Central Asia Temperature and Precipitation Data, 1879-2003, Version 1

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Williams, M. W. and V. G. Konovalov. 2008. *Central Asia Temperature and Precipitation Data, 1879-2003, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5NK3BZ8. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/G02174

TABLE OF CONTENTS

1 OVERVIEW

This data set updates and expands the NOAA Global Historical Climate Network (GHCN) of quality controlled meteorological records, focusing on the Northern Tien Shan and Pamir Mountain Ranges of Central Asia. It is compiled primarily from meteorological measurements conducted by the National Hydrometeorological Services (NHMS) of the Central Asian countries. Precipitation data are monthly sums bias corrected for gauge type and for wetting, but not for wind. The correction factors, K1 for gauge type, K2 for wind, and K3 for wetting, are included as separate files. Temperature data are monthly means, for example, the mean of daily temperatures for that month, where daily temperature is defined as the average of all observations for each calendar day. For many stations, average maximum and average minimum temperature are supplied as well. These are derived from daily maximum and minimum temperatures. The station metadata for this data set are station histories, population, vegetation, and topography. Data were subjected to rigorous quality control and homogeneity assessment procedures, consistent with those used for the GHCN.

There are records from 298 stations. The period of record covered by each station is variable, and in the period from 1985 to 1995, there was a sharp reduction in the number of operating stations. The earliest record is from 1879. Records are updated through 2003 where data are available. Most stations have almost 100 years of observations. Records are from stations in Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan.

The Global Historical Climate Network product and data from Central Asia

Climate research relies heavily on meteorological data from surface stations. Usually, stations are part of a national network, and some or all of these stations may be in the World Meteorological Organization network of stations, reporting observations over the Global Telecommunication System (GTS). The instrumental records from surface stations are generally kept to assess local conditions and are not always suitable for climate studies. For instance, the records are often not digitized or are not readily available outside of the country in which they were measured. An uneven distribution of stations introduces network biases that have significant effects on estimated temperature trends. Instrumental records also often contain data errors resultant from the data recording and archiving processes. These errors take many forms (e.g., outliers, truncations). In addition, instrumental records are subject to inhomogeneities caused by many factors. For example, stations are relocated with attending changes in surrounding topography or vegetation, new instrumentation is introduced, personnel changes may result in subtle changes in operating procedures, and populations may grow up around stations, introducing a heat island effect. Such inhomogeneities introduce non-climatic variation into historical records.

The GHCN of monthly data is often used for climate studies (Jones and Moberg 2003) because the data records it contains have been quality checked. Both historical and near-real-time GHCN data undergo rigorous quality assurance reviews. These reviews include preprocessing checks on source data, time series checks that identify spurious changes in the mean and variance, spatial comparisons that verify the accuracy of the climatological mean and the seasonal cycle, and neighbor checks that identify outliers from both a serial and a spatial perspective. From GHCN Monthly Verison 2.

Version 1 of the GHCN was released in 1992 and Version 2 in 1997. Version 2 enhancements include:

- data for additional stations to improve regional-scale analyses, particularly in previously data sparse areas
- the addition of maximum/minimum temperature data, to provide climate information not available in mean temperature data alone
- detailed assessments of data quality to increase the confidence in research results
- rigorous and objective homogeneity adjustments to decrease the effect of non-climatic factors on the time series
- detailed metadata such as population, vegetation, and topography that allow more detailed analyses to be conducted
- an infrastructure for updating the archive at regular intervals so that current climatic conditions can constantly be put into historical perspective (Peterson and Vose, 1997).

There are relatively few high elevation climate station records in the GHCN: most stations are below 500m elevation. This gap is significant, because mountainous areas drive synoptic climatology, and climate cannot be modeled well in these areas without station data. To address this problem for Central Asia, data from state or country official Monthly Reference Books, or from other data collections, were acquired. Most observations were originally recorded on paper at the meteorological station. Most records were provided to the data set authors in digital form, but some required digitization. In some cases digital archives of daily data or operational data from the GTS were used. Appendix 1 lists references for data sources, as well as some related references.

GHCN v.1 has not been updated since 1992, and the Central Asia records generally end in 1988. GHCN v.2 is being updated regularly by the National Climatic Data Center (NCDC), but of the relatively few Central Asia stations in GHCN v.2, most have records only to 1993 or earlier. The authors of this data set have filled a gap in the climatological record by making high elevation station data from Central Asia available after extensive quality control. They have accounted for inhomogenieties in the station air temperature data by standardizing data and removing values more than three standard deviations outside the mean. Precipitation data have been corrected (for gauge type and wetting but not for wind) using standard methods and coefficients. The authors evaluated for spatial homogeneity in temperature data by using multifactor equations based on latitude, longitude, and elevation.

Table 1 shows the GHCN stations to which this data set adds data. Table 2 also gives an overview of the data coverage in this data set. However, Appendix 7 has more detailed coverage information. In Table 1 and Table 2, the country codes correspond to these countries: 211 = Kazakhstan, 213 = Kyrgyzstan, 227 = Tajikistan, 229 = Turkmenistan, and 231 = Uzbekistan. In all, this data set adds to 75 stations already existing in GHCN v.1 or v.2, and also provides data from 223 new stations.

In Table 1, a cross marks those stations for which station records were extended or gaps were filled by this data set. From an unpublished report by Williams and Konovalov, 2005.

