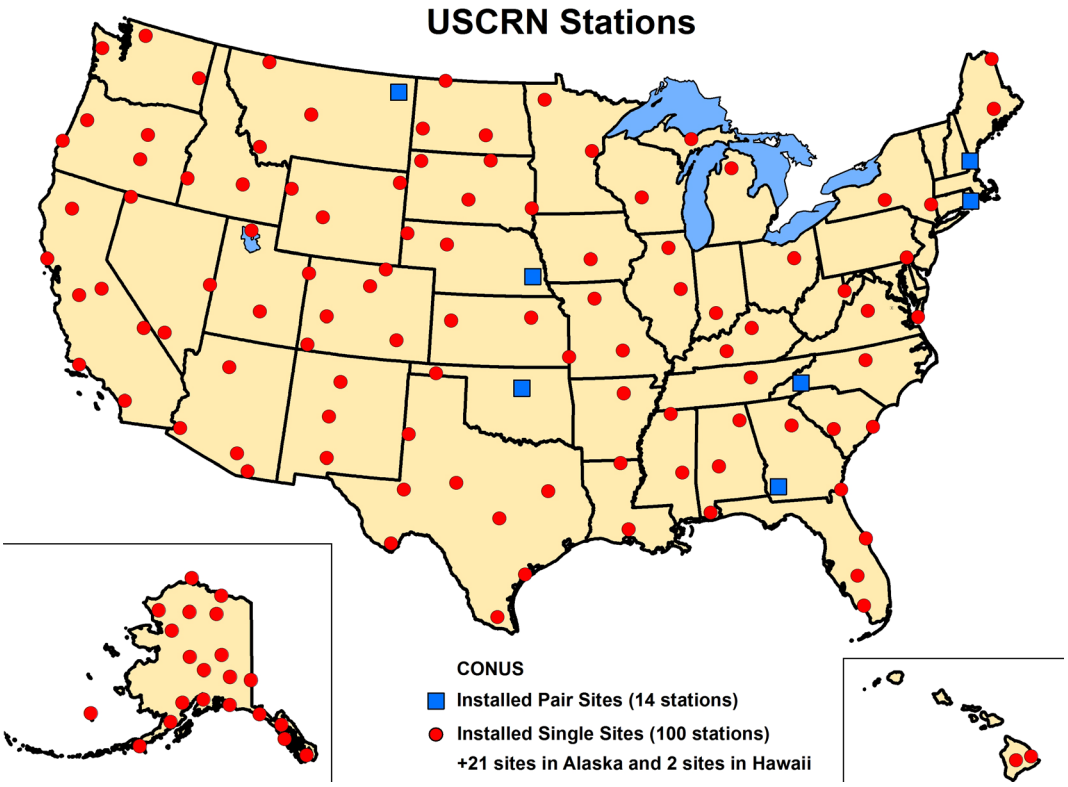
US Climate Reference Network

Data Management Plan
August 2017



Compiled by the
U.S. Climate Reference Network (USCRN)
Program

Compiled on behalf of NOAA by:

OAR/Air Resources Laboratory (ARL)/Atmospheric Turbulence and Diffusion Division (ATDD) 456 S. Illinois Avenue
Oak Ridge, TN 37830

Howard J. Diamond, USCRN Program Manager
e-mail: howard.diamond@noaa.gov

C. Bruce Baker, Director,
ARL/ATDD
e-mail: bruce.baker@noaa.gov

Report posted on the USCRN Website at

[This report supersedes the original data management plan for USCRN from September 2012]

https://www.ncdc.noaa.gov/crn/documentation.html

U.S. Climate Reference Network Data Management Plan

August 2017

1. General Description of Data to be Managed

1.1. Name of the Data, data collection Project, or data-producing Program:

U.S. Climate Reference Network (USCRN)

1.2. Summary description of the data:

Raw USCRN data are transmitted continuously via the Geostationary Orbiting Environmental Satellite (GOES) datastream, and sometimes via NOAAPort, to the National Centers for Environmental Information (NCEI). The raw USCRN data collected from station dataloggers in the field during site maintenance visits. These data are placed on an anonymous FTP site from where they are pulled by NCEI. These raw data are also in a different format than data transmitted through the GOES datastream.

