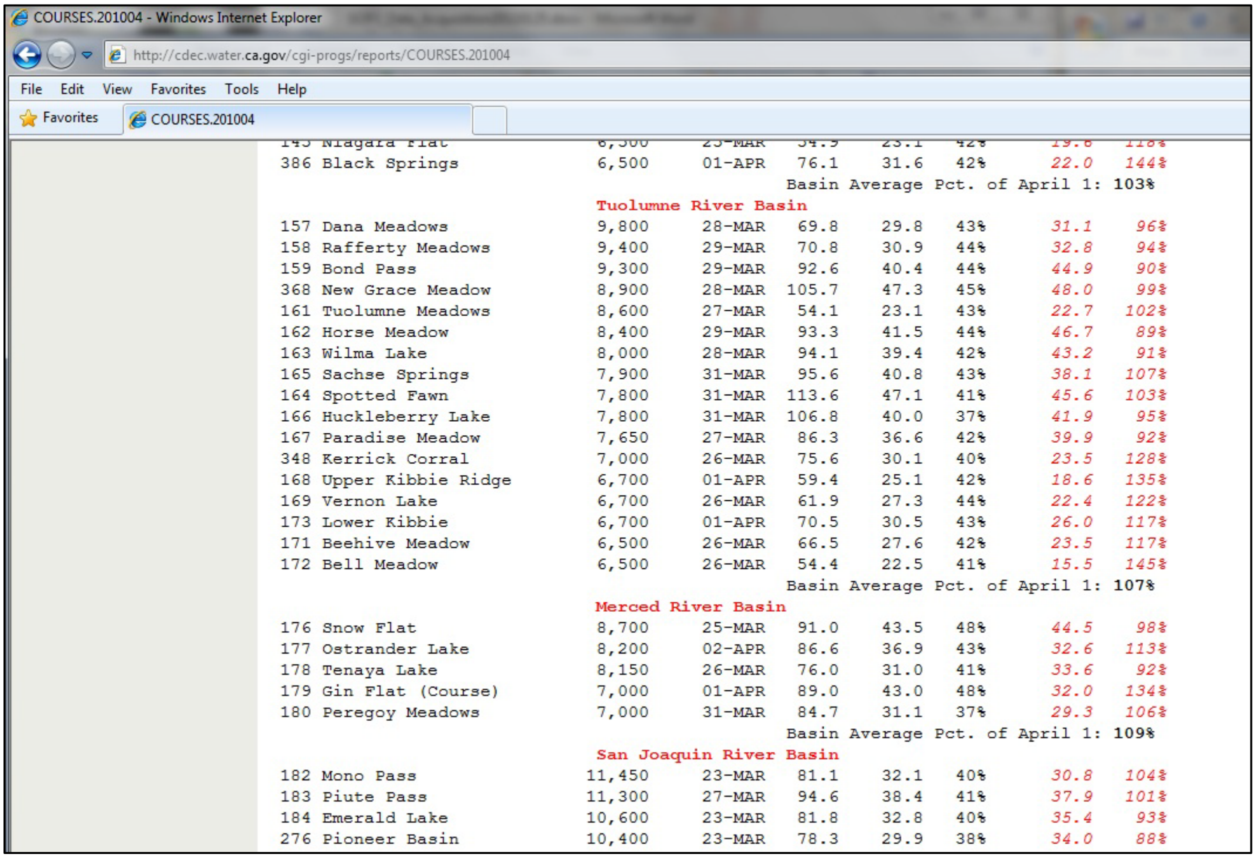
Figure SOP 3.3. Example of snow course data returned by querying a watershed or region.

On the browser's menu bar select Edit/Select All (Ctrl+A) to highlight all of the text. Copy (Ctrl+C) and paste (Ctrl+V) the text into any text editor (e.g., Notepad or Wordpad) Delete the line starting with Group: and save the file as SIEN_Snowcourse_*watershed/region_date*.txt in J:\sien\monitoring_projects\weather_climate\data\tabular\SnowCourseFromCDEC.

Import the saved text file into SIEN_SnowcourseData_CDWR.accdb, an MS Access database application developed specifically for this purpose. Code and queries in the database will be used to process, format, then export and upload the data to the IMD Climate database following guidance provided in IMD SOP: Submitting Data for Upload. The SIEN snowcourse database application will process the entire historical record for each snow course and upload it to the IMD climate database as a new version every year.

Steps for Downloading Summary Statistics for Snow Courses and Watersheds

In addition to providing historic snow course data by watershed/region, the CDWR reports the snow water content for each snow course for the first of each month, usually January through June. Typically, the April 1 snow water content (SWC) data is referred to in reports because it is often the month with the peak SWC. To retrieve the April 1 SWE for all courses and averages for the major watersheds, download the “course report” for April. The course reports are at: http://cdec.water.ca.gov/cgi-progs/rpts1/COURSES. The CDWR provides watershed and course summaries based on historic data as shown in Figure SOP 3.4. This information can easily be placed in an annual status report. The course report contains all snow courses that were surveyed in a particular watershed that year. Not all courses are surveyed each year. The top of the report will note the “reference period” to which the present year is compared to get the Percent of Average.



145 Niagara Flat	6,500	25-MAR	54.9	23.1	42%	19.6	118%
386 Black Springs	6,500	01-APR	76.1	31.6	42%	22.0	144%
		Basin Average Pct. of April 1: 103%					
	Tuolumne River Basin						
157 Dana Meadows	9,800	28-MAR	69.8	29.8	43%	31.1	96%
158 Rafferty Meadows	9,400	29-MAR	70.8	30.9	44%	32.8	94%
159 Bond Pass	9,300	29-MAR	92.6	40.4	44%	44.9	90%
368 New Grace Meadow	8,900	28-MAR	105.7	47.3	45%	48.0	99%
161 Tuolumne Meadows	8,600	27-MAR	54.1	23.1	43%	22.7	102%
162 Horse Meadow	8,400	29-MAR	93.3	41.5	44%	46.7	89%
163 Wilma Lake	8,000	28-MAR	94.1	39.4	42%	43.2	91%
165 Sachse Springs	7,900	31-MAR	95.6	40.8	43%	38.1	107%
164 Spotted Fawn	7,800	31-MAR	113.6	47.1	41%	45.6	103%
166 Huckleberry Lake	7,800	31-MAR	106.8	40.0	37%	41.9	95%
167 Paradise Meadow	7,650	27-MAR	86.3	36.6	42%	39.9	92%
348 Kerrick Corral	7,000	26-MAR	75.6	30.1	40%	23.5	128%
168 Upper Kibbie Ridge	6,700	01-APR	59.4	25.1	42%	18.6	135%
169 Vernon Lake	6,700	26-MAR	61.9	27.3	44%	22.4	122%
173 Lower Kibbie	6,700	01-APR	70.5	30.5	43%	26.0	117%
171 Beehive Meadow	6,500	26-MAR	66.5	27.6	42%	23.5	117%
172 Bell Meadow	6,500	26-MAR	54.4	22.5	41%	15.5	145%
		Basin Average Pct. of April 1: 107%					
	Merced River Basin						
176 Snow Flat	8,700	25-MAR	91.0	43.5	48%	44.5	98%
177 Ostrander Lake	8,200	02-APR	86.6	36.9	43%	32.6	113%
178 Tenaya Lake	8,150	26-MAR	76.0	31.0	41%	33.6	92%
179 Gin Flat (Course)	7,000	01-APR	89.0	43.0	48%	32.0	134%
180 Peregoy Meadows	7,000	31-MAR	84.7	31.1	37%	29.3	106%
		Basin Average Pct. of April 1: 109%					
	San Joaquin River Basin						
182 Mono Pass	11,450	23-MAR	81.1	32.1	40%	30.8	104%
183 Piute Pass	11,300	27-MAR	94.6	38.4	41%	37.9	101%
184 Emerald Lake	10,600	23-MAR	81.8	32.8	40%	35.4	93%
276 Pioneer Basin	10,400	23-MAR	78.3	29.9	38%	34.0	88%

Figure SOP 3.4. Example of a monthly snow course report that provides summary statistics for by snow course and watershed.

Steps for Downloading USGS Streamgage Station Data and Annual Hydrographs

SIEN will utilize the graph function of the USGS NWIS website which produces report quality hydrographs for streamgages. We do not intend to archive the tabular data for each of the streamgages, although the procedures are outlined here in case it is necessary to examine the data for some reason.

The following are the procedures to obtain the streamgage hydrographs for each station for the annual status reports:

- Go to the site: http://waterdata.usgs.gov/nwis/dv/?referred_module=sw.
- Check the box titled “Site number” and hit submit

- Enter the station ID number (see Table SOP 3.3) and hit enter. Click on the site number on the next page.

- Choose “graph with stats”, enter the appropriate time frame and “go”.

- Click on “create presentation quality graph”, hit enter and save the image for inclusion in the annual report.

Table SOP 3.3. USGS streamgages to be included in annual climate reports.

Station Name	Station Number
Merced River at Happy Isles Bridge	11264500
Merced River at Pohono Bridge	11266500
Tuolumne River above Hetch Hetchy Reservoir	11274790
Middle Fork San Joaquin in DEPO	11224000
Kern River near Kernville	11186000

It is not necessary to download and save the tabular data, unless you want to create a graph or figure different than the one that can be created through the steps above. If the tabular data is needed, first follow the first three steps above. Next, select discharge and enter the dates of interest, then hit Go (Figure SOP 3.5). Copy the tabular data into an excel spreadsheet and save.

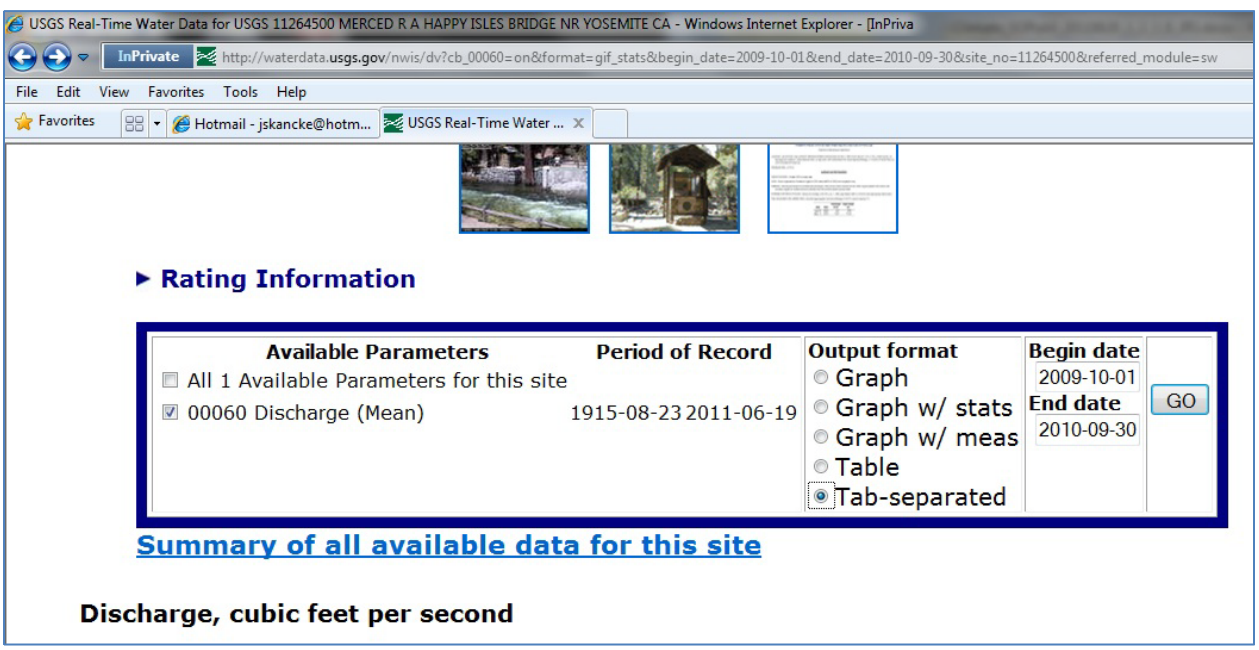
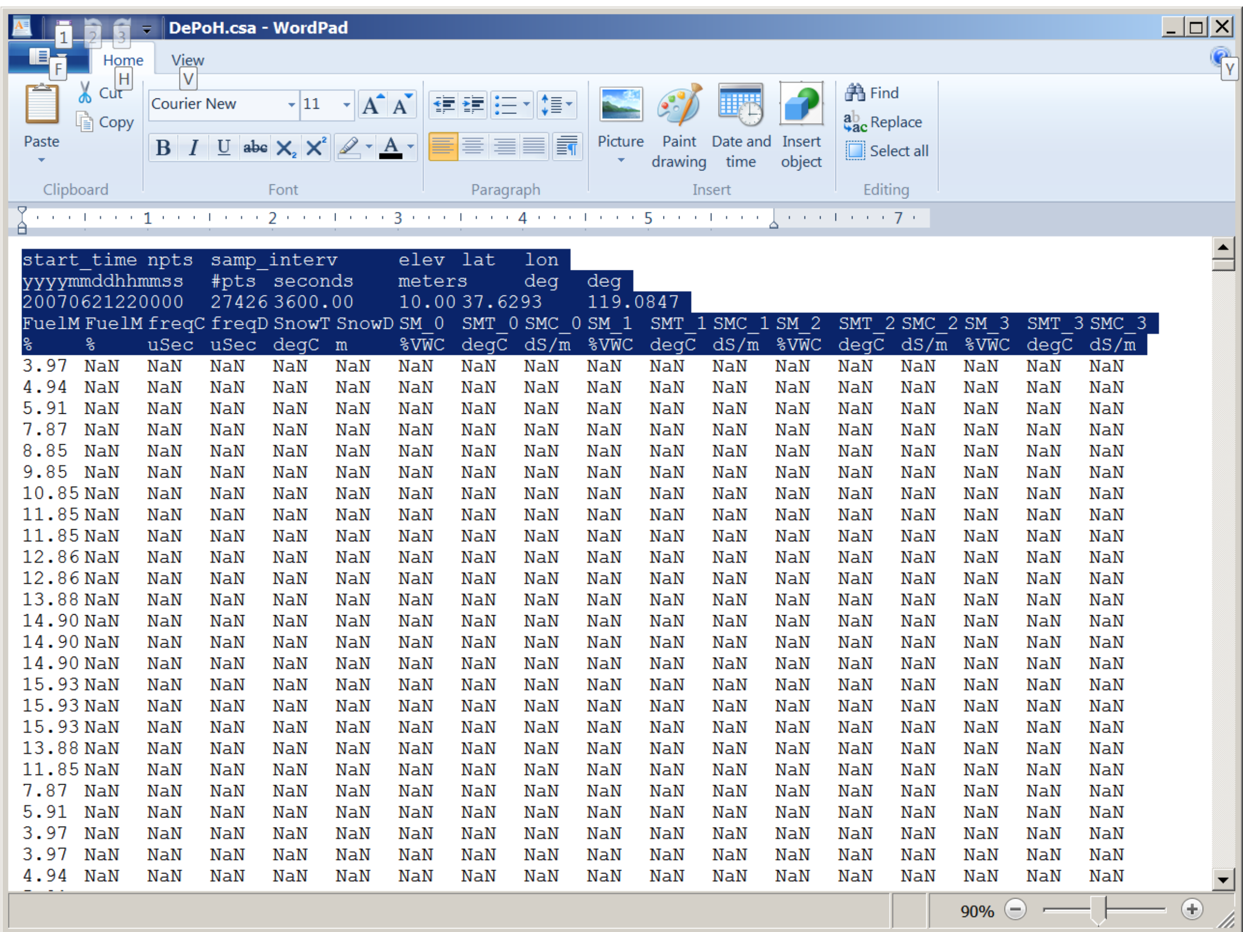