Table 1. Central Asia Stations in GHCN v.1 and GHCN v.2

		Meteorological		Air temperature		Precipitation		Max air temperature	Min air temperature
NN	Country	station	Alt, m	GHCN v.1	GHCN v.2	GHCN v.1	GHCN v.2	GHCN v.2	GHCN v.2
1	211	Almaty	847	+	+	+	+	+	+
2	211	Aralskoe more	62	+	+	+	+		
3	211	Bakhty	221			+			
4	211	Balhash	398	+	+	+	+		+
5	211	Chimkent	606			+			
6	211	Fort Shevchenko	-25	+	+	+	+		
7	211	Gur'ev	-24	+	+	+	+		
8	211	Ili	455			+			
9	211	Irgiz	117		+	+	+		
10	211	Kalmykovo	1		+		+		
11	211	Karaganda	554	+	+	+	+		+
12	211	Kara-tyurek	2600			+			
13	211	Karkaralinsk	810			+			
14	211	Karsakpai	505	+	+	+	+		
15	211	Katon-Karagai	1081			+			
16	211	Kazalinsk	66	+	+	+	+		
17	211	Kizildzhar	361	+		+			
18	211	Kokpekty	512		+	+	+		
19	211	Kyzyl-orda	128	+	+		+		
20	211	Mointy	582			+			
21	211	Panfilov	641		+		+		
22	211	Sam	86	+	+	+	+		
23	211	Semipalatinsk	202	+	+	+	+		+
24	211	Taldy-Kurgan	601			+			
25	211	Turgai	133	+	+	+	+		
26	211	Turkestan	207	+	+		+		
27	211	Uch-Aral	397		+		+		
28	211	Uchtobe	421			+			
29	211	Uil	128		+		+		
30	211	Urdzhar	489			+			
31	211	Uyuk	373			+			
32	211	Zaisan	603		+		+		
33	213	Bishkek	823	+	+		+	+	+
34	213	Dzhergetal	1800			+			
35	213	Irkeshtam	2819			+			
36	213	Naryn	2039	+	+	+	+	+	+

37	213	Osh	1016			+			
38	213	Przhevalsk	1716			+			
39	213	Rybach'e	1660			+			
40	213	Talas	1217			+			
41	213	Tien-Shan	3614			+			
42	213	Tokmak	816			+			
43	227	Altynmazar	2782	+		+			
44	227	Dushanbe	790	+	+		+	+	+
45	227	Horog	2075	+	+		+	+	+
46	227	Irht	3290			+		+	
47	227	Kalai-Khumb	1284			+			
48	227	Kurgan-tyube	426		+	+	+	+	+
49	227	Leninabad	425	+	+	+	+	+	+
50	227	Murgab	3576			+			
51	227	Pendjikent	1016			+			
52	227	Tavildara	1616			+			
53	229	Ashhabad	227	+	+	+	+		+
54	229	Bairam-ali	240	+	+		+		+
55	229	Chardzhou	193	+	+	+	+		+
56	229	Gasan-kuli	23	+	+		+		+
57	229	Kerki	241			+			
58	229	Kizyl-arvat	97	+	+		+		+
59	229	Krasnovodsk	89	+	+	+	+		+
60	229	Kushka	57	+	+	+	+		+
61	229	Repetek	29			+			
62	229	Serahs	279		+		+		+
63	231	Andizhan	476	+		+			
64	231	Buzaubai	97	+					
65	231	Chimbai	65	+	+	+	+		
66	231	Dzhizak	344			+			+
67	231	Fergana	578	+	+	+	+	+	
68	231	Muinak	68			+			
69	231	Nukus	75			+			
70	231	Samarkand	726	+	+	+	+	+	+
71	231	Tamdy	263	+	+	+	+	+	+
72	231	Tashkent	477		+		+	+	+
73	231	Termez	309	+	+		+	+	+
74	231	Turtkul	109			+			
75	231	Urgench	100			+			
		In total		34	40	58	40	13	23

In Table 2, the **NN** column has the number of stations that can now be added to the GHCN, the **Years/st** column is the number of station years for each country, and the **% gaps** column is the percentage of the time covered by the records for that country where data are absent, from Williams and Konovalov (2005).

Table 2. Data Set Coverage Statistics

Country code	Mean air temperature			Precipitation			Max air temperature			Min air temperature		
	NN	Years/st	% gaps	NN	Years/st	% gaps	NN	Years/st	% gaps	NN	Years/st	% gaps
211	27	2261	2.2	47	3628	3.7	12	613	1.8	12	639	1.8
213	45	2404	2.3	53	3007	2.5	22	755	3.3	22	767	3.0
227	46	2503	1.3	40	2167	1.3	43	2069	1.6	34	1455	2.0
229	12	1145	6.3	14	1251	4.2	0	0	0	8	709	9.3
231	64	3708	2.2	116	6308	3.9	30	1367	1.7	11	685	2.0
In total	194	12021		270	16361		107	4804		87	4255	
≥1000 m	97			116			76			65		

1.1 Differences with GHCN Data

These data have undergone rigorous quality assessment like data in the GHCN. However there are some differences: GHCN data have been quality controlled by a variety of methods (Peterson et al., 1998) and data points that fail the QC tests are not included in the data set. In addition to the quality controlled data, GHCN includes a quality controlled and homogeneity adjusted version. This version makes historical data homogeneous with present day observations by adjusting for non-climatic discontinuities. Homogeniety adjusted data can be better for looking at regional scale long term trends, but using adjusted data has some disadvantages. In particular, adjusted station data can be less representative of local microclimate than non-adjusted data (Peterson and Vose, 1997). This Central Asia data set has only one version of temperature data, and these data have not been adjusted for homogeneity following GHCN practices. Instead, temperature values are checked for inconsistencies relative to station record statistics as described in a later section.

For any given station, the GHCN may have more than one temperature time series (Peterson and Vose, 1997) and each time series taken singly may not cover the entire history of the station. One reason for this is that there may be more that one series, and generally the method of calculating the monthly mean temperature at a station is not readily available. (Peterson and Vose, 1997) point out that differences in temperature attributable to calculating the mean using two different methods at a particular station can be greater than the temperature difference between two neighboring stations. Since it is not known how the differences came to be, all the series are retained. In this Central Asia data set, there is only one record for each station. This is because data were originally recorded following precise instructions, therefore it was rare to have more than one version of a station record.

The GHCN data set includes raw precipitation as well as adjusted data, while this data set consists of only adjusted precipitation data.

1.2 History of Meteorological Observations in the Former Soviet Union

All meteorological stations in the former Soviet Union measured climate parameters using the same instruments and sampling frequency. Observations were recorded by hand in field notebooks. Monthly observations books were then mailed to the Computing and Processing Department of the National Hydrometeorological Services (NHMS) for each state, where they were stored. Data were then entered onto magnetic tape using a format that translates as Language of Data Description. The main archival facility for Central Asia prior to 1991 was the Regional Computing and Processing Center at Tashkent, Uzbekistan. Prior to 1991, a subset of data from Central Asia was archived at the Russian Center of Hydrometeorological Information in Obninsk. The data set authors worked with the Regional Computing and Processing Center at Tashkent, Uzbekistan to download archived meteorological information from magnetic tape to ASCII. Other data sources were used as well.