Derived geophysical parameters with other quality indicators are processed from the raw data by the USCRN Team and produce monthly, daily, hourly, and sub-hourly data files of air temperature, precipitation, soil temperature and moisture, solar radiation, and skin (ground) temperature data. The raw data input may include PDA backup device files and/or raw data from GOES and NOAAPort. These data records are versioned based on the processing methods and algorithms used for the derivations, and data are updated when more accurate and complete raw data become available from stations' datalogger storage PDA backup files.

1.3. Is this a one-time data collection, or an on-going series of measurements?

On-going.

1.4. Actual or planned temporal coverage of the data:

The network was initially commissioned in 2004, and the full conterminous U.S. network was completely commissioned in 2009. Stations are now being installed in Alaska, and as of this point, a total of 21 stations have been installed with 19 of those being formally commissioned as of FY 2017. Data collected and summarized on timescales of 5-minutes, hourly, daily, and monthly[1]. The network is intended to provide national climate scale data for the next 50-100 years.

1.5. Actual or planned geographic coverage of the data:

Conterminous U.S., Alaska, and Hawaii.

[1] It is planned that a total of 29 stations in Alaska will be commissioned by 2023

1.6. Type(s) of data:

Digital numeric data and Metadata including photographs

1.7. Data collection method(s):

Data are collected from automated *in-situ* climate stations via satellite communication and direct downloads from station datalogger during annual maintenance visits.

1.8. If the data are from a NOAA Observing System of Record, indicate name of system:

U.S. Climate Reference Network

1.8.1. If data are from a another observing system, please specify:

N/A

2. Points of Contact for this Data Management Plan (author or maintainer)

2.1. Name: **Howard J. Diamond**

2.2. Title: **U.S. Climate Reference Network Program Manager**

2.3. Affiliation or facility: **OAR/Air Resources Laboratory/Atmospheric Turbulence and Diffusion Division**

2.4. E-mail address: **howard.diamond@noaa.gov**

2.5. Phone number: **301-427-2475**

3. Responsible Party for Data Management

3.1. Name: **Jay Lawrimore**

3.2. Title: **Chief, Dataset Section, Center for Weather and Climate, NOAA's National Centers for Environmental Information (NCEI)**

3.3. E-mail address: **jay.lawrimore@noaa.gov**

3.5. Phone number: **828-271-4750**

4. **Resources** – *Programs must identify resources with their own budget for managing the data they produce.*

4.1. Have resources for management of these data been identified?

Yes.

4.2. Approximate percentage of the budget for these data devoted to data management (specify percentage or "unknown"):

10-15%

5. Data Lineage and Quality

NOAA has issued Information Quality Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information which it disseminates.

5.1. Processing workflow of the data from collection or acquisition to making it publicly accessible (describe or provide URL of description):

See formal workflow description at https://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/CRN_Ingest_Functional_Spec.pdf

5.1.1.

5.2. If data at different stages of the workflow, or products derived from these data, are subject to a separate data management plan, provide reference to other plan:

N/A

5.3. Quality control procedures employed (describe or provide URL of description):

See https://www.ncdc.noaa.gov/crn/documentation.html for several QC procedures.

6. Data Documentation

6.1. Does metadata comply with EDMC Data Documentation directive?

Yes. NCEI designed a metadata repository tailored to the needs of USCRN and its stakeholders. The Integrated Station Information System (ISIS) provides archive and access for all metadata associated with the USCRN network including station id, geolocation information, site selection information, agency partners, dates of installation, commissioning, and maintenance. All event details associated with scheduled and unscheduled maintenance visits and any factor that would affect the performance of a station is included. Included are details regarding instrumentation including manufacturer, model and serial numbers, and calibration coefficients for each sensor on the USCRN platform.

In addition to discovery-level metadata, we have additional metadata or other documentation is necessary to fully describe the data and ensure its long-term usefulness. During site survey, installation, and each subsequent annual maintenance visit, dozens of aspects of the USCRN instrumentation, site, and surrounding area are photo documented. Photographs and other documentation such as engineer site visit forms and site survey forms are stored in NCEI's Image and Publications (IPS) system. This system supports digital imaging of photography, maps, forms, etc. The photographs are reviewed by project scientists to identify changes to the station environment and other factors that may alter the true climate signal.