Figure SOP 3.5. USGS National Water Information System web portal.

Obtaining Data from the Meteorological Station at Devils Postpile

The meteorological station at Soda Springs meadow in DEPO is jointly operated by the Scripps Institution of Oceanography (Scripps) and the California Department of Water Resources (CDWR). See the Devils Postpile Weather Data Inventory (Balmat and Scott 2010) for further information on the sensors present and parameters collected. The Scripps staff member responsible for the DEPO station is Douglas Alden and he is in communication with the Western Regional Climate Center (WRCC) to establish a procedure for sending the data from the station for permanent archiving each year. The WRCC may format the data in a way that will be easier to handle than the format that is provided by Scripps. Until that agreement is established, it is necessary to request the data directly from Douglas. The data is collected in 1-minute increments. When the data is received from Douglas, follow these steps:

- Send request to Doug Alden (dalden@ucsd.edu) on June 01
- After receiving the data, open the file in a text editor (e.g., Notepad or Wordpad) (Figure SOP 3.6).
- Select and delete the header information in the file so only data remains.




```
start_time npts samp_interv elev lat lon
yyyymmddhhmmss #pts seconds meters deg deg
20070621220000 27426 3600.00 10.00 37.6293 119.0847
```

FuelM	FuelM	freqC	freqD	SnowT	SnowD	SM_0	SMT_0	SMC_0	SM_1	SMT_1	SMC_1	SM_2	SMT_2	SMC_2	SM_3	SMT_3	SMC_3
%	%	uSec	uSec	degC	m	%VWC	degC	dS/m	%VWC	degC	dS/m	%VWC	degC	dS/m	%VWC	degC	dS/m
3.97	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4.94	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5.91	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7.87	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12.86	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12.86	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13.88	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14.90	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14.90	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14.90	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
15.93	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
15.93	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
15.93	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13.88	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11.85	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7.87	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5.91	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3.97	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3.97	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4.94	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Figure SOP 3.6. Tabular data from the DEPO-AWS station.

- Save the file as a text file using the appropriate file naming convention: e.g., DEPO_RawWxData_ SCRIPPS_*date_username*.txt
- Open DEPO_WxData_Scripps_Vxx.accdb (development to be completed by the SIEN Data Manager, Spring 2012)
- Run the queries to flag missing values and perform basic QA/QC to check for logical inconsistencies (minimum temperatures greater than maximum temperatures, etc.)
- Use queries in the database application to convert the data collected at 1-minute intervals to hourly and daily time periods.
- Use queries in the database application to calculate the average of the daily min and max to get the monthly min and max. Note where > 3 days are missing from any month and exclude.
- Use queries in the database application to calculate the average of the monthly mean min and max to get the annual mean min and max (convert values to F).
- Precipitation data is reported as an accumulated value and resets to zero when Scripps downloads the data. There is no set schedule for data download. Thus, it is necessary to scroll through the data and note when the accumulated precipitation has reset to zero. If it has not reset to zero in the middle of the month, subtract the first value for each month from the last value to get the monthly precipitation.
- Alternatively, data may be processed using the Aquarius Workstation whiteboard. Aquarius software is specifically designed to analyze continuously collected time-series data. The NPW Water Resources Division (WRD) recently deployed a 5-user Aquarius license that can be accessed by remote desktop. The software provides procedures to conduct QA/QC, perform data reduction , and export data summaries. The sequence of procedures can be saved for reuse. This alternative may prove to be a more efficient way of processing the DEPO data and will be compared with current procedures. We will provide additional details on using Aquarius software in a revised version of this SOP if Aquarius becomes the preferred method for processing the data from the DEPO station.

The Scripps staff will perform QA/QC and summary statistics on the data from the station, but the schedule for such tasks may be significantly delayed compared to SIEN's schedule for producing annual reports. When the data is acquired for trend reports, it will be necessary to get the quality-checked data as well as any documentation.

The CDWR collects duplicate precipitation and temperature data that is available through the CDEC website. The sensors undergo less rigorous calibration and maintenance than the Scripps sensors so the data should be used only for filling or comparison. The website link to the CDWR data is: http://cdec.water.ca.gov/cgi-progs/staMeta?station_id=DPO.

Acquiring Data from Air Quality Stations

Although data from these stations will not be utilized annually for reporting, the air quality monitoring programs have stations in YOSE and SEKI that can provide additional data for periodic trend reports. The stations do not have long records, but in some cases, are located near COOP stations so the data could be used to fill in missing records if needed.

Temperature and total daily precipitation data is available for National Atmospheric Deposition Program (NADP) sites Hodgdon Meadow (YOSE) and Giant Forest (SEKI) through the website: http://nadp.sws.uiuc.edu/sites/sitemap.asp?state=CA.

Hourly data are available for CastNet sites Turtleback Dome (YOSE) and Ash Mountain (SEKI) at : http://www.epa.gov/castnet/data.html. One must download the table for all sites for the entire year then sort, cut and paste SIEN specific data into new table.

Hourly temperature and precipitation data from the Gaseous Pollutant monitoring sites Lower Kaweah and Ash Mountain - SEKI, Turtleback Dome, Schoolyard, Mobile, and Village -YOSE are available at: http://12.45.109.6/data.aspx.

Tips for acquiring NWS-COOP data without the assistance of IMD

As of 2012, the NPS IMD accepts requests to acquire, compile, and manage station data from NWS-COOP stations. These tips demonstrate how to acquire these data on your own in case you need original provider data or if IMD is unable to provide compiled data for requested stations.

During initial data acquisition, multiple years or the complete station history will need to be acquired. In subsequent years, data for individual water year or calendar year may be acquired and combined with past data already acquired stored in the local database. However, since data from previous years or decades may be updated or replaced by the provider, project staff must annually review the source website for indications of changes and updates. Updates and changes can be found in associated metadata files (e.g., station history for COOP network). If an important update occurs since the data was originally acquired, download and replace the entire station history data set.

Download methods may differ depending on the requirements and expertise of the project staff and as the providers' online distribution systems change over time. Therefore, the instructions below are guidelines to aid users who will follow current step-by-step directions at each source web site to identify, select, and obtain data. All source websites provide several methods to retrieve a single or multiple stations in a download file. The preferred method will vary depending on the number of stations desired and the most efficient means to retrieve the data. Project staff will document acquisition steps for each transaction to store with metadata in the climate protocol folder on the shared network drive.

Save the file(s) using the appropriate local naming convention or as follows:

[*Source name*]_[*StationID or State*]_[year(yyyy)].txt.

Files in text format are preferred to support data transformations, uploading to other systems, and archiving.

Project staff should maintain an electronic tracking log for all downloads with the transaction date and the name of the person who conducted the download. The tracking log may be in a database or a simple Microsoft Excel file that is stored with or linked to the downloaded text files. Tracking log review and maintenance activities may be incorporated with annual data quality control procedures.

Steps for Downloading COOP Data

(1) Navigate to the NCDC weather observation station locator website (Figure SOP 3.7): http://www.ncdc.noaa.gov/oa/climate/stationlocator.html

Figure SOP 3.7. NCDC weather observation station locator website.

(2) Click on Location Search Tool

(3) Enter the station name in the box and select a data product.

- Daily GHCND will give the daily COOP observations.
- Precipitation 15-minute and Precipitation Hourly are not options for COOP stations.

(4) Click on the (+) sign next to the station to add it to the box on the right side of the screen (Figure SOP 3.8). Click **CONTINUE** in the box on right hand side of the screen.

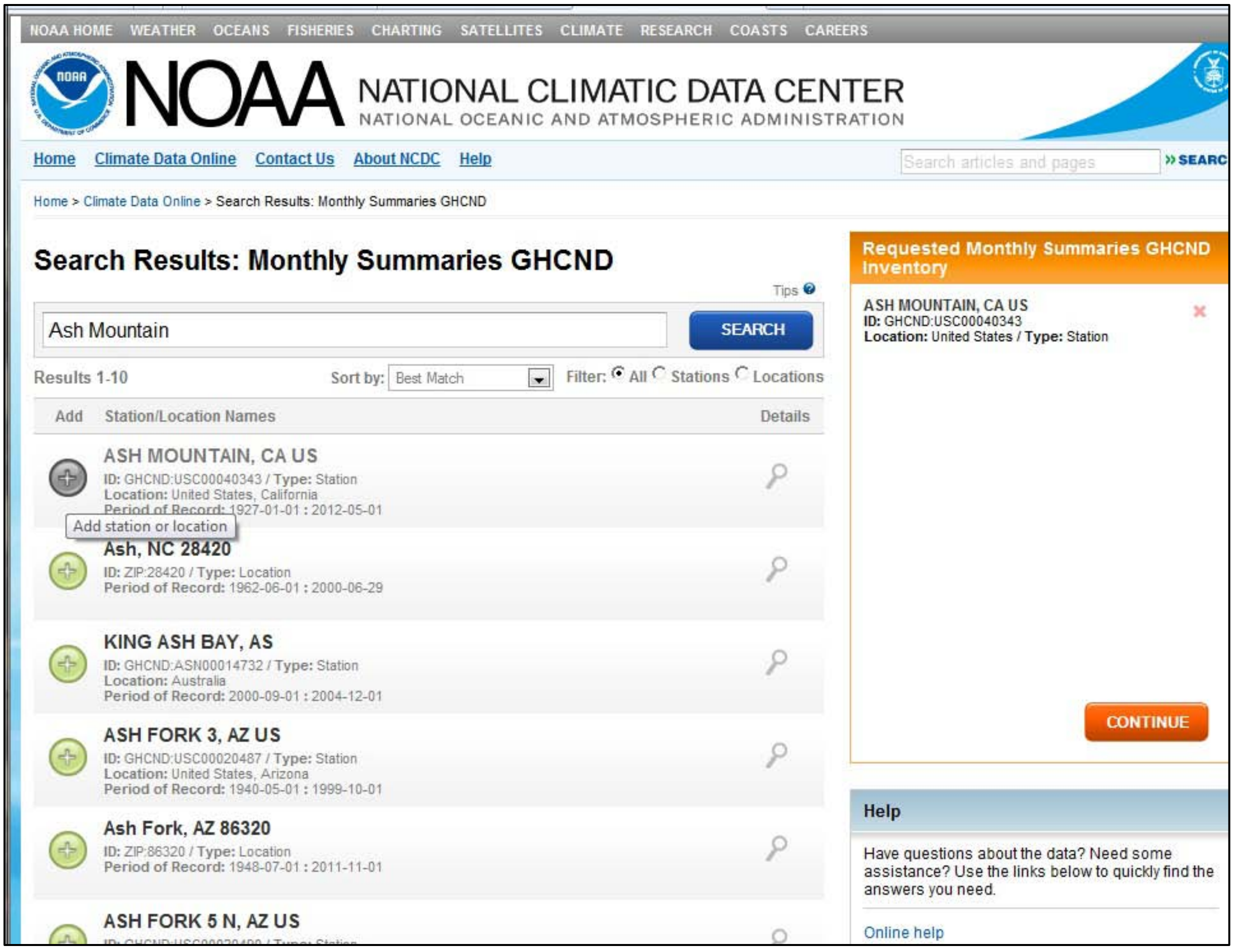


Figure SOP 3.8. NCDC data retrieval website.

(5) Fill out the information on in the “complete order information” box.

(6) A request submittal confirmation email will arrive followed by an email with a link to download the actual data.

Steps to Acquire B-91 Forms

B-91 Weather Station Records are the original paper forms used by observers in the NWS-COOP network and can be downloaded as pdf files. They may be used to infill missing data or to better QA/QC stations of interest. B-91 forms contain all measurements made from one station during one month. If data are corrected based on the forms, a new version of the dataset must be uploaded to the Climate Database as following “IMD SOP: Submitting Data for Upload”.