Data are no longer sent to the Russian Center of Hydrometeorological Information in Obninsk nor the center in Tashkent. Access to the meteorological information collected since 1991 differs by country within Central Asia. The following information was current in 2005.

- **Kazakhstan**: Meteorological data is free for non-profit entities registered with the Federal **Administration**. Some but not all of the data is saved in electronic format.
- **Krygyztan**: Meteorological information is stored only in paper format, there is only one master copy, and is only available to employees of the Kyrgyzhdromet.
- **Tadjikstan**: Same as Krygyztan.
- **Turmenistan**: Meteorological data is free for non-profit entities; it is stored only in paper format.
- **Uzbekistan**: Same as Kazakhstan.

After the breakup of the Soviet Union in 1991, most meteorological stations continued operation. There have been some reductions since that time. For example, in 1980 there were 312 stations measuring air temperature and precipitation amount in Central Asia. By 1996, the number of stations decreased to 248.

2 DETAILED DATA DESCRIPTION

2.1 Format

The data are stored as tab-delimited ASCII text format, Microsoft Excel, and PDF.

Each data sheet contains a table with columns containing the following:

- Three-digit WMO country code
- Five-digit WMO station number
- Station or post name
- Coordinates (longitude and latitude in decimal degree)
- Altitude (meters above sea level)
- Year
- Temperature (degrees Celsius) or precipitation (mm) values for each month, in separate columns with Roman numerals for months

The metadata sheet for each files contains the same contents as listed above along with the first and last year of record, duration of record in years, percentage of the record with no data (data gaps), and percentage of data completeness for each month for each station.

Missing data are indicated by -999.0.

2.2 File Naming Convention

Table 3. File Naming Convention and Contents

File Name	Description
Taver_v1.xls	Excel file with a Taver sheet containing monthly average temperature by station and year, and a Info_Taver sheet containing information about each station record. The metadata sheet Info_Taver contains the same contents as listed above along with the first and last year of record, duration of record in years, percentage of the record with no data (data gaps), and percentage of data completeness for each month and for each station.
Taver_v1_1.txt	The Taver sheet of the Excel file, as a tab delimited text file.
Taver_v1_2.txt	The Info_Taver sheet of the Excel file, as a tab delimited text file.
Tmin_Tmax_v1.xls	Excel file with four sheets for monthly mean minimum temperature, information about the monthly mean minimum temperature records, monthly mean maximum temperature, and information about the monthly maximum temperature record. The format is the same as for Taver_v1.xls.
Tmin_v1_1.txt Tmin_v1_2.txt Tmax_v1_3.txt Tmax_v1_4.txt	Text files for the above sheets.
Precip_v1.xls	Excel file with a Precip sheet containing monthly precipitation sums by station and year, and a Info_prec sheet containing information about each record including start and end year and percentage complete. The format is the same as for Taver_v1.xls.

File Name	Description
Precip_v1_1.txt	The Precip sheet of the Excel file, as a tab delimited text file.
Precip_v1_2.txt	The Info_prec sheet of the Excel file, as a tab delimited text file.
Appendix_1.pdf	Adobe PDF file with a table of all the sources used for data, and related references, in Russian (Cyrillic) as well as English.
Appendix_2.xls	Excel file of station metadata such as location and surroundings. The first six columns are for the data files. INFO (column G) has PRCP to indicate when the stations' precipitation gauge with Nipher shield was replaced by a Tretyakov precipitation gauge, with the year, month and day of the change in columns H, I, and J. The INFO column also has a MOVE indicator. If MOVE is followed by a year in column H and –9s in columns I and J, 0 in K, and –99 in L, the station has remained in the same spot. -9 is also used for missing data. If MOVE is followed by other information, that information is on the year, month, and day of the move, the cardinal direction of the move, and the distance in km. If the Distance column contains 0, the relocation was by less than 100m. Column M gives information about the surroundings, while columns N-Q give the population for stations in inhabited areas for census years 1959, 1979, 1979, and 1991, in thousands.
Appendix_2_1.txt	A tab delimited text file based on Appendix_2.xls.
Appendix_3.pdf	Adobe PDF file with a table of all the sources used for the metadata in Appendix 2, in Russian (Cyrillic) as well as English. These references are also in Appendix 1.
Appendix_4.xls	Statistics for Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan on long-term mean monthly air temperatures (mean, standard deviation, minimum, maximum, utmin and utmax).
Appendix_4_Kaz.txt Appendix_4_Kyr.txt Appendix_4_Taj.txt Appendix_4_Tur.txt Appendix_4_Uzb.txt	Text files for the sheets in Appendix_4.xls.
Appendix_5.xls	Charts of anomalies in mean monthly maximum and minimum temperature, 1961-1990, in a Microsoft Excel file. The plus or minus 3 sigma line is also shown. There is a chart for each month, and each chart is based on all the stations in the data set. Note: There is not a tab-delimited file for this spreadsheet since there are primarily graphs contained in this file.
Appendix_6.1.pdf	Excel file with precipitation bias correction coefficients K1.1
Appendix_6.2.pdf	Excel file with precipitation bias correction coefficients K2.1

File Name	Description
Appendix_6.3.pdf	Excel file with precipitation bias correction coefficients K3.1
Appendix_7.xls	A summary of the location and data coverage of each station in the database such as: country number country code station name longitude and latitude altitude first and last years duration of years percentage of gaps and the percentage of monthly data completeness for both temperature and precipitation data
Appendix_7_1.txt	A tab delimited text file based on Appendix_7.xls.
1In these PDF files, the column named Type refers to the site descriptions given in Table 4 (Williams and Konovalov, 2005).	