6.2. What Name of organization or facility providing metadata hosting: **NOAA/NCEI**

6.3. URL of metadata folder or data catalog, if known:
https://www.ncdc.noaa.gov/isis/stationlist?networkid=1

6.4. Process for producing and maintaining metadata *(describe or provide URL of description)*:

USCRN standard operating procedures include monitoring data for potential problems with station instrumentation and equipment. Tracking and resolution of such problems is managed in NCEI's Anomaly Tracking System, an internal web-based system that enables engineers and quality control specialists the ability to log system anomalies and track their progress toward resolution. In addition, metadata central to the USCRN network as a whole, are maintained as part of the Archive Dataset Documentation. The storage and access of these metadata vary by the type. All are under the authority of the NCEI Data Administrator. These metadata include information about the data ingest, data processing, and data storage, commissioning information, site survey processes, and technical manuals for the suite of instruments, software and dataset documentation. Furthermore, software used to process the data are identified as Configuration Items, and are under formal NCEI Configuration Management. When actions are taken that change the configuration of the network as a whole, such as modifying the Quality Control algorithms, an explanation is added to the Archive Dataset Documentation.

7. Data Access

NAO 212-15 states that access to environmental data may only be restricted when distribution is explicitly limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements. The EDMC Data Access Procedural Directive contains specific guidance, recommends the use of open-standard, interoperable, non-proprietary web services, provides information about resources and tools to enable data access, and includes a Waiver to be submitted to justify any approach other than full, unrestricted public access.

7.1. Do these data comply with the Data Access directive? **Yes.**

7.1.1. If the data are not to be made available to the public at all, or with limitations, has a Waiver (Appendix A of Data Access directive) been filed? **N/A**

7.1.2. If there are limitations to public data access, describe how data are protected from unauthorized access or disclosure: **N/A**

7.2. Name of organization of facility providing data access: **NOAA/NCEI**

7.2.1. If data hosting service is needed, please indicate: **N/A**

7.2.2. URL of data access service, if known: **https://www.ncdc.noaa.gov/crn/data.html**

7.3. Data access methods or services offered: **HTTP and FTP**

7.4. Approximate delay between data collection and dissemination: **One-Hour**

7.4.1 If delay is longer than latency of automated processing, indicate under what authority data access is delayed: **N/A**

8. Data Preservation and Protection

` *The NOAA Procedure for Scientific Records Appraisal and Archive Approval14 describes how to identify, appraise and decide what scientific records are to be preserved in a NOAA archive.*

8.1. Actual or planned long-term data archive location: **NOAA/NCEI**

8.1.1. If World Data Center or Other, specify: **World Data Center for Meteorology at NCEI.**

8.1.2. If to Be Determined, Unable to Archive or No Archiving Intended, explain: **N/A**

8.2. Data storage facility prior to being sent to an archive facility (if any): **None**

8.3. Approximate delay between data collection and submission to an archive facility: **One-Hour**

8.4. How will the data be protected from accidental or malicious modification or deletion prior to receipt by the archive? Discuss data back-up, disaster recovery/contingency planning, and off-site data storage relevant to the data collection:

USCRN data are transmitted via GOES satellite to NCEI each hour and available to the public on the USCRN website within the hour. In addition preliminary temperature and precipitation observations are satellite transmitted to MADIS for distribution in support of forecasting and other near real-time requirements. There is no hold on the data provided at NCEI however the quality of the data are higher than the feed to MADIS because data collected at NCEI are subjected to the full suite of automated processing and quality control algorithms associated with calculating official USCRN measurements from the triplicate sensor configuration. Also USCRN data are stored on the on-site datalogger and retrieved during annual maintenance visits. These data are used to confirm observations collected via satellite transmission.

NCEI's data holdings are archived on a Hierarchical Data Storage System (HDSS), which includes a tape robotics system for data archived on tape. NCEI provides direct online access to these data though the Archive Information Request System (AIRS). Other services for archived data are available elsewhere on the NCEI website, so this is not a compilation of all HDSS-archived datasets. NCEI is a recognized world-class data center with over 25 petabytes of data stored. In that they employ the latest and best practices in ensuring that data back-up, disaster recovery/contingency planning, and off-site data storage practices relevant to the data collection.

9. Additional Line Office or Staff Office Questions

N/A

End of Report

USCRN Station in Denali National Park and Reserve in Alaska