It is also possible to download the pdfs of the original data collection forms from each COOP station. These might be helpful to double check if missing data was truly not collected or was entered incorrectly.

(1) Navigate to:
http://www7.ncdc.noaa.gov/IPS/coop/coop.html?foreign=false&_page=0&jsessionid=BA2FE892A96CA766C581D36CE57F1B58&state=CA&_target1=Next+%3E.

(2) Choose station of interest (Figure SOP 3.9), click on NEXT

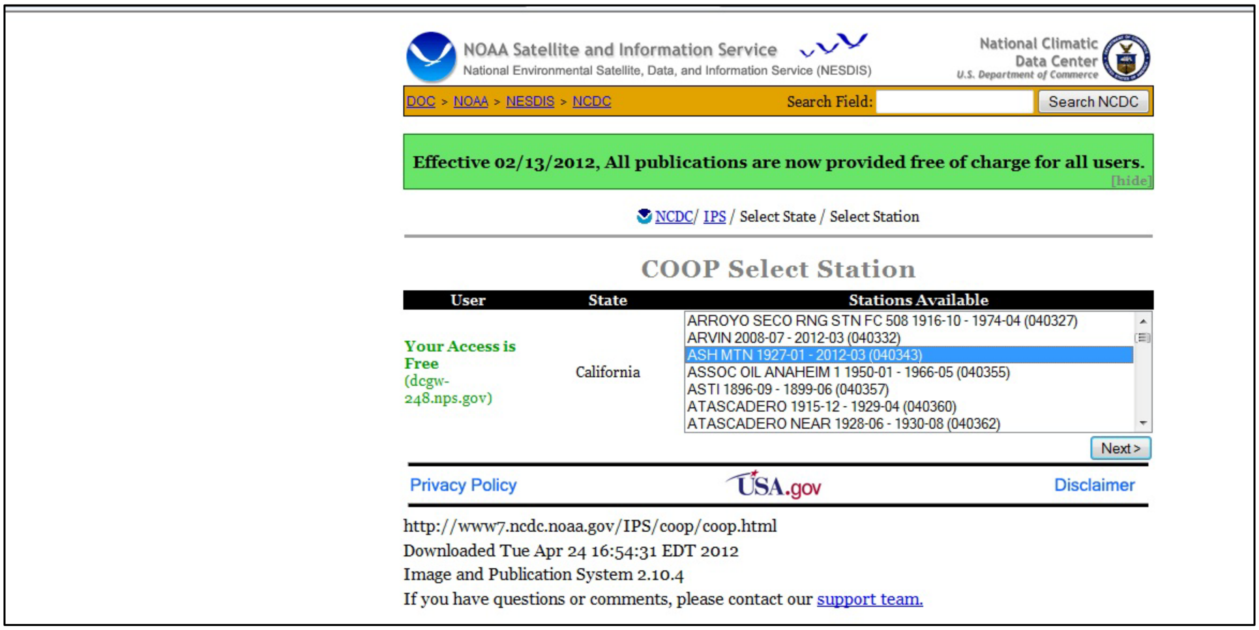


Figure SOP 3.9. NOAA COOP program B-91 datasheet website.

(4) Choose Year and Month of interest from list, click on NEXT.

(5) Click on file, it will open as a pdf. Check that it is the correct file and then save file as .pdf and name it "STATION_YEAR_MONTH.pdf". Store the pdf in the correct calendar year folder on the shared drive at: J:\sien\monitoring_projects\climate\data\tabular\2012\B_91Forms.

Acquiring Geospatial Data

Instructions for acquiring and processing geospatial data from PRISM are found in "SIEN SOP 4: Processing Geospatial Data" and "IMD SOP: Climate Grid Analysis Toolset".

Note: The Inventory and Monitoring Division is developing the NPScape web application that may be an adequate replacement for the processes outlined in the SOPs listed above.

Literature Cited

Balmat, J. and D. K. Scott. 2010. Weather data inventory, Devils Postpile National Monument. Natural Resource Report NPS/SIEN/NRDS—2010/113. National Park Service, Fort Collins, Colorado.

Frakes, B. and S. Kingston. 2011a. Connecting to the I&M Enterprise climate database for the purpose of data retrieval, summary and/or analysis. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2167699 (accessed 30 March 2012).

Frakes, B. 2012a. Requesting essential parameter values from COOP, SNOTEL, streamgage, and snow course data for acquisition, compilation, basic quality control, and distribution by the NPS Natural Resource Program Center. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166927 (accessed 30 March 2012).

Frakes, B. 2012b. Submitting data for upload to the I&M Enterprise climate database. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166928 (accessed 30 March 2012).

Rocky Mountain Climate Working Group. 2011. Standard operating procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sierra Nevada Network Climate Reporting Protocol

SOP 4: Processing Geospatial Data for Status Reports

Version 1.0

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Figures

Introduction

This SOP is meant to act as a supplement to the "IMD SOP: Climate Grid Analysis Toolset" (Sherrill and Frakes 2011). The Climate Grid Analysis Toolset (CGAT) is a suite of GIS Python scripts developed to facilitate efficient analysis of PRISM and SNODAS geospatial climatic datasets. SIEN will use the tools for PRISM but not SNODAS. The CGAT tools are packaged as two sets of python scripts requiring only basic understanding of python/computer programming and a moderate level of proficiency in GIS.

Before proceeding with this SOP, go to the Integrated Resource Management Application (IRMA) for the most recent version of Sherrill and Frakes (2011) as well as updated scripts and the executable file to install PythonWin. PythonWin is a program that works with Python to process the PRISM files. The SOP, scripts and executable file can be found at https://irma.nps.gov/App/Reference/Profile/2171809. Update the files to the network drive. Read through Sherrill and Frakes (2011) and check that the most recently updated scripts are compatible with the version of ArcGIS currently installed on your computer.

PRISM

Given the relatively sparse network of climate stations and complex topography in the Sierra Nevada Mountains, grid-based estimates of precipitation and temperature can be used to provide an overview of climatic conditions in the area. These estimates are generated via a statistical modeling technique that interpolates precipitation values between actual climate observing stations while also accounting for the effects of aspect and elevation. Known as the Parameter-elevation Regression on Independent Slopes Model (http://www.prism.oregonstate.edu/), this approach has a long history of use in the western United States, and it has been shown to provide highly robust products in a wide variety of studies (Daly et al. 2008).

Using the Climate Grid Analysis Tools

Start-Up

1) Ensure that all files for the year to be analyzed are stored on your computer along with files from the previous 30 years – or from a 30 year reference period such as 1971-2000 for PRISM. Follow the instructions below to download the files. Files will be stored from year to year on the network drive, so each year there should only be a need to download the files for the current year. Create a set of temporary folders locally on your computer for the files. You will need a folder titled "PRISM", along with three folders within titled "temp", "tmin", and "tmax". Pull the historic files from the SIEN server to the local folders you created. To download the most recent files:

    a. Install the FileZilla program. FileZilla is a program that assists with the transfer of files from ftp sites.

    b. Important: First visit http://prism.oregonstate.edu/products/matrix.phtml?view=data. http://prism.oregonstate.edu/products/matrix.phtml?vartype=tmax&view=data. And http://prism.oregonstate.edu/products/matrix.phtml?vartype=tmin&view=data to check that the data are in their "final" format rather than "provisional. The ftp site used to obtain the files does not indicate the status of the files, it only indicates that they are available.

c. The ftp site for PRISM data is: ftp://prism.oregonstate.edu/pub/prism/us/grids/. If the final versions of the data are available, proceed with the next steps.

d. Open FileZilla. As shown in Figure SOP 4.1, insert the ftp site address listed above in the box "host". Type 21 into the box "port" and click "quick connect". The username and password will be filled in for you after you click "quick connect". Navigate in the left window to the local drive on your computer where the files will be stored. C:/PRISM is a good temporary location, with a separate folder for ppt, tmin, and tmax. Drag the monthly PRISM files from the right window to the left window. File transfer should commence. After the CGAT python scripts have been run from the local drive, copy the PRISM files over to the network drive for storage. \\Inpsekihqgis1\sekigis\sien\monitoring_projects\weather_climate\data\spatial\original_data.

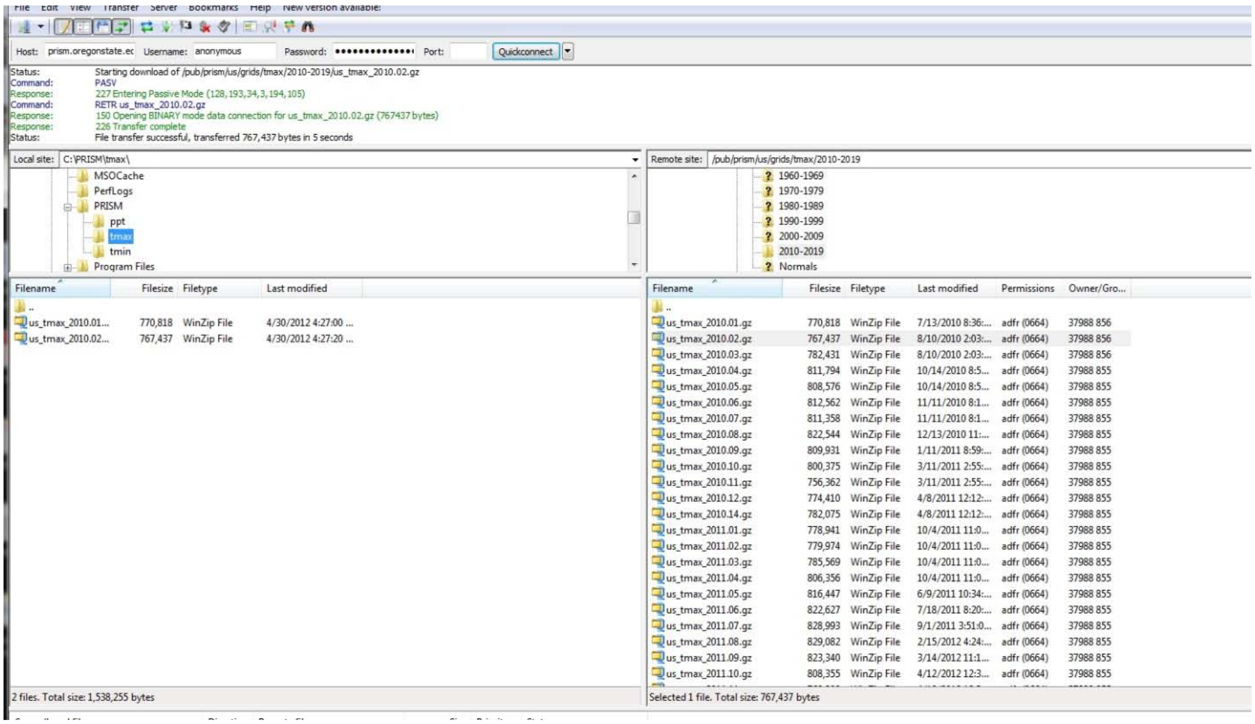


Figure SOP 4.1. FileZilla application used for geospatial file transfer.

Running the CGAT

1. Ensure that PythonWin is installed on the computer you are using. The executable file to install PythonWin is located in the product profile zip file in IRMA. You can check which version of Python you have by viewing the Python folder on the C drive of your computer. The folder will be named Python##. The numbers indicate the version of Python so a folder titled Python26 means you're running version 2.6. You will need to run the executable file of PythonWin that corresponds to your version of Python. For example, the executable PythonWin file for Python26 is pywin32-217.win32-py2.6.exe. If the correct version is not located in the IRMA product profile, contact IMD for help locating the correct version of PythonWin.

2. The Python scripts will not work when directed to files on the network drive, it is necessary to direct Python to the files to your computer while you are performing the analyses. Most files can and should be deleted from these local folders and moved to the network drive after all the steps listed below have been completed because the files are large and take up memory on the computer hard drive.

3. Place the CGAT Python scripts locally on your computer. In the past, python scripts have not worked properly from the network drive. You should have already moved the most recent scripts to the network drive and will just need to copy them to your hard drive. You must have both PRISM_MS_Arc10.py and FunctionsPRISM_Arc10.py. You will not need to open the functions file (unless you are troubleshooting).

Creating PRISM image files

The CGAT for PRISM will allow for the creation of annual or monthly average files as well as files which show the annual or monthly percent of precipitation or temperature for the current year compared to a given time frame (percentiles). It is common to compare the current year to either the previous 30 years or the 30 year "normal" which, for the 2009 report, would be 1971-2000. For the creation of annual reports after 2010, the 30-year normal will be 1981-2010. It is best to use the same 30-year normal timeframe that the station data will be compared to. To begin, right click on the PRISM_MS_Arc10.py (which you should have placed somewhere on your computer) and choose Edit with PythonWin.

(1) Enter the correct information under "Start of Parameters Requiring Set Up".

   a. Switch the workspace to the correct folder on your computer.

   b. Enter the BeginYear and EndYear. These are the years for the "30-yr normal" (i.e. 1981-2010). Note: The current year must be within the begin date and end date. Therefore, if you want to compare 2012 to 1981-2010, you must have a begin date of 1981 and an end date of 2012 even though you would like to ignore the year 2011. Next, you must ensure that only the files for the years 1981-2010 as well as 2012 are in the PRISM files folder. Pull the 2011 files out of the working PRISM files folder and store them on the network drive.

   c. Enter a park extent. An appropriate "parkExtent" for the SIEN network is -120.6W, 36.0S, -117.0E, 38.5N. The park extent does not affect the python calculations so the extent can be made larger or smaller. A much larger park extent will take more time to process.

   d. Enter the appropriate set-up parameters for Variable, AverageType, and TimeStep. Refer to the CGAT SOP for assistance in setting the correct parameters.