Table 4. Classifications and Site Descriptions of Stations/Posts

Type	Site Description	Estimation
1a	Site bounded by solid, not ventilated fence. Clearing in dense forest. Site in mountain depression.	Most favorable for measurements.
1b	Site bounded by openwork vegetation or buildings	Vertical profile of wind is uniform.
2a	Half-protected site in city, village, or hilly place. Protection by buildings or vegetation from one, two, or three sides of horizon.	Very different influence on precipitation.
2b	Outlying district of inhabited locality. Half-protected site, bounded by dense obstacles such as separate buildings or groups of trees. Narrow gorge.	Most unsuccessful conditions for measurement of precipitation.
3	Open site inside of small houses or separate trees.	Saving logarithmic profile of wind.
4	Open sea shore. Mouth of large river. Island or summit of mountain.	Regional features of place are most important.

2.3 Measurement History

As a rule, meteorological observations were performed according to the Manual for Hydrometeorological Stations and Posts, Gidrometeoizdat, 1985 (translated from the Russian "Nastavlenie gidrometeorologicheskim stantsyyam i postam." Vypusk 3, chast’ 1). Stations are sites that have a more complete program and observations than at posts. Stations have a WMO number and posts may not. A standard meteorological site was 26 m by 26 m, and was located on relief typical of the area. It was more than 100 m distant from any bodies of water, and at a distance 20 times the height of any obstruction such as trees or a building.

Changes in methods and instruments for air temperature and precipitation measurements over the period of record that can lead to inhomogeniety in the data record are described by Williams and Konovalov (2005). See that paper, and the references therein, for more information.

2.4 Air Temperature

The thermometer type used at each station remained the same throughout the period of record. Minimum temperature was consistently measured with an alcohol thermometer, whereas hourly and maximum temperatures were each collected with separate mercury thermometers. When the air temperature approached the freezing point of mercury (-38.9°C), either an alcohol thermometer, or in some cases a minimum thermometer alcohol column, was used in place of the mercury thermometer. Information on when thermometers were replaced is available from the site histories at each station. Refer to Appendix 2.

The type of shelter or screen was not standardized until about 1930. No guidelines on the type of shelter existed prior to 1912. In 1912 thermometers were housed in Stevenson screens. From 1920 to 1930, depending on the station, the Stevenson screens were replaced with the current screens. The Russian translation for these screens is meteorological small house. In 1928, additional guidelines were issued regarding the exact dimensions of the shelters and their mounting heights. From 1930 on, most stations had their thermometers sheltered in roughly the same fashion. Potential inhomogeneities in the temperature data set were caused by differences in the times when hourly measurements were recorded. Prior to 1936, hourly measurements for computing daily mean temperature were taken at 0700, 1300, and 2100 Local Mean Time (LMT). Because of the lack of nighttime observations, daily mean temperature was probably overestimated at some stations, depending on location. Beginning in 1936, all thermometers (hourly, minimum, and maximum) were checked at 0100, 0700, 1300, and 1900 LMT at most stations. From 1966 to the present, all thermometers were checked at 3-hour intervals beginning at midnight Moscow winter Legal Time (MLT). (MLT is three hours later than Greenwich Mean Time). Table 5 provides a summary of this information.

Table 5. Summary of Air Temperature Recording Methods and Instrumentation. From (Williams and Konovalov 2005)

Year	Recording method/instrumentation implemented
1881	Measurements for computing daily mean temperature taken at 0700, 1300, and 2100 LMT; mercury thermometer used; because of alack of nighttime observations, daily mean temperatures were probably overstated.
1881	Daily minimum temperature thermometer checked at 0900 LMT; alcohol thermometer used.
1881	Daily maximum temperature thermometer checked at 0900 LMT; mercury thermometer used.
1881	No regulations regarding type of shelter surrounding thermometers.
1883	Daily minimum temperature thermometer checked at 0700 and 2100 LMT (lower value chosen); multiple measurements taken only to determine approximate time of occurrence of minimum.
1891	Daily maximum temperature thermometer checked at 1300 and 2100 LMT (higher value chosen); multiple measurements taken only to determine approximate time of occurrence of maximum.
1912	Official meteorological instructions recommended use of Stevenson screen to shelter thermometers; practice not implemented at all stations.
1920	Official meteorological instructions recommended use of current screen to shelter thermometers; practice implemented over next ten years.
1928	Official meteorological instructions specified exact size/height of screens.
1936	Measurements for computing daily mean temperature taken at 0100, 0700, 1300, and 1900 LMT (or at 0700, 1300, 1900, and 2100 LMT); bias in daily mean temperature dropped to ~0.2 C; daily maximum and minimum thermometers may or may not have been checked each hour.
1966	Measurements for all temperature variables collected at 3-hour intervals beginning at midnight MLT; bias in daily mean temperature eliminated.

2.5 Precipitation

By the late 1800s there were several hundred precipitation gauges in what was to become the former Soviet Union. In the late 1880s, it was determined that precipitation gauges underestimated the true amount of precipitation, especially precipitation in solid form. In 1891, the Nipher wind shield was introduced, and installation of the shield began.

The need for meteorological observations at airports was responsible for the move of some stations to airports in the 1930s. As new stations came online after the 1930s, many were located at airports. The unobstructed terrain characteristic of airports resulted in the precipitation instruments becoming more exposed to wind, potentially causing undercatch to increase. Even with the Nipher shield, undercatch of precipitation during the winter was still a problem.

Over the period 1946 to 1960, rain gauges with Nipher shields were replaced with the Tretyakov gauge. Tretyakov's precipitation gauges reduce errors caused by undercatch in windy conditions and the blowing of solid precipitation out of the gauges during snowstorms. Nipher gauges record considerably less solid precipitation than Tretyakov gauges. For this reason data acquired using Nipher shield gauges cannot be used with data acquired using Tretyakov gauges without adjustments to the data record. This correction, K1, is typically applied to monthly totals using coefficients derived for individual stations. The amount of snow recorded at each station increased from 5 - 40 percent after installation of the Tretyakov gauge. The date that this new gauge replaced the old is included in the station history metadata.

At the stations, a meteorologist measured the amount of precipitation one, two or four times every 24 hours. The duration and type of precipitation were observed continuously. The amount of precipitation was measured to a precision of 0.1 mm.