(2) Enter "yes" for Percentiles and the PercentileDate. At this time set "zonal" to no. Don't worry about the "polygonfield" and "zonefield".

   a. When creating percentile calculations for each month of a year, it is necessary to run each month separately. For example, for the 2009 annual report, one ArcMap document will be produced showing the precipitation (ppt), maximum temperature (tmax), and minimum temperature (tmin) percentiles for each month of 2009. First, for January, enter 200901 as Percentile Date. A "month_percentile_200901.img" file will be created in same folder where the PRISM files are stored. Repeat this step with 200902 through 200912. Each run takes approximately 5 minutes. A total of 36 runs will be performed, 12 each for ppt, tmax, and tmin percentiles.

(3) Click on Run. Do this for a single parameter (ie. ppt), then follow the directions below before going through the steps for the next parameter.

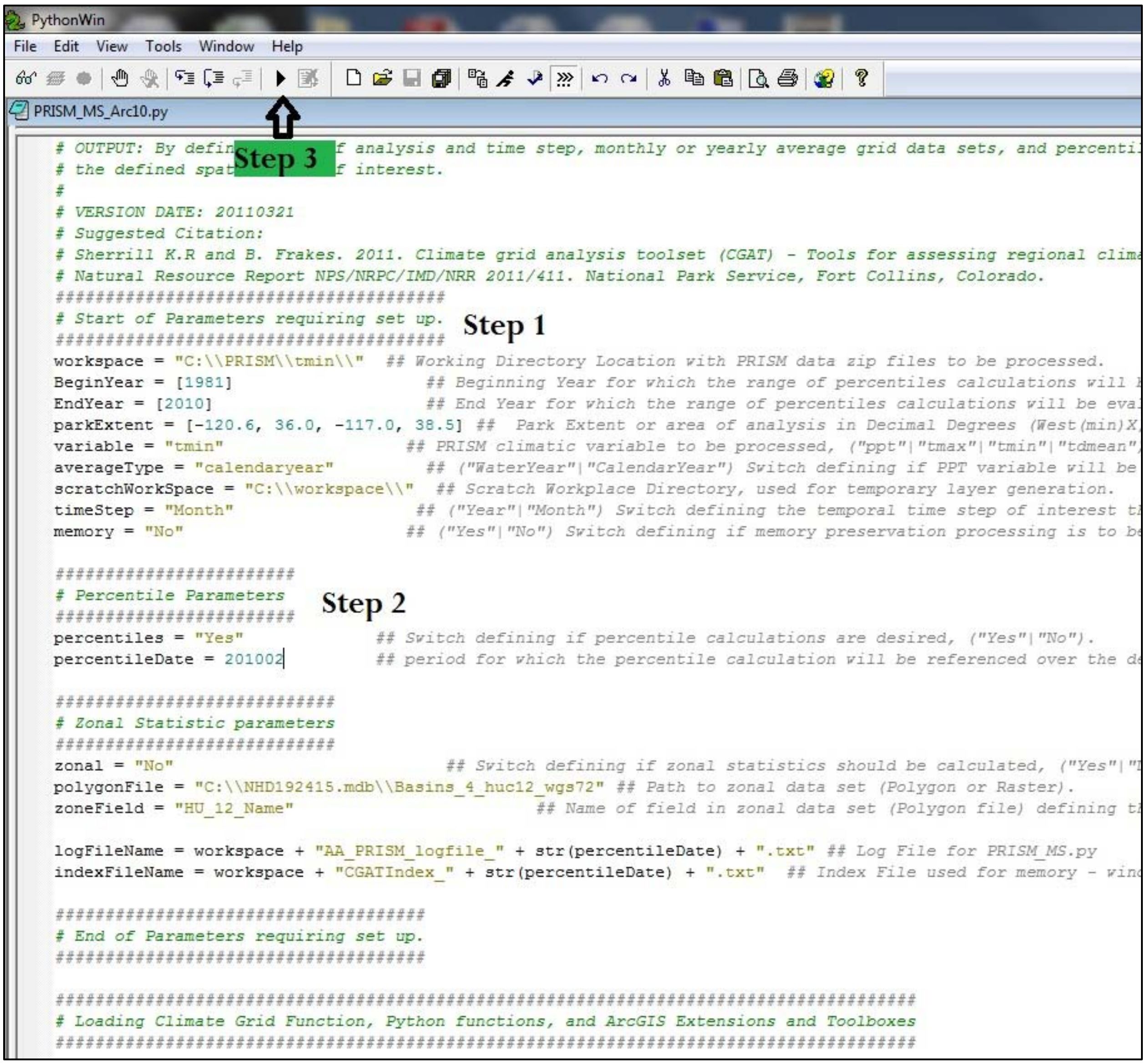


Figure SOP 4.2. Set up parameters requiring modification from the original script.

When the scripts are run, a variety of files will be placed in the folder with the PRISM files. Most of these will be deleted, only those labeled month_percentile_YYYYMM.img will be kept. After you've finished running each month for the parameter of interest follow these steps:

1) create a folder on the SIEN server for the current year at \\Inpsekihqgis1\sekigis\sien\monitoring_projects\weather_climate\data\spatial\derived_data\PRISM)

2) copy the files month_percentile_YYYYMM.img from the local folder to the network drive

3) rename each IMG file appropriately so the file may be linked to the ArcMap project (See Figure SOP 4.3). It is important to rename the files right away with the correct parameter because the original output files from Python do not include the parameter. So, the ppt files with have the same name as the tmin files until you rename them.

Note: It can be helpful to add a readme file which states the dates you chose for the climate "normal" when you ran the phthon scripts or create a folder named "1981_2010Normals" or "1971_2000Normals". For example:
J:\sien\monitoring_projects\climate\data\spatial\derived_data\PRISM\2010\1981_2010Normals

After you've moved and renamed the 12 image files for each parameter (36 total files) to the SIEN drive, you can delete all the additional files that have been created by Python from your computer. Be sure to transfer over the PRISM files that you downloaded from the PRISM ftp site first, and then you can remove those files as well.

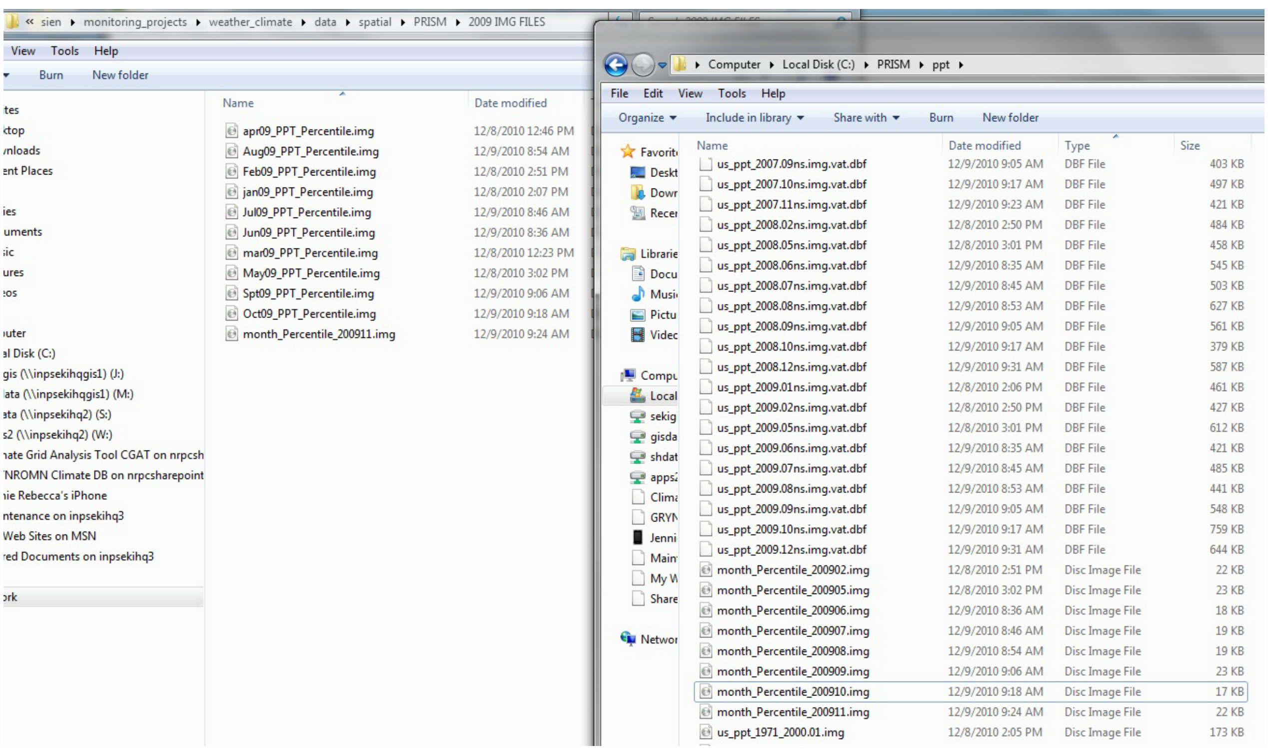


Figure SOP 4.3. IMG file is produced and transferred to the network drive then renamed.

Steps for Creating Figures with the Processed PRISM Data using ArcMap

Several templates have been created which can be saved with the current reporting year and modified

(J:\sien\monitoring_projects\weather_climate\data\spatial\GIS_Projects//PRISM/PRISM templates). Open the template, then save to J:\sien\monitoring_projects\weather_climate\data\spatial\Derived_data in the correct yearly folder with the file name modified to include the creation date (Figure SOP 4.4), or modify the project from the previous year, rename and file in the current year folder.

Change the source for each layer/month to the correct .img file, then export the project as a .jpg to the current year's annual report folder J:\sien\monitoring_projects\weather_climate\products\reports\.

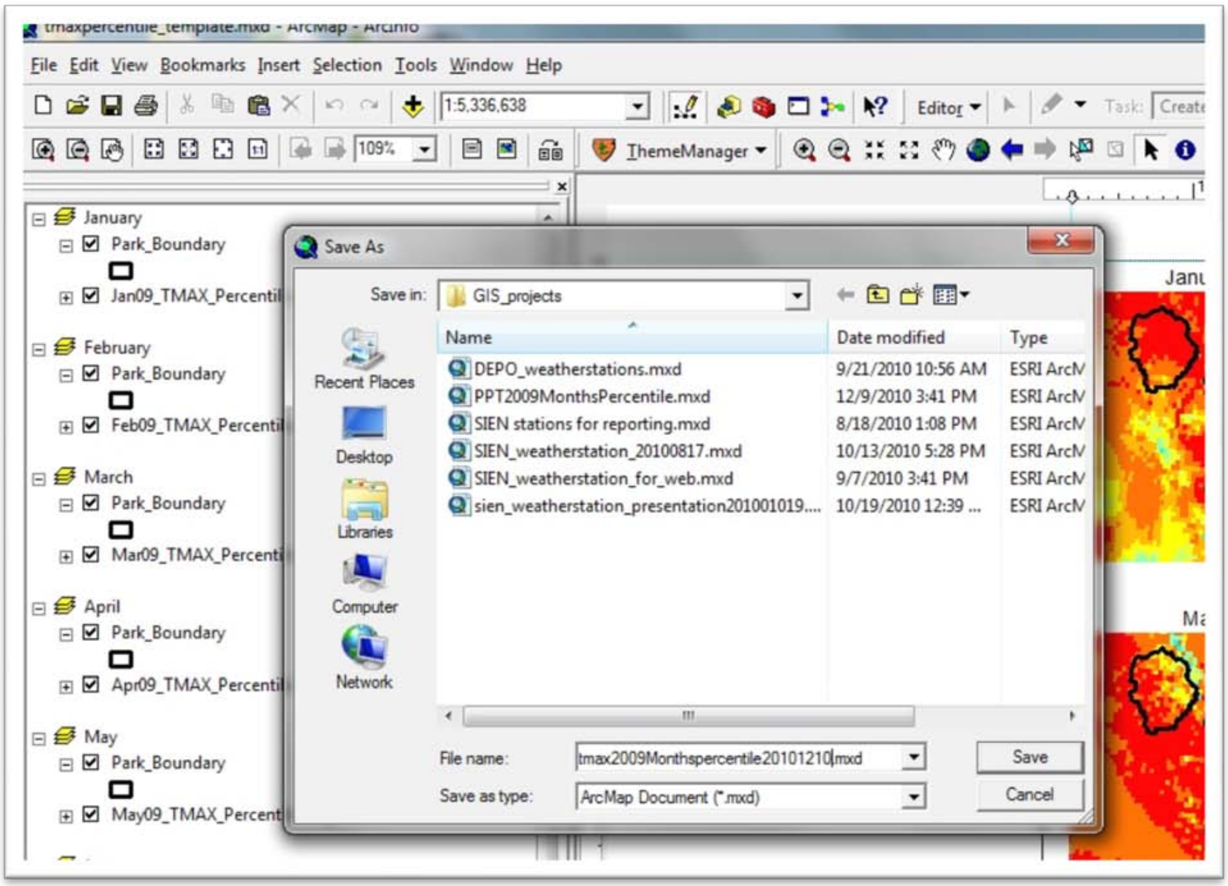


Figure SOP 4.4. ArcMap template and file renaming procedure.

Literature Cited

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. A. Pasteris. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. International Journal of Climatology 28:2031-2064.

Sherrill, K. R. and B. Frakes. 2011. Climate grid analysis toolset – Tools for assessing regional climatological trends: Standard operating procedure (Version 1.2). Natural Resource Report. NPS/NRSS/IMD/NRR—2011/411. National Park Service, Fort Collins, Colorado.

Sierra Nevada Network Climate Reporting Protocol

SOP 5: Annual Status Report

Version 1.0

This document has been adopted in part from 'Rocky Mountain Climate Protocol SOP Status Reports' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Figures

Page

Tables

Page

Introduction

This SOP provides a framework for the annual climate status report which summarizes and interprets 1) temperature, precipitation and drought metrics within and around a park unit for each calendar year, and 2) snowpack, snow water equivalent, and streamflow for each water year (October – September). The primary audiences for annual status reports are park staff and managers, researchers, and collaborators.

Due to a lag-time in the availability of quality-controlled and checked data, we expect the reports to be initiated approximately six months after the end of the water year (approximately the following April). Preliminary data may be available prior to April, but may not have gone through QA/QC processes and should not be used for published reports.

Format

Report Template

The annual reports should be published in the Natural Resource Technical Report Series (NRTR)**.** The directions and details for submission can be found at http://www.nature.nps.gov/publications/NRPM/**.**

The report should be titled: "Climate Status Report for the Sierra Nevada Network parks, YYYY". A brief description of the report content is provided below.

Executive Summary: A brief (~300 words) abstract of the annual report highlighting where and if the year departs from normal/recent conditions. It should be concise and written for a broad audience including superintendents and the public.

Introduction: State the purpose and goals of the report, the definition of the water year, and general information about the network.

Data and Methods: This should describe and provide web links to the original data sources, including the version identifier for compiled data used from the NPS Climate Database or other providers. Describe the quality control measures taken as well as the calculation of any indices (e.g., standard precipitation index) or statistical tests used. Details may also be placed in an appendix.

Results and Discussion: Presenting each parameter of interest (e.g., temperature, precipitation, snowpack, drought and streamflow) in separate sections is straightforward and easy to interpret, with the narrative integrating across variables placed in a separate section or the summary. PRISM data should be presented for the entire network. For each parameter, intra and inter-annual variation are discussed.

Conclusion: The conclusion briefly and comprehensively synthesizes the results and discussion, and provides integration across climate variables. It should contain more detailed information than the Executive Summary.

Units for Data Analysis and Graphics

For annual status reports, use English (°F and inches) units.

Baselines for Comparison

When possible, use mean values from the established 30-year normal period (i.e., 1971-2000) for comparison to the current year's climate. Calculate the 30-year normal by using the table-valued functions in the Climate Database which are available in "IMD SOP: Connecting to the I&M Enterprise Climate Database" (Frakes and Kingston 2011). Compare the calculated 30-year normal from the database to the calculated value on the Western Regional Climate Center (WRCC) website (Table SOP 5.1). In cases where data for the 30-year normal period are not available, we use alternative comparisons such as the mean from the period of record or a recent 10-year period or we present the annual values without comparison to historic averages. Data may be reported in the form of percentiles, which are often appropriate because climate data are not always normally distributed, are skewed by outliers, and can be on different scales. Percentiles are used for all intra-annual departures because averages and normal ranges of a parameter change for each month. Likewise, percentiles are an effective way to normalize spatially distributed variations.

Table SOP 5.1. List of stations used for annual status reports with links to "30-year normals".

Station Name	WRCC website link to data
Ash Mountain	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca0343
Lodgepole	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca5026
Grant Grove	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca3551
South Entrance Yosemite	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca8380
Yosemite Valley	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca9855
Hetch Hetchy	http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca3939
Tuolumne	Summary data not yet available on the WRCC website
DEPO Soda Springs	Summary data not yet available on the WRCC website

Producing the Annual Status Report

Production of the report involves these basic steps:

1. Download the current NPS Natural Resource Technical Report Series template and save as a document for the park(s) and reporting period;
2. Follow the Data Acquisition SOP and work with the network data manager to obtain data that has undergone required quality control;
3. Analyze the current water and calendar year data and produce graphics using SigmaPlot, ArcMap, or website specific functions (i.e., produce hydrographs through the USGS NWIS website);
4. Insert graphics, tables, summary statistics and data into the template;
5. Consult with a climatologist as necessary and add text to interpret the current year climate data and graphics (methods and introduction will likely remain the same each year);
6. Follow PWR peer review guidelines;
7. Revise report as needed and submit for publication in the Natural Resource Report Series.

Content of Annual Status Report

Data and Methods: Overview of Data

1. Map and provide metadata (e.g., period of record) for all stations used in the report. Report data completeness for each station (e.g., Garman 2009).

2. Describe the data flow from original providers to NPS systems, along with data processing performed to prepare the data for summary reporting. Clearly explain how park staff can directly obtain or request the data used in the report.

Results and Discussion

Listed here are suggested figures and tables to be included in annual status reports. During any given year, the Physical Scientist may choose to include or exclude a selection of these figures and tables. Each annual report must include a summary of each of the five selected climate parameters: temperature, precipitation, snow pack, streamflow, and drought.

Temperature

1. Overview for the network
   a. Monthly Tmin and Tmax PRISM maps for each zone or region as described in “IMD SOP Climate Grid Analysis Tools (CGAT)” (Sherrill and Frakes 2011) and “SIEN SOP 4: Processing Geospatial Data”. Consider using one map per month/variable of the larger region along with and park boundaries. See example in Ashton et al. (2010).

2. Station specific data
   a. Tables of average minimum and maximum daily temperatures, and departures from 30–year normals for each month for each of the selected NWS-COOP stations if available.
   b. For NWS-COOP stations, plot monthly mean maximum and minimum along with 30-year normals, or as deviations from normals (Figure SOP 5.1).

3. Annual Timing Variables
   a. Narrative discussing temperature during the past year, including discussion of annual timing and integrative variables: mean temperature and accumulated growing degree days. Optional: show the number of days with Tmin >90th percentile over the period of record (Figure SOP 5.2).
   b. Table with first and last freeze and frost dates, number of days below 0ºF and 32ºF or above 80ºF and 90ºF for COOP stations (e.g., Table 3 in Gray et al. 2010).

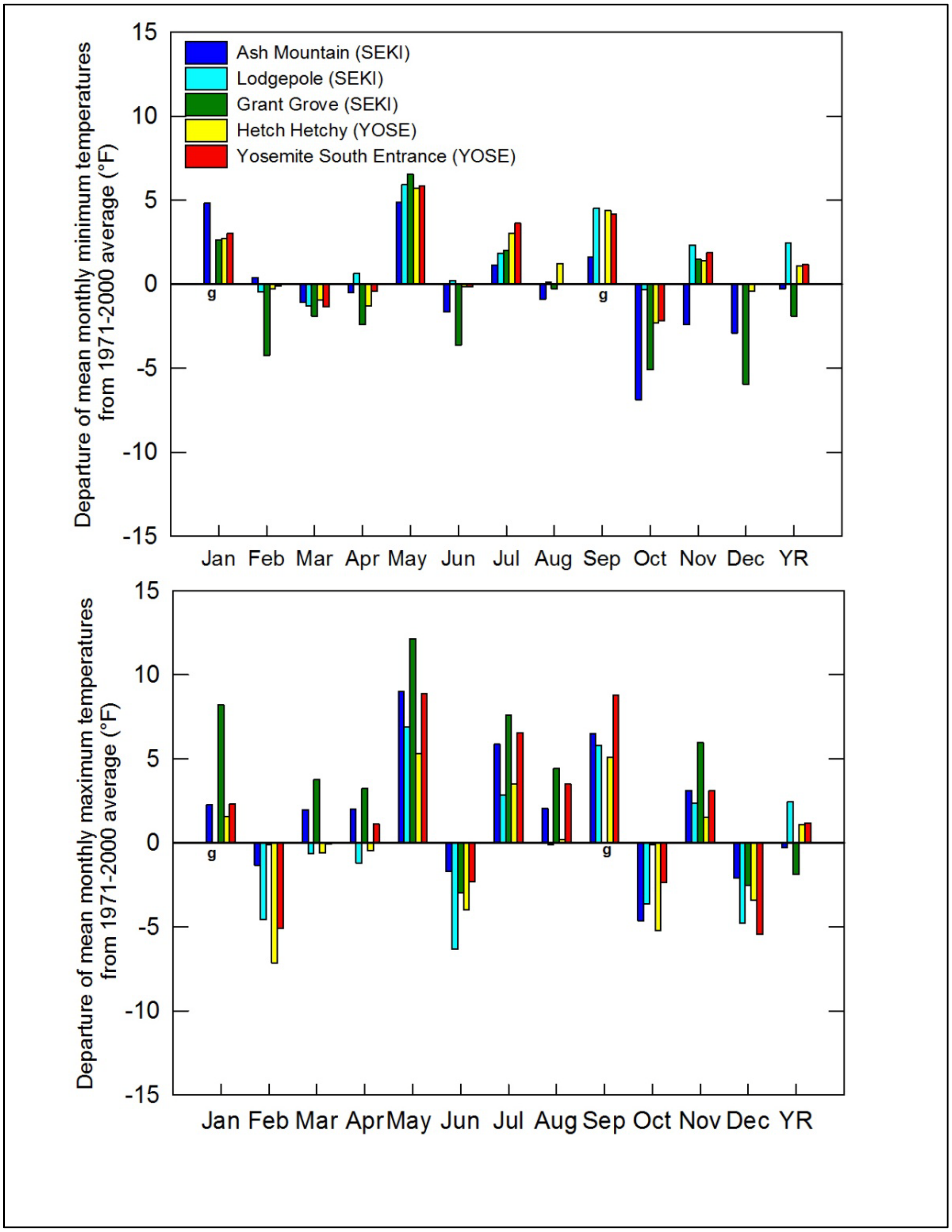


Figure SOP 5.1. Example table to be used in annual status reports that shows the departure of annual temperatures from the 30-year climate normal.

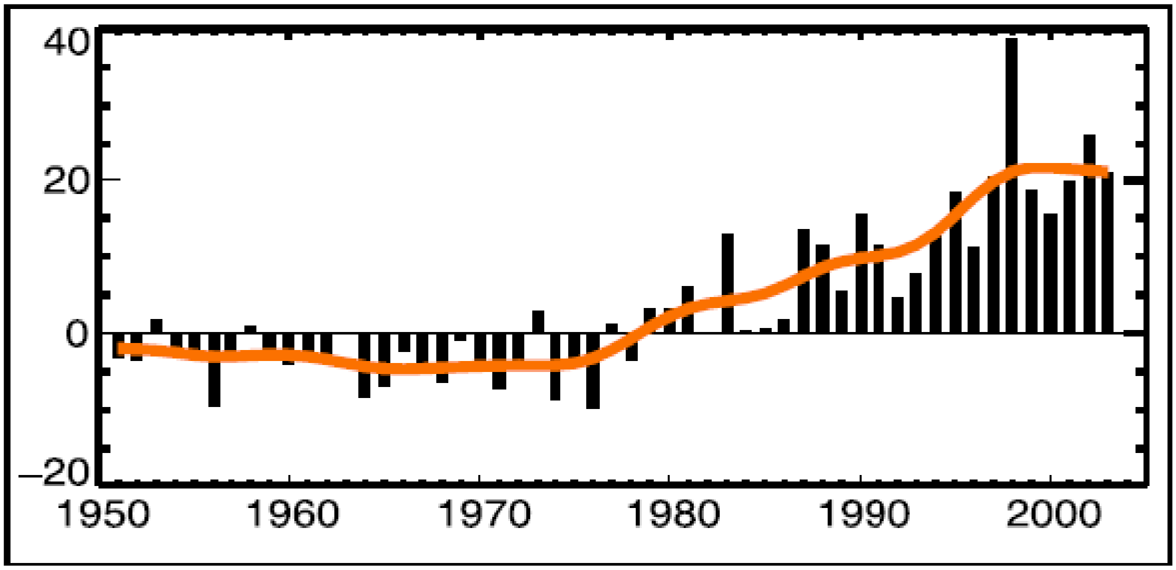


Figure SOP 5.2. Historical change in temperature event frequency. This graph shows the number of days per year with Tmin >90th percentile. The orange solid line represents the moving average. Image taken from Kittel (2009).

4. Daily event structure (Optional: these may be too much detail for an annual status report)
    a. Evaluate daily event structure with frequency distribution plots. As described by Kittel (2009), the analysis of daily records can reveal the characteristic structure of weather events and the frequency of extreme events.
    b. Plot year by day (one box plot per month) for Tmin and Tmax (e.g., Garman 2009) for NWS-COOP stations.

Precipitation

1. Overview for the region
    a. Use the CGAT (Sherrill and Frakes 2011) and SIEN SOP 4 to produce a time series from the PRISM data source to create monthly and annual anomaly histograms for recent 10+ years compared to the 30-year normals. See example anomaly histogram in Ashton et al. 2009.
    b. Monthly total precipitation compared to the 30-year normal PRISM maps for the region (Figure SOP 5.3).