Changes in the time of observation occurred in 1936, 1966, and 1986. Prior to 1936, precipitation was measured once a day at 0700 LMT. From 1936-65, gauges were checked twice daily at 0700 and 1900 LMT. Beginning in 1966, gauges were checked four times a day, relative to MLT, which varied with the time zone of the region. In 1986, observations were again reduced to twice daily.

Concomitant with the increased daily sampling in 1966, a wetting correction was applied. The wetting correction accounts for systematic moisture losses due to moisture remaining on the sides and bottom of the precipitation gauge bucket when precipitation is decanted to the precipitation measurement glass. The amount of precipitation added to the total by the wetting correction depends on the number of observations per day. This is particularly important correction because a volumetric method of measurement for non-recording precipitation gauges is used. Unfortunately, the wetting correction for 1966 was insufficiently tested. In 1967 a better wetting correction based on controlled laboratory experiments was implemented. The correction adds 0.2 mm for liquid and mixed precipitation and 0.1 mm for solid precipitation for each individual measurement. The data published in reference books since 1967 use this improved wetting correction. The data for 1966 use the less reliable method for that year. Precipitation data published in reference books prior to 1966 do not have a wetting correction. Table 6 provides a summary of the history of changes in precipitation measurement practices.

Table 6. Summary of Precipitation Recording Methods and Instrumentation. From Williams and Konovalov (2005)

Year	Recording method/instrumentation implemented
1881	Rain gauge measurements taken at 0700 LMT; snowfall converted to a liquid total by melting snow in gauge; type of gauge and shielding not standardized.
1883	Official meteorological instructions recommended that cross-shaped zinc strips be inserted into the gauge to prevent snow from drifting; change probably not implemented at all stations.
1887	Official meteorological instructions recommended surrounding the gauge with the funnel-shaped Nipher's shield; change probably not implemented at all stations.
1892	Official meteorological instructions recommended erecting a fence around the gauge; change probably not implemented at all stations.
1902	Official meteorological instructions recommended erecting a double fence around the gauge; change probably not implemented at all stations.
1936	Rain gauge measurements taken at 0700 and 1900 LMT; daily total rainfall obtained by summing all measurements for the calendar day.
1946-1960	Old-style gauge (Nipher's shield) replaced with the Tretyakov-type gauge.
1966	Rain gauge measurements taken at 0300, 0900, 1500, and 2100 MLT in time zone 2; at 0300, 0600, 1500, and 1800 MLT in zones 3-5; at 0300 and 1500 MLT in zones 6-8; at midnight, 0300, 1200, and 1500 MLT in zones 9-11; and at 2100, 0300, 0900, and 1500 MLT in zone 12; wetting corrections ≤0.2 mm applied to each hourly measurement (Because four observations per day were collected at stations in time zones 2-5 and 9-12, four corrections were counted in the daily total; therefore, total daily corrections are higher for stations in these areas.)
1986	Rain gauge measurements at 0300 and 1500 MLT discontinued at all stations except those in time zone 2.

2.6 Station Metadata

The history of each station was acquired from publications such as History and Physical-Geographic Description of Stations and GaugePosts, published as part of the Reference Book on the Climate of the USSR (1966-1969). Appendix 3 gives specific metadata sources. Historical metadata includes station relocation date(s), the distance and direction of any such move(s), and date(s) for any changes in instrumentation such as the switch from Nipher shields to Tretyakov precipitation gauges.

Metadata consisting of information about the present environment of the station is included in Appendix 2, following the protocol in Peterson and Vose (1997). Station name, latitude, longitude, and elevation, following the coding from the current World Meteorological Organization station listings (WMO, 1996) is provided, in addition to:

- Population: Population metadata provides a valuable tool for climate analysis. Knowing whether a station is located in a rural area or near a large city allows data users to avoid the effects of urban heat islands, one of the criteria in the initial selection of the Global Observing System Surface Network (Peterson et al., 1997). Population values were obtained from 1990 census data. To conform with GHCN protocols, each site is classified as: (i) rural (not associated with a town larger than 10,000 people); (ii) small town (10-50,000 people); and (iii) urban (city with a population greater than 50,000).

- Airport locations: Many stations were moved to airport locations over the last several decades. If a station is located at an airport, this information along with the distance from its associated city or town is included.
- Topography: Each station is classified as flat, hilly, or mountainous. Mountain valley stations and the few ridge or mountaintop stations are distinguished.
- Vegetation: Vegetation classification in the surrounding area for each station included forested, agricultural, clear, open, marsh, ice, and desert.

2.7 Temporal and Spatial Coverage and Resolution

Data are available as monthly values (monthly means for air temperature and monthly sums for precipitation). The temporal coverage varies by station, but the longest range spans from 1879-2003.

The data fall within this bounding box:

Northernmost Latitude: 50° N
Southernmost Latitude: 35° N
Easternmost Longitude: 85° E
Westernmost Longitude: 55° E

Figure 1 shows the locations of the stations in this data set.