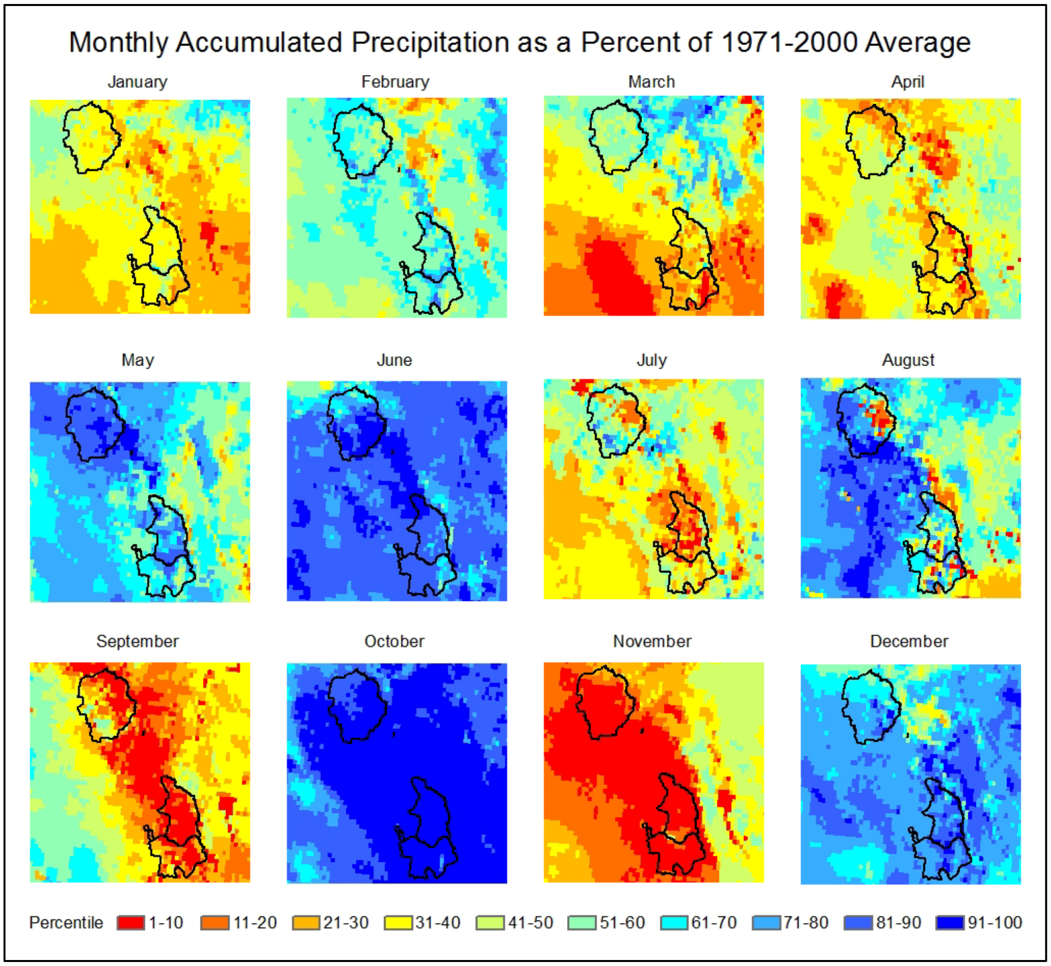


Figure SOP 5.3. Example of a figure to be included in an annual status report that displays monthly accumulated precipitation as a percent of the 30-year average.

2. Station specific data
    a. Tables of total precipitation and departures from 30–year normals per month for each of the selected NWS-COOP stations if available (see Gray et al. 2010).
    b. For NWS-COOP stations, plot year by month along with 30-year normals, or as deviations from normals.
3. Annual Timing Variables
    a. Table or narrative including annual timing and integrative variables: number of days with precipitation, intervals between precipitation events, frequency of precipitation events that exceed a threshold.
4. Daily event structure (OPTIONAL: This may be too much detail for an annual status report).
    a. Plot year by day (one box plot per month) for total precipitation (see Garman 2009) for COOP stations.
    b. OPTIONAL: Frequency distribution graphics (see Kittel 2009).

5. Interannual variability – provisional analysis
    a. OPTIONAL: Append current year's value to long-term plots by year from Trends Report, without updating smoothing function.

Snowpack and Snow Water Equivalent

SIEN is currently working with cooperators to explore additional ways to summarize and report snowpack on an annual basis.

1. Report the April 1 snow water content (SWC) as a percent of historic values for the main SIEN watersheds from the California DWR snow course report - http://cdec.water.ca.gov/cgi-progs/rpts1/COURSES
2. Graphic showing the total monthly and annual snow depth and departures from the 30-year average for NWS-COOP stations that receive snow.

Drought

1. Given the difficulty of accurately defining drought and the strengths and weakness of various drought indices, we will report on two indices, SPI and the U.S. Drought Monitor.
    a. SPI reported for 1, 3, 6, 12 and 24 months. See drought.unl.edu/dm
    b. Drought Monitor maps of the region for each season. See http://lwf.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html

Streamflow

1. For key stations provide graphics or tables of derived annual hydrograph timing variables: peak and minimum flow dates, and snowmelt onset
2. Hydrographs of mean daily flows relative to the average over the period of record.

Correlation among Atmospheric Indices and Climate Parameters

1. OPTIONAL: Update station and circulation values in plots from Trend Report teleconnection analyses
2. OPTIONAL: Provide table showing status of atmospheric indices for the past year with a brief description of implications for the network.

Schedule of Annual Tasks

There are multiple tasks associated with the creation of each annual climate status report. Some tasks could be completed earlier in the year, to ease the work load in April and May when finalized NWS-COOP data will be available from the Climate Database and figures can be compiled into the report.

February

(These tasks could be done in March, but if done here, would give a head start of the annual report)

- Obtain the April course report from CDEC for the previous year and create table
- Obtain PRISM files and run CGAT to produce figures for the annual report (confirm that all months have "green light" and are not in preliminary status on the PRISM website)
- Obtain drought information and create figures. Drought.unl.edu/dm and http://lwf.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html.

March

- Obtain DEPO data from Douglas Alden at SCRIPPS. Format and upload to the Climate Database (Data Manager to complete this task).
- Send the data request form to IMD or an email stating that our request is identical to the year before.

April and May

- Obtain hydrographs from NWIS
- Check WRCC website to see if 1981-2010 30-year normals are available. Calculate 30-year normals using the data from our Climate Database and compare to WRCC.
- Obtain COOP data from Climate Database, perform any needed QA/QC and create figures for the report (refer to SOP 3).
- Compile statistics from Climate Database into SigmaPlot figures or into tables.
- Compile all figures and tables into the report. Send report for peer review and publish.

Interpreting Climate Data

The narrative should provide a succinct interpretation of the year's climate and assist with understanding the dynamics of other park resources. Highlighting the departures from normal conditions may reveal links between climate and other vital signs.

Annual status reports should generally not include discussion of trends, which are covered in depth by the trend reports. Percentiles and ranks can be used to interpret those months or years that are different from average. Generally, values between 20% and 80% are considered average and should be interpreted as such. Discussion and conclusions should highlight conditions that differ from average and present an integrated narrative of conditions across the domain. Network staff may consult with a climatologist (e.g., Kelly Redmond – Regional Climatologist at the Western Regional Climate Center) while writing the report, and/or have a climatologist review early reports to improve the interpretation of any significant or abnormal results.

Limitations of Station and Raster Data

Annual status reports summarize point data from specific locations (weather stations) as well as interpolated raster data for park landscapes, climate zones, and larger regions. SIEN has carefully selected stations from existing climate monitoring programs to provide the best possible foundation for understanding climate status and trends in and around parks, given our use of existing installations only. Given our approach, it is also important for annual status reports to disclose and discuss limitations and exceptions to the reported information, including the following factors that directly affect station data and any derived, generalized, or interpolated values for landscapes between and around stations. Some of these include:

- Spatial representation
  - Effects of micro-scale climate – differences resulting from proximity to water, trees, vents, parking lots, walls, etc. that may affect representativeness of measurements
  - Effects of meso-scale climate – the location of the station relative to individual storm events, especially in summer when convective storms unevenly distribute rain
  - The limited number of stations within SIEN parks and lack of a probabilistic sampling design constrain our analyses and observations to the point scale (selected weather stations) rather than all areas of the park. The limited number of stations also reduces the ability of PRISM models to accurately represent conditions in all areas of the park
- Data Errors – errors related to the instrumentation or recording of the information
- Shortness and discontinuity of records – some stations do not have 30 years worth of observations and/or have substantial data gaps.

To produce robust and continuous climate records and to address some of the factors affecting station data, existing raster data products are processed and reported for daily maximum and minimum temperature and total daily precipitation (PRISM - Daly 2006) and snowpack (SNODAS). Annual status reports should state that while these raster datasets provide useful indicators of climate conditions within a park, they are not representative of any single point within a park and may not correctly represent climate signals relevant to certain ecological processes.

Publishing Reports

Once the report is completed, the project lead will submit it to appropriate park staff and others for review. After review comments and suggestions have been addressed, the network program manager will follow PWR peer review guidelines for publication in the Natural Resource Report Series. All reports will also be prepared in accordance with guidance from the Natural Resource Publications Office at: http://www.nature.nps.gov/publications/NRPM/index.cfm.

Literature Cited

Ashton, I., B. Frakes, D. Pillmore, and J. Burke. 2009. Annual climate report for Grant-Kohrs Ranch National Historic Site, 2007. Natural Resource Report NPS/ROMN/NRR—2009/101. National Park Service, Fort Collins, Colorado.

Ashton, I. W., L. O'Gan, and K. Sherrill. 2010. Climate monitoring in Glacier National Park: Annual report for 2009. Natural Resource Technical Report NPS/ROMN/NRTR—2010/388. National Park Service, Fort Collins, Colorado.

Daly, C. 2006. Guidelines for assessing the suitability of spatial climate data sets. Int. J. Climatol. Vol 26:707-721

Frakes, B., and S. Kingston. 2011. Connecting to the I&M Enterprise Climate Database for the Purpose of Data Retrieval, Summary and/or Analysis. Available from: https://nrinfo.nps.gov/Reference.mvc/Profile?code=2167699 (accessed 20 June 2011).

Garman, S. L. 2009. Climate Monitoring in the Northern Colorado Plateau Network: Annual Report 2007. Natural Resource Technical Report NPS/NCPN/NRTR‒2009/216. National Park Service, Fort Collins, Colorado.

Gray, S. T, C. M. Nicholson, and M. D. Ogden. 2010. Greater Yellowstone Network: Climate of 2008. Natural Resource Report NPS/GRYN/NRR‒2010/173. National Park Service, Fort Collins, Colorado.

Kittel, T. 2009. The Development and Analysis of Climate Datasets for National Park Science and Management: A Guide to Methods for Making Climate Records Useful and Tools to Explore Critical Questions. Final draft report (Dec. 10, 2009) prepared for the National Park Service Inventory and Monitoring Program. University of Colorado, Institute of Arctic and Alpine Research, Boulder, Colorado.

Rocky Mountain Climate Working Group. 2011. Standard Operating Procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sherrill K. R. and Frakes B. 2011. Climate grid analysis toolset – Tools for assessing regional climatological trends: Standard operating procedure (Version 1.2). Natural Resource Report. NPS/NRSS/IMD/NRR—2011/411. National Park Service, Fort Collins, Colorado.

Sierra Nevada Network Climate Reporting Protocol

SOP 6: Trend Analysis and Reporting

Version 1.0

This standard operating procedure is part of the Sierra Nevada Climate Protocol, but is designed to be printed and viewed as a separate document. This document has been adopted in part from 'Rocky Mountain Climate Protocol SOP: Trend Analysis and Reporting' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Background and Approach to Trend Reporting

As described in the protocol narrative, the Climate Variability and Trends Report presents analyses of inter-annual variability, long term historical trends, and teleconnections with hemispheric climate patterns (e.g., the PDO) for temperature, precipitation, snowpack, drought and streamflow. Trend reports will be published in a peer-reviewed climatology journal and/or as a Natural Resource Report through the NPS Natural Resources Stewardship and Science (NRSS) publication series. Additional analyses, graphics or text not included in the report or publications may be posted on the SIEN website.

Kittel et al. (2010) state "agency station data products, as released, are not suitable to describe variability and trends of climate in a scientifically defensible manner. Rather, the analysis of climate variability and trends requires a substantial, but worthwhile investment in data quality control". Therefore, the overall objectives of the trend reporting process are: (1) to provide scientifically-defensible analyses of variability and trends of five climate parameters (temperature, precipitation, snow water equivalent, drought, and streamflow), and (2) to create high-quality climate datasets that can be used as covariate data in the analysis of other vital signs. The purpose of the report is to provide park management, researchers and park interpreters with reliable, periodic and pertinent assessments of changes in park climates. Important differences between the climate status reports and the climate trends reports are: the reporting time frames (annual vs. every 5 years), the number of stations used, and the level of QC (limited vs. extensive).

To describe the variability and trends of a climate in a scientifically-defensible manner requires a substantial investment in quality control (Kittel et al. 2010). Implementing techniques in keeping with climate-community standards is beyond the scope of what can be done with current I&M network staff. Therefore, preparation of the trend reports will require collaboration and/or contracting with a climatologist. Because climate change is a high-profile, contentious topic, in the absence of such expertise, reliance on less than high-quality data for these monitoring goals would pose a strong risk to the I&M Program's credibility. The consensus of outside experts at the 2009 Climate Data Analysis Workshop (Kittel et al. 2010) was that, unless able to implement a protocol on par with climate community-standards, the networks should not attempt to evaluate and report on long-term climate trends.

There are several criteria for successfully reporting on inter-annual variability and trends. First, the dataset creation methods must match the analysis requirements. As described in Kittel et al. (2010), dataset development must be highly integrated with the scientific questions being asked and the intended climate analyses. The specific monitoring objectives of this protocol determine the framework of and analyses needed in the trends report. Each monitoring objective will be addressed analytically or an explanation will be provided as to why the analysis objective could not be met. Other criteria for successfully reporting on inter-annual variability and trends include that the dataset correction and analysis techniques are defensible and follow well-established, best practices of the climate science community and that the development process is transparent and well-documented (Kittel et al. 2010). Ultimately, the report should be suitable for publication in a peer-reviewed climatological journal.

This SOP provides an overview of the trends report preparation process and the anticipated scope and content of trends reports. It is designed to supplement guidance from the I&M network staff during the preparation of the trends reports by a collaborating or contracted climatologist.

Trends Report Data Processing

In considering the overall processing of trends reports, it should be recognized that two products will result from this effort:

> (1) trends reports that will be available in print and online. These reports also provide context for the climate status reports that are completed annually; and
>
> (2) high quality climate datasets that have undergone rigorous quality control measures that are appropriate for the intended analyses. These datasets must be thoroughly documented to track all quality control procedures.

Data development, analysis and reporting tasks include:

- Identification of data sources and station selection per climate zone or reporting region
- Data acquisition
- Data quality control
- Data analyses
- Reporting
- Documentation, archiving and providing access to data

Preparation of the trends reports begins with identifying the key stations that represent the parks in the network (SIEN SOP 1: Station Selection). Reporting will be based on the network parks and surrounding region. After a few key stations per reporting unit have been identified, data should be downloaded and processed according to anticipated analyses (Figure SOP 6.1; Kittel et al 2010, Kittel 2009). To avoid data correction techniques interfering with analyses, it is helpful to lay out correction technique processing steps against the climate dynamical features being evaluated and the assumptions of analytical techniques being applied (Kittel et al. 2010). As previously noted in the protocol narrative, this will lead to different lineages of data being created to address different sets of questions, with different corrections applied or bypassed (Kittel et al. 2010; Kittel 2009).

In identifying specific techniques for data analysis, we refer to The Development and Analysis of Climate Datasets for National Park Science and Management (Kittel 2009) and the current climatology literature. Kittel (2009) provides overall strategies for working with climate data and a primer on techniques for handling data errors, inhomogeneities, and missing values; event, variability and trend analyses; and regional and teleconnection analyses (Figure SOP 6.2). The specific analysis techniques used in trends reports should be subject to ongoing evaluation so that the protocol and standard operating procedures remain current with the climate community's well-established best practices. However, any substantial change to analysis methods must be carefully considered if it will warrant reprocessing of the historical or baseline dataset. Whenever possible, data analysis will be performed using data from the Climate Database and data that have gone through extensive quality control procedures will be uploaded to the database with a new version number and metadata documenting the changes made. Any analysis software and code will be made available to other networks or parks using similar climate protocols.

Trends Report Content

The content of trends reports should generally follow that of a scientific publication and is outlined below. This outline, as well as Kittel et al. (2010), provide general guidance on the scope and content of trends reports; however the specific presentation of the reports may be different from what is suggested here. The narrative of the report should integrate information across variables, and should highlight results at the park or climate zone scale as well as across the network.

A. **Executive summary**
B. **Introduction**
C. **Data acquisition, quality control, and analysis methods**
D. **Results and discussion**
    1) Variability/teleconnections and trends - narrative of the 5 year period from a spatial and temporal perspective.
    2) Daily structure/extreme value analysis - discussion of variation and change in the probability of extreme events and other features of daily frequency distribution.
        a) Interannual variability – discussion exploring temporal patterns by zone (reporting unit) and the domain, integrating across variables and noting spatial connections to:
        b) Region – regional coherence
        c) Hemispheric circulation – teleconnections
        d) Regime shifts – discussion exploring temporal patterns, integrating across variables and noting spatial connections to:
            i) Region – regional coherence
            ii) Hemispheric circulation – teleconnections
    3) Long-term trends – discussion exploring temporal patterns by zone (reporting unit) and the domain, integrating across variables and noting spatial connections to:
        a) Region – regional coherence
        b) Hemispheric circulation – teleconnections
E. **Integrative summary and conclusions**
F. **Literature cited**

A key element of successfully reporting on interannual variability and trends as well as providing useful datasets is making the dataset development process transparent to network and outside users. As described by Kittel et al. (2010) transparency is accomplished through:

- Thorough, up-to-date documentation for each dataset lineage that includes processing methods and their assumptions for station selection, quality control and correction, temporal and spatial aggregation. The intended uses of each dataset as well as any caveats regarding the data should be clearly stated.
- Version control – particularly given that SIEN calls for multiple dataset lineages. Version control should be sufficient to reverse any correction, and should include archiving the original raw data.
- Open online access for the final datasets. This open access permits critical review by outside users, giving another level of quality control.

Upon completion of the trends report, all datasets and the associated metadata are transferred to the data managers to archive and make available directly or by request. Regular communication between the SIEN Data Manager and the climatologist(s) performing the data QC and analysis will be critical to ensure the proper transfer and handling of datasets and metadata.

Requirements for Individuals Completing Trend Reports for SIEN

SIEN will work with a climatologist to complete the first trend report for the network. The methods will be documented and incorporated into this SOP. SIEN may elect to complete some of the analyses independently for subsequent reports if the methods are replicable. All data that has been quality controlled will be formatted and uploaded to the Climate Database as outlined in the IMD SOP: Submitting Data for Upload (Frakes 2012).

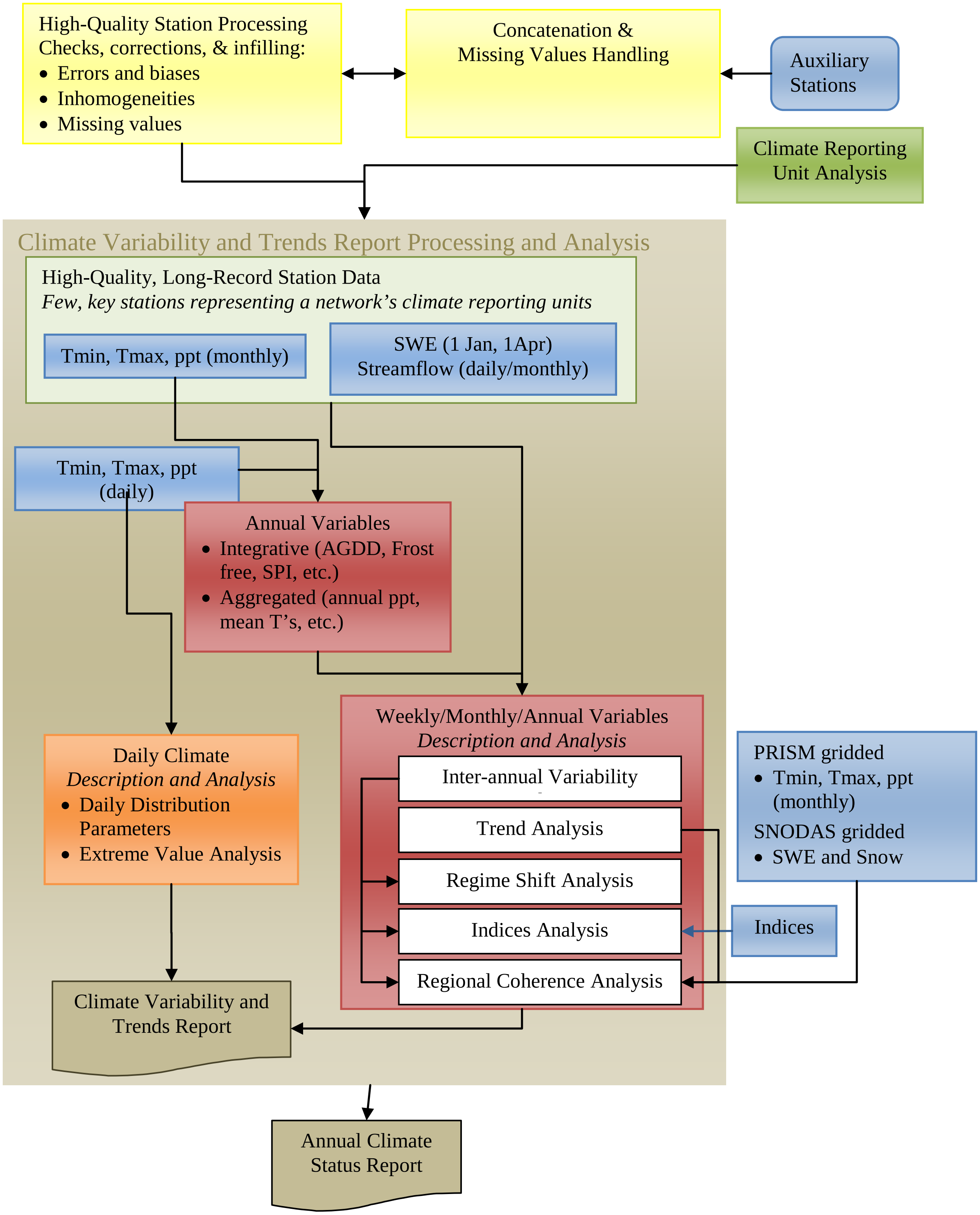


Figure SOP 6.1. Overview of the processing for variability and trends reports (Kittel et al. 2009).

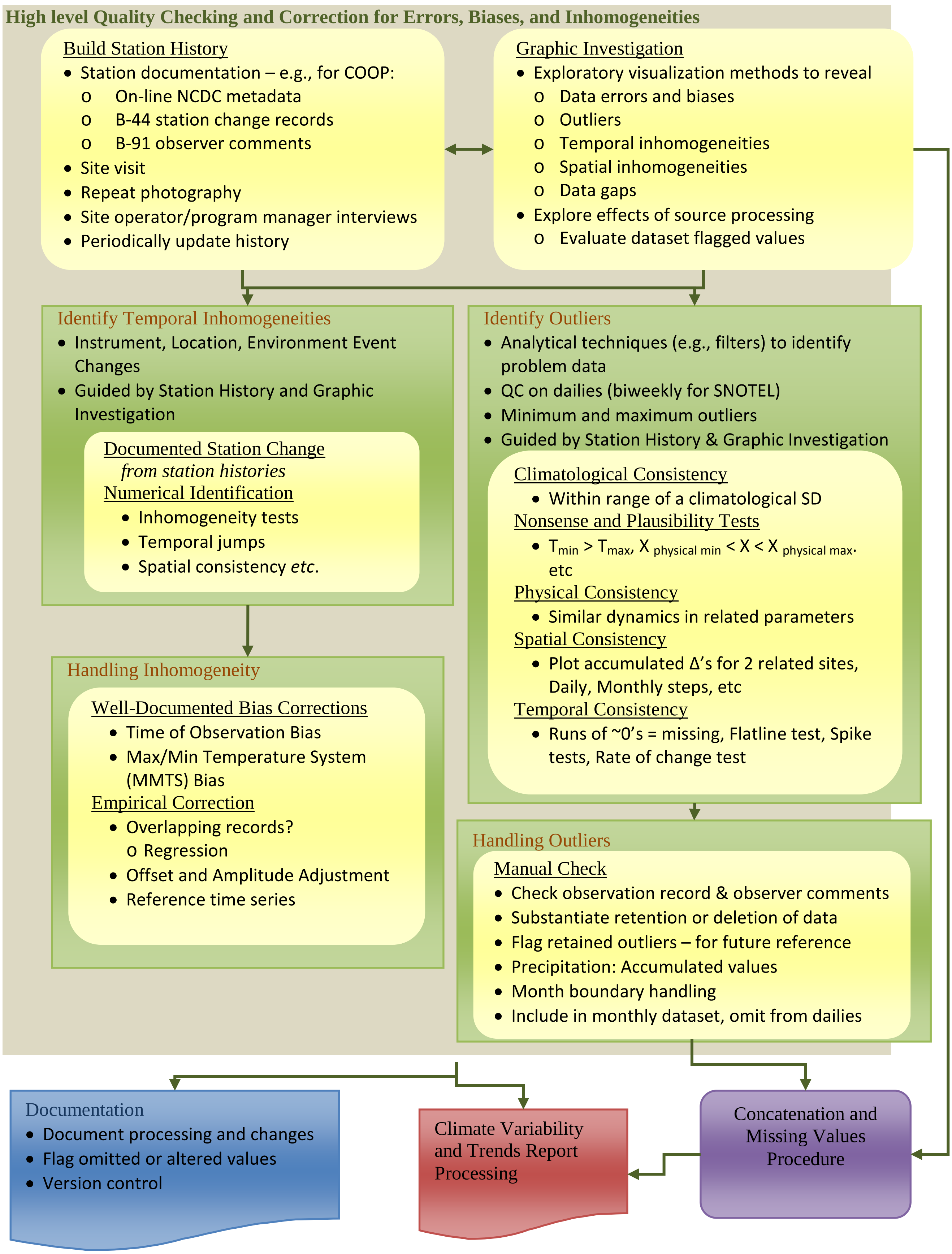


Figure SOP 6.2. Overview of quality control procedures for trends reports (Kittel et al. 2009).

Literature Cited

Frakes, B. 2012. Submitting data for upload to the I&M Enterprise climate database. Unpublished report. National Park Service, Fort Collins, Colorado. Available from: https://irma.nps.gov/App/Reference/Profile?code=2166928 (accessed 30 March 2012).

Kittel, T. 2009. The development and analysis of climate datasets for National Park science and management: A guide to methods for making climate records useful and tools to explore critical questions. Final draft report (Dec. 10, 2009) prepared for the National Park Service Inventory and Monitoring Program. University of Colorado, Institute of Arctic and Alpine Research, Boulder, Colorado.

Kittel, T., S. Ostermann-Kelm, B. Frakes, M. Tercek, S. Gray, and 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, Montana, 7–8 April 2009. Unpublished final report (18 February 2010). National Park Service, Greater Yellowstone Network, Bozeman, Montana.