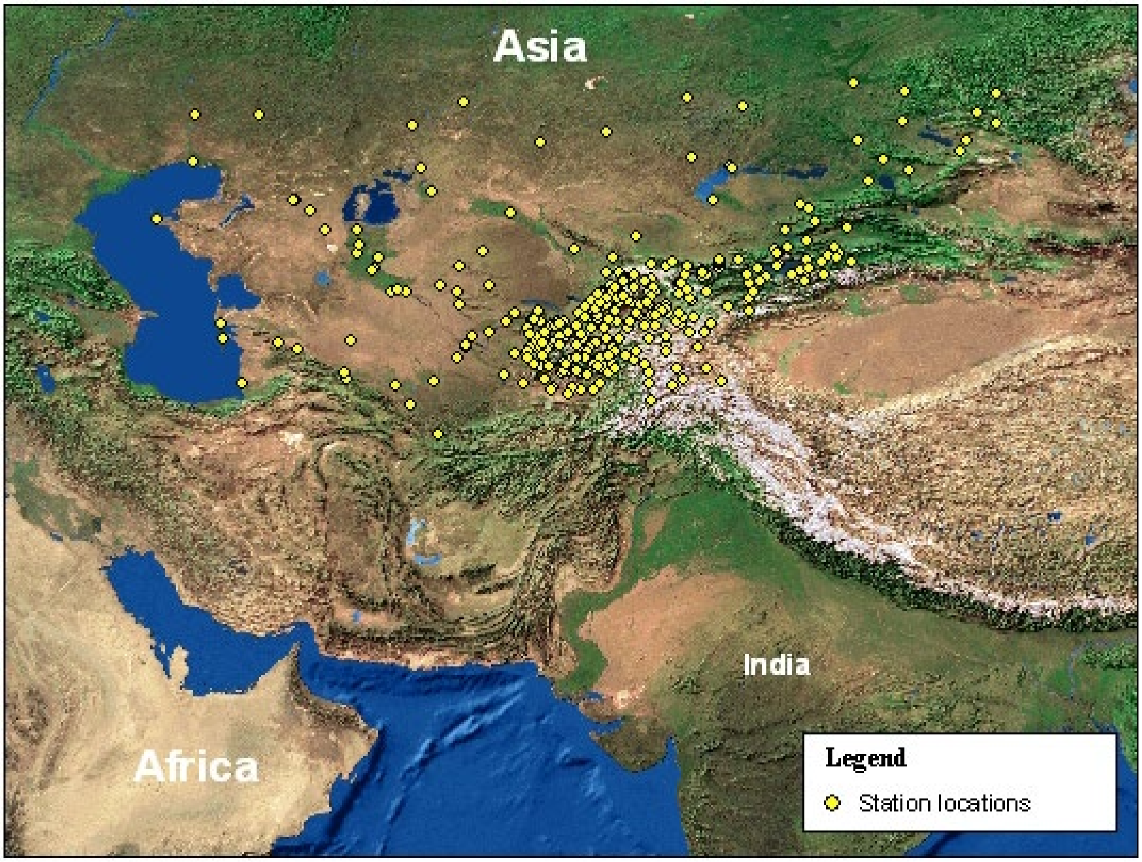

Figure 1. Spatial coverage map for Central Asia Precipitation and air temperature stations.

2.8 Parameter or Variable

2.8.1 Sample Data Record

In each Excel file, there is a Data and an Info worksheet. For example, in the average air temperature data files, these worksheets are named Taver and Taver_info. The worksheets with info in the name contain the metadata that correspond to the data. Refer to Tables 7 through 12 for an example of the beginning sections of the worksheets for the Taver_v1.xls, Precip_v1.xls, and Tmin_Tmax_v1.xls files.

Table 7. Taver Worksheet for the Average Air Temperature Data File

Country	Code	Name	Long	Lat	Alt	Years	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
211	36870	Almaty	76.93	43.23	847	1879	-9.8	-6.1	4.5	12.4	18.8	24.3	24.6	21.2	14.5	7.5	-1.5	-4.1
211	36870	Almaty	76.93	43.23	847	1880	-10.2	-10.9	2.0	7.0	15.8	21.3	22.3	20.1	14.0	10.6	1.5	-6.7
211	36870	Almaty	76.93	43.23	847	1881	-4.8	6.0	1.1	13.8	16.3	18.6	22.7	21.1	14.9	6.4	-2.0	-11.4
211	36870	Almaty	76.93	43.23	847	1882	-8.4	-8.9	-2.1	9.7	16.2	20.0	21.5	20.4	15.3	3.2	-1.1	-7.7
211	36870	Almaty	76.93	43.23	847	1883	-10.7	-14.0	2.8	7.9	16.3	21.3	21.8	21.6	14.3	6.7	-3.4	-6.0
211	36870	Almaty	76.93	43.23	847	1884	-6.6	-7.0	-2.6	12.0	15.5	20.0	22.6	20.0	14.9	6.6	-3.9	-5.1
211	36870	Almaty	76.93	43.23	847	1885	-10.5	-11.8	1.5	11.3	16.8	22.3	**22.4**	23.2	16.6	8.4	0.5	-4.2
211	36870	Almaty	76.93	43.23	847	1886	-999.0	-13.7	1.3	8.2	13.6	19.3	22.5	-999.0	-999.0	6.5	-2.7	-6.7
211	36870	Almaty	76.93	43.23	847	1887	-13.2	-7.5	0.7	13.5	12.5	19.1	22.2	19.6	14.9	10.8	2.3	-4.9
211	36870	Almaty	76.93	43.23	847	1888	-4.5	-5.7	5.6	11.3	17.0	20.7	25.0	21.8	15.4	10.9	0.5	-3.7
211	36870	Almaty	76.93	43.23	847	1889	-11.9	-5.3	2.2	11.3	12.2	21.7	22.0	21.1	14.2	5.1	-3.9	-11.5
211	36870	Almaty	76.93	43.23	847	1890	-7.9	-10.0	-6.7	9.0	15.0	19.5	23.0	19.8	14.4	8.2	1.9	-4.4
211	36870	Almaty	76.93	43.23	847	1891	-12.6	-12.2	-1.9	7.7	15.4	19.4	21.5	21.7	15.7	6.6	-1.1	-3.3
211	36870	Almaty	76.93	43.23	847	1892	-5.6	-5.7	-6.2	9.7	16.8	19.8	22.7	21.8	15.7	8.2	-3.5	-2.8
211	36870	Almaty	76.93	43.23	847	1893	-13.1	-9.9	3.8	12.2	15.8	22.8	23.8	22.6	18.5	7.8	0.8	-2.4
211	36870	Almaty	76.93	43.23	847	1894	-11.4	-3.3	3.2	7.3	14.9	22.2	24.0	20.8	16.0	5.7	-1.7	-9.9

Table 8. Info_Taver Worksheet for the Average Air Temperature Data File

211-KAZ	213-KYR	227-TAD	229-TUR	231-UZB				Dur	% of gaps	% of monthly data completeness										
Country	Code	Station	Long	Lat	Alt	F_year	L_year	years	year	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
211	36870	Almaty	76.93	43.23	847	1879	2003	125	0.0	97.6	99.2	99.2	98.4	99.2	99.2	99.2	99.2	98.4	99.2	99.2
211	35746	Aralskoe more	61.65	46.78	62	1905	2003	99	5.2	88.5	91.7	91.7	89.6	91.7	90.6	90.6	92.7	90.6	92.7	90.6
211	35796	Balhash	75.08	46.80	350	1932	2003	72	0.0	98.6	97.2	98.6	98.6	97.2	97.2	98.6	98.6	98.6	98.6	98.6
211	36879	Bolshoe Almaty Ozero	76.98	43.07	2516	1932	1995	64	0.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
211	38001	Fort Shevchenko	50.25	44.55	-25	1848	2003	156	1.9	98.1	98.1	98.1	98.1	98.1	98.1	98.1	98.1	98.1	98.1	97.4
211	35700	Gur'ev	51.85	47.02	-24	1881	2003	123	4.2	95.0	95.0	95.0	95.8	95.8	95.8	95.8	95.8	95.8	95.8	95.8
211	35542	Irgiz	61.27	48.62	114	1924	1995	72	0.0	98.6	98.6	97.2	98.6	97.2	97.2	98.6	98.6	98.6	97.2	97.2
211	36885	Issyk	77.47	43.37	1098	1938	1988	51	0.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
211	35406	Kalmykovo	51.87	49.05	1	1925	2003	79	1.3	90.8	94.7	90.8	90.8	88.2	92.1	92.1	94.7	94.7	92.1	89.5
211	35394	Karaganda	73.13	49.80	555	1936	2003	68	0.0	100.0	100.0	100.0	98.5	100.0	98.5	98.5	98.5	98.5	100.0	97.0
211	35663	Karsakpai	66.75	47.83	488	1926	1994	69	0.0	94.2	97.1	89.9	97.1	100.0	98.6	97.1	94.2	97.1	97.1	94.2
211	35849	Kazalinsk	62.12	45.77	68	1881	2003	123	0.0	95.0	96.7	96.7	95.8	95.8	95.0	96.7	96.7	96.7	95.0	95.8
211	36535	Kokpekty	82.37	48.75	512	1894	2003	110	12.1	85.0	83.2	87.9	86.9	86.9	86.0	84.1	85.0	85.0	86.0	84.1
211	38082	Kyzyl-orda	65.51	44.85	128	1891	2000	110	2.7	95.5	94.5	94.5	93.6	94.5	95.5	94.5	95.5	93.6	94.5	94.5
211	35576	Kzyl-zhar	69.65	48.30	361	1937	1995	59	0.0	93.2	96.6	94.9	98.3	96.6	98.3	94.9	96.6	98.3	98.3	100.0

Table 9. Precip Worksheet for the Precipitation Data File

Country	Code_WMO	Name	Long	Lat	Alt	Years	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
211	36870	Almaty	76.93	43.23	847	1879	25.0	24.0	18.0	87.0	32.0	23.0	36.0	37.0	29.0	41.0	31.0	43.0
211	36870	Almaty	76.93	43.23	847	1880	34.0	20.0	67.0	61.0	145.0	20.0	29.0	49.0	4.0	25.0	24.0	32.0
211	36870	Almaty	76.93	43.23	847	1881	14.0	51.0	50.0	102.0	62.0	89.0	40.0	28.0	73.0	47.0	110.0	21.0
211	36870	Almaty	76.93	43.23	847	1882	68.0	34.0	33.0	154.0	126.0	77.0	24.0	28.0	33.0	123.0	12.0	16.0
211	36870	Almaty	76.93	43.23	847	1883	30.0	4.0	37.0	58.0	37.0	89.0	29.0	4.0	20.0	30.0	12.0	37.0
211	36870	Almaty	76.93	43.23	847	1884	46.0	46.0	48.0	51.0	150.0	29.0	22.0	0.0	41.0	16.0	16.0	17.0
211	36870	Almaty	76.93	43.23	847	1885	10.0	7.0	34.0	57.0	5.0	100.0	30.0	6.0	19.0	45.0	79.0	48.0
211	36870	Almaty	76.93	43.23	847	1886	83.9	18.0	109.0	123.0	208.0	120.0	43.0	26.1	42.3	74.0	60.0	46.0
211	36870	Almaty	76.93	43.23	847	1887	25.0	20.0	48.0	102.0	82.0	58.0	68.0	53.0	20.0	26.0	52.0	47.0
211	36870	Almaty	76.93	43.23	847	1888	29.0	45.0	58.0	29.0	68.0	28.0	23.0	5.0	31.0	36.0	71.0	38.0
211	36870	Almaty	76.93	43.23	847	1889	15.0	17.0	81.0	139.0	89.0	75.0	57.0	8.0	0.0	62.0	48.0	52.0
211	36870	Almaty	76.93	43.23	847	1890	42.0	33.0	21.0	105.0	82.0	38.0	14.0	67.0	48.0	45.0	57.0	5.0

Table 10. Info_Precip Worksheet for the Precipitation Data File

211-KAZ	213-KYR	227-TAD	229-TUR	231-UZB				Dur	% of gaps				% of monthly data completeness								
Country	Code	Station	Long	Lat	Alt	F_year	L_year	years	year	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
211	36870	Almaty	76.93	43.23	847	1879	2002	124	0.8	96.8	97.6	97.6	98.4	98.4	99.2	99.2	99.2	97.6	97.6	97.6	98.4
211	35746	Aralskoe more	61.67	46.78	62	1906	2000	95	5.2	88.4	90.5	90.5	90.5	91.6	89.5	87.4	89.5	89.5	91.6	86.3	89.5
211	36622	Ayaguz,town	80.45	47.93	653	1896	1989	94	26.6	69.1	68.1	70.2	71.3	69.1	70.2	70.2	70.2	69.1	70.2	70.2	70.2
211	36736	Bakhty	82.72	46.66	221	1928	1999	72	0.0	95.8	97.2	90.3	91.7	94.4	94.4	93.1	98.6	91.7	97.2	93.1	93.1
211	35796	Balhash	75.00	46.90	398	1932	1999	68	0.0	95.6	97.1	97.1	97.1	95.6	95.6	95.6	97.1	97.1	97.1	98.5	94.1
211	36879	Big Almaty Lake	76.98	43.07	2516	1932	1991	60	0.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
211	38337	Blinkovo	70.15	42.23	1122	1906	1991	86	0.0	94.2	94.2	94.2	90.7	93.0	94.2	94.2	94.2	93.0	97.7	97.7	96.5
211	36665	Buran	85.20	48.00	409	1926	1994	69	2.3	95.7	95.7	97.1	94.2	97.1	97.1	97.1	94.2	95.7	97.1	98.6	97.1
211	38328	Chimkent	69.6	42.3	606	1919	1999	81	0.0	97.5	93.8	96.3	97.5	95.1	92.6	95.1	95.1	97.5	96.3	95.1	93.8
211	38001	Fort Shevchenko	50.25	44.55	-25	1891	2000	110	5.4	89.1	90.9	90.0	89.1	90.9	90.9	82.7	81.8	81.8	82.7	88.2	86.4
211	35700	Gur'ev	51.85	47.02	-24	1892	2000	109	31.1	68.8	67.0	67.9	67.9	67.9	67.0	67.0	67.0	67.9	67.0	67.9	67.9
211	36881	Ili	77.40	44.10	455	1933	1989	57	1.8	93.0	93.0	93.0	94.7	94.7	94.7	96.5	96.5	98.2	98.2	98.2	98.2
211	35542	Irgiz	61.27	48.62	117	1891	1999	109	11.4	83.5	85.3	81.7	82.6	83.5	85.3	81.7	86.2	84.4	82.6	83.5	81.7