Rocky Mountain Climate Working Group. 2011. Standard operating procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sierra Nevada Network Climate Reporting Protocol

SOP 7: Sharing Climate Data, Information and Reports

Version 1.0

This standard operating procedure is part of the Sierra Nevada Climate Protocol, but is designed to be printed and viewed as a separate document. This document has been adopted in part from 'Rocky Mountain Climate Protocol SOP Sharing Climate Data, Information and Reports' (Rocky Mountain Climate Working Group 2011).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Introduction

Effectively delivering relevant climate information to parks involves systems, processes, and services that will necessarily evolve to meet the information needs of park staff. Networks will publish and share printed protocols, related standard operating procedures, and scheduled reports in the NPS Integrated Resouce Management Applications (IRMA) data store. Published climate status reports and climate trend reports for network parks will list and describe data sources used, including data quality reviews and results, and how to access both original and quality-checked data sources. Additional climate related data and information may be provided online as network and park staff identify, evaluate, obtain, and appropriately use climate-related data for understanding land and resource management issues and concerns. Some data sets may be accessed by linking directly to the online service of the original data provider (PRISM data is an example). Park personnel and others may require data and information beyond what is used to meet the stated objectives in this protocol. Parks are encouraged to partner directly with networks and the NPS Inventory and Monitoring Division to evaluate and meet these additional requirements.

Sharing Data Used for Reporting

The ‘Data and Methods’ section of each status report describes data used in report preparation (see SIEN SOP 5: Annual Status Reports). Data used in climate trend analysis and reporting are described in the report section dealing with data acquisition, quality control, and analysis methods (see SIEN SOP 6: Trend Analysis and Reporting).

SIEN will maintain instructions for accessing data used for climate reporting on the network website and/or SharePoint sites following current NPS Inventory and Monitoring Division guidelines for consistently distributing information about resource monitoring topics (vital signs) (National Park Service 2009). We will make every attempt to:

- avoid duplicating information that is available from other trusted and stable sources,
- briefly describe the procedures used to identify, obtain, quality-check, process, and use the source data for reporting,
- provide a link to the complete SOP or other documentation for further details
- explain how data acquired, processed, and used for reporting may differ from other available data sources and other reporting procedures so that the reader understands the purpose and limits of data used to meet SIEN climate protocol objectives.

Sharing Published Material

We will publish scheduled status reports and trend reports in the NPS Natural Resource Report Series, upload these to the NRInfo Reference application, and create links to them from including the SIEN internal and external web and SharePoint sites.

Sharing Tools for Data Acquisition, Processing, and Reporting

Many of the data automation and reporting tools developed to meet the protocol objectives are useful for working with the same or other data sources to inform other analyses outside the scope of network activities.

SIEN will provide links to Python code, R scripts, and other useful automation tools on network web and/or SharePoint sites. To prevent duplication and ensure synchrony, SIEN will always identify and link to the master repository of these tools rather than providing a local copy.

SIEN will maintain and provide adequate documentation and instructions for using the available tools.

Sharing Dynamic Information Online

SIEN may develop online content to supplement scheduled reports and help meet park information needs with more frequent and dynamic reporting elements, and/or that centralize relevant information from multiple sources. For example, the network website could provide static or dynamic summary and 'departure from normal' graphs for common parameters such as precipitation and max and min temperature by month. Before developing this type of online content, we will carefully assess whether the work to provide and maintain such provisional and/or final products is feasible for network staff to accomplish and will truly meet park information needs.

Sharing Extended Data and Information Resources

SIEN will provide and maintain a list of links on network web and SharePoint sites to other online climate resources and services. Some examples are listed in Table SOP 7.1. The SIEN web page will include text that encourages users to develop a complete understanding of the relevance, applicability, and limitations of products and information from these web sites that are supplemental to the status and trends reports. The Sierra Nevada Inventory and Monitoring Network website has a page dedicated to each vital sign that will be monitored. The climate page along with external links can be found at:
http://science.nature.nps.gov/im/units/sien/monitoring/WeatherAndClimate/WeatherAndClimate.cfm

Table SOP 7.1. List of online climate resources and services.

General Topic	Websites
Climate forecasts and advisories for California	http://alerts.weather.gov/cap/ca.php?x=1
Climate data	http://www.ncdc.noaa.gov/oa/ncdc.html
Drought information	http://drought.unl.edu/DM/DM_west.htm http://www.drought.gov/portal/server.pt/community/drought.gov/202/area_drought_information?mode=2&region=west&x=13&y=11 http://lwf.ncdc.noaa.gov/oa/climate/research/drought/drought.html http://www.drought.unl.edu/monitor/monitor.htm http://www.cpc.ncep.noaa.gov/products/expert_assessment/seasonal_drought.html
Snow	Latest Snow Sensor Report http://cdec.water.ca.gov/cgi-progs/snow/PAGE6 Latest Statewide Summary of Snow Water Equivalents http://cdec.water.ca.gov/cgi-progs/snow/DLYSWEQ Monthly Snow Course Report http://cdec.water.ca.gov/cgi-progs/snow/COURSES
Precipitation	Latest Northern Sierra 8-Station Precipitation Index http://cdec.water.ca.gov/cgi-progs/queryDaily?s=8SI&d=today Northern Sierra 8-Station Precipitation Tabulation Table http://cdec.water.ca.gov/cgi-progs/products/8-Stations_Tab.pdf Latest San Joaquin 5-Station Precipitation Index http://cdec.water.ca.gov/cgi-progs/queryDaily?s=5SI&d=today Current WY Precipitation Summary http://cdec.water.ca.gov/cgi-progs/precip/PRECIPSUM
Streamflow information for the state	http://waterwatch.usgs.gov/new/index.php?r=ca&m=real&w=real%252Cmap Forecast of Unimpaired Runoff http://cdec.water.ca.gov/cgi-progs/iodir?s=b120 Daily Full Natural Flow http://cdec.water.ca.gov/cgi-progs/snowsurvey_ro/FNF Monthly Full Natural Flow http://cdec.water.ca.gov/cgi-progs/snowsurvey_ro/FNFSUM Seasonal Full Natural Flow http://cdec.water.ca.gov/cgi-progs/snowsurvey_ro/FLOWOUT
Park specific information	http://www.nps.gov/yose/planyourvisit/climate.htm http://www.nps.gov/seki/planyourvisit/weather.htm http://www.nps.gov/depo/planyourvisit/weather.htm

Creating and Maintaining Shared Information

Follow the schedule in Table SOP 7.2 to develop and maintain the required online resources and content.

Table SOP 7.2. List of SIEN online content, assignments and schedules.

Content Element	Responsible Party	Schedule
Overall site design and development	SIEN Data Manager and Science Communication Specialist	Initial and ongoing
Coordinating site content with other online resources	SIEN Physical Scientist & Data Manager working with providers of content on other web sites	Initial and ongoing
List of climate and weather resources	SIEN Physical Scientist	Prepare initial list and update annually or more frequently if necessary
Reporting schedule	SIEN Physical Scientist	Initial and ongoing
Climate status reports	SIEN Physical Scientist, Data Manager and Science Communication Specialist	Provide status reports for posting as soon as they are published/available. Continuously ensure that all current and past status reports are available online.
Climate trends reports	SIEN Physical Scientist, Data Manager and Science Communication Specialist	Provide trends reports for posting as soon as they are published/available. Continuously ensure that all current and past trends reports are available online.
Climate monitoring protocol	SIEN Physical Scientist, Data Manager and Science Communication Specialist	Continuously ensure that all current and past Sierra Nevada Climate Monitoring protocols are available online.
Climate monitoring procedures (SOPs)	SIEN Physical Scientist, Data Manager and Science Communication Specialist	Continuously ensure that all current and past standard operating procedures for the Sierra Nevada Climate Monitoring protocol are available online.
Static data sets used in climate trends reports	SIEN Data Manager	Ensure that fully qualified and documented data sets used in climate trends reporting are accessible directly online or by request to a project contact. Input and derivative data sets meeting these criteria will be made available concurrently with the release of each climate trends report.

Literature Cited

National Park Service. 2009. I&M Network data management web authoring. Available from: http://www1.nrintra.nps.gov/im/datamgmt/webdev/index.cfm (accessed 9/28/2009).

Rocky Mountain Climate Working Group. 2011. Standard operating procedures for the Rocky Mountain Climate Protocol. March 2011 Version. Available from: https://nrinfo.nps.gov/Reference.mvc/DownloadDigitalFile?code=426590&file=NPS_Rocky_Mountain_Climate_Monitoring_SOPs_March_2011.pdf (accessed March 2011).

Sierra Nevada Network Climate Reporting Protocol

SOP 8: Protocol Revisions

Version 1.0

This standard operating procedure is part of the Sierra Nevada Climate Protocol, but is designed to be printed and viewed as a separate document.

This SOP has been adapted from the multi-network white pine monitoring protocol (McKinney et al. 2012).

Change History

Previous version #	Revision date	Revised by	Changes	Justification	Page #'s affected	New version #

Procedures

This monitoring protocol is an actively evaluated and updated document that reflects the latest procedures of the monitoring program. Revisions are expected, and can involve only minor changes with little overall impact or occasional major revisions and course corrections. Evaluation and revision of the protocol is directed by the project leader on an annual basis in association with season close-out. The narrative as well as each SOP has a revision history log whereby changes can be recorded. Older versions of the narrative and SOPs should be archived to ensure proper legacy of past work is maintained. Each revision will require the updating of the version number. Minor changes are recorded as decimal numbers (e.g., 1.0, 1.1, 1.2, etc…). Major changes are recorded as a change in the primary number of the protocol version (e.g., 1.0, 2.0, 3.0, etc…). In some cases, major revisions to the protocol may prompt the need for additional peer-review. The project lead and Program Manager will coordinate this with the Pacific West Region (PWR) I&M Program Manager.

Timing

SOPs can be changed at any time, however it is often most convenient for the changes to be made during reporting and project close-out and to take effect at the beginning of the next calendar year. Changing an SOP during a data collection period should be avoided. However, testing of proposed new methodologies with the existing approaches is highly desirable and provides invaluable information necessary for evaluating whether the new methods are superior and how their measurements will correlate with the previous methods.

Instructions

The following procedures will ensure that both minor and major revisions to this document will align with the monitoring protocol and data management.

1. Discuss proposed changes with other project staff prior to making modifications. It is important to consult with the Data Manager prior to making changes because certain types of changes may jeopardize dataset integrity unless they are planned and executed properly. Also, because certain changes may require altering the database structure or functionality, advance notice of changes is important to minimize disruptions to project operations. Consensus should be reached on who will be making the changes and in what timeframe.
2. Make the agreed-upon changes in the appropriate protocol document. Note that the protocol is split into separate documents for each appendix and SOP. Also note that a change in one document may necessitate other changes elsewhere in the protocol. For example, a change in the narrative may require changes to several SOPs; similarly, renumbering an SOP may mean changing document references in several other documents. Also, the project task list and other appendices may need to be updated to reflect changes in timing or responsibilities for the various project tasks.
3. Document all edits in the Revision History Log embedded at the beginning of the protocol narrative and each SOP. Record changes only in the document being edited (i.e., if there is a change to an SOP, log those changes only in that document and not in the narrative). Record the date of the changes (i.e., the date on which all changes were finalized), author of the revision, a change description and the paragraph(s) and page(s) where changes are made, a brief reason for making the changes, and the new version

number. Version numbers increase incrementally by tenths (e.g., version 1.1, 1.2) for minor changes. To ensure that minor errors noted or recommendations are not lost, changes should be made within 30 days of when they are noted, once the network team has reviewed and approved the recommended changes. Major revisions will be designated with the next whole number (e.g., version 2.0, 3.0).

4. Circulate the changed document for internal review among project staff and cooperators. Minor changes and clarifications will be reviewed in-house. When significant changes in methodology are suggested, revisions will first undergo internal review by the project staff. Additional external review, including, but not limited to, National Park Service staff with appropriate research and statistical expertise, will be required.
5. Upon ratification and finalizing changes:
    a. Ensure that the version date (last saved date field code in the document header) and file name are updated properly throughout the document.
    b. Make a copy of each changed file to the protocol archive folder (i.e., a subfolder under the Protocol folder in the project workspace).
    c. The copied files will be renamed by appending the revision date in YYYYMMDD format. In this manner, the revision date becomes the version number and this copy becomes the "versioned" copy to be archived and distributed.
    d. The current, primary version of the document (i.e., not the versioned document just copied and renamed) does not have a date stamp associated with it.
    e. To avoid unplanned edits to the document, reset the document to read-only by right-clicking on the document in Windows Explorer and checking the appropriate box in the Properties popup.
    f. Inform the Data Manager so the new version number(s) can be incorporated into the project metadata.
6. As appropriate, create PDF files of the versioned documents to post to the Internet and share with others. These PDF files will have the same name and be made from the versioned copy of the file.
7. Send a digital copy of the revised monitoring protocol to the Project Lead. The revised monitoring protocol should also be forwarded to all individuals who had been using a previous version of the affected document. Ensure that field staff has a hardcopy of the new version.
8. The Network Data Manager will place a copy of the revised protocol in the proper folder on Network shared drives. In addition, the Network Data Manager will archive the previous version on the archive drives.

The Network Data Manager will post the revised version and update the associated records in the proper I&M databases, including but not limited to Reference Services, Network Intranet and Internet websites, and the Protocol database.

Literature Cited

McKinney, S. T., T. Rodhouse, L. Chow, A. Chung-MacCoubrey, G. Dicus, L. Garrett, K. Irvine, S. Mohren, D. Odion, D. Sarr, and L. A. Starcevich. 2012. Monitoring white pine (*Pinus albicaulis*, *P. balfouriana*, *P. flexilis*) community dynamics in the Pacific West Region - Klamath, Sierra Nevada, and Upper Columbia Basin Networks: Narrative version 1.0. Natural Resource Report NPS/PWR/NRR—2012/532. National Park Service, Fort Collins, Colorado.

NPS 963/115248, June 2012

National Park Service
U.S. Department of the Interior

Natural Resource Stewardship and Science
1201 Oakridge Drive, Suite 150
Fort Collins, CO 80525

www.nature.nps.gov

EXPERIENCE YOUR AMERICA ™