Table 11. Tmax Worksheet for the Maximum Temperature Data File

Country	Code_WMO	mst	Long	Lat	Alt.m	years	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
211	36870	Almaty	76.93	43.23	847	1915	-999.0	-999.0	-999.0	-999.0	23.6	29.1	30.1	30.7	26.7	14.6	8.8	4.6
211	36870	Almaty	76.93	43.23	847	1916	0.1	-3.5	5.9	17.1	22.8	24.3	29.4	30.8	-999.0	14.5	0.6	0.4
211	36870	Almaty	76.93	43.23	847	1917	2.7	4.1	5.1	14.7	27.6	26.5	30.5	29.5	23.5	14.5	6.9	-2.3
211	36870	Almaty	76.93	43.23	847	1919	-999.0	1.9	5.8	18.4	22.8	-999.0	-999.0	29.8	20.7	17.9	8.8	0.9
211	36870	Almaty	76.93	43.23	847	1920	-999.0	-999.0	3.6	13.8	20.2	26.8	28.7	27.5	21.9	14.9	3.8	-6.4
211	36870	Almaty	76.93	43.23	847	1921	-0.3	0.9	4.1	14.9	17.5	23.5	28.0	26.9	22.7	15.5	7.1	2.2
211	36870	Almaty	76.93	43.23	847	1922	-1.6	0.4	9.0	17.4	21.8	26.1	28.6	27.6	24.3	17.5	8.6	2.8
211	36870	Almaty	76.93	43.23	847	1923	-4.3	0.6	6.7	15.8	21.0	26.4	27.6	32.7	20.8	15.2	7.9	-0.5
211	36870	Almaty	76.93	43.23	847	1924	0.2	0.8	4.8	16.9	18.2	24.4	30.6	30.3	23.5	15.5	7.2	1.1
211	36870	Almaty	76.93	43.23	847	1925	-0.8	-3.0	10.5	15.0	22.6	26.9	30.6	29.4	25.0	17.3	7.5	8.2
211	36870	Almaty	76.93	43.23	847	1926	-0.5	1.1	10.7	14.4	21.8	26.2	28.6	29.9	23.7	15.9	4.3	-0.4
211	36870	Almaty	76.93	43.23	847	1927	-2.4	-1.3	4.0	19.3	21.7	27.9	32.1	29.7	22.7	18.9	6.9	1.5
211	36870	Almaty	76.93	43.23	847	1928	-2.3	2.6	4.0	15.1	22.8	27.6	28.4	27.7	20.6	15.8	6.7	-4.2

Table 12. Info_Tmax Worksheet for the Maximum Temperature Data File

211-KAZ	213-KYR	227-TAD	229-TUI	231-UZB				Dur	% of gaps				% of monthly data completeness								
Country	Code	Station	Long	Lat	Alt	F_year	L_year	years	year	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
211	36870	Almaty	76.93	43.23	847	1915	2003	89	0.0	96.6	97.8	98.9	97.8	100.0	98.9	98.9	98.9	98.9	100.0	100.0	100.0
211	36879	Big Almaty Lake	76.98	43.07	2516	1932	1989	58	0.0	98.3	98.3	98.3	96.6	96.6	98.3	98.3	98.3	96.6	96.6	98.3	98.3
211	36891	Chilik	78.25	43.60	606	1933	1965	33	0.0	93.9	97.0	97.0	97.0	97.0	97.0	100.0	100.0	100.0	97.0	97.0	90.9
211	36885	Issyk	77.47	43.37	1098	1937	1965	29	0.0	89.7	86.2	93.1	96.6	96.6	96.6	100.0	100.0	100.0	100.0	96.6	96.6
211	36907	Kegen'	79.23	43.02	1845	1931	1989	59	20.3	74.6	76.3	76.3	76.3	78.0	78.0	76.3	78.0	74.6	76.3	78.0	78.0
211	36846	Kugaly	78.67	44.48	1410	1932	1989	58	1.7	89.7	91.4	91.4	91.4	91.4	93.1	94.8	98.3	94.8	93.1	94.8	94.8
211	36889	Mynzhilky	77.07	43.08	3017	1936	1989	54	0.0	98.1	98.1	98.1	96.3	96.3	98.1	98.1	96.3	96.3	96.3	96.3	100.0
211	36953	Narynkol	72.45	42.72	1806	1947	1989	43	0.0	97.7	97.7	97.7	95.3	97.7	97.7	97.7	97.7	95.3	97.7	100.0	100.0
211	36905	Podgornoe	79.47	43.33	1273	1934	1989	56	0.0	83.9	87.5	91.1	94.6	91.1	89.3	91.1	89.3	85.7	85.7	89.3	83.9
211	36952	Sarydzhas	79.60	42.92	1252	1940	1989	50	0.0	96.0	96.0	96.0	92.0	96.0	96.0	98.0	96.0	94.0	98.0	100.0	100.0

3 DATA ACQUISITION AND PROCESSING

The data were acquired from books, online sources and publicly available data sets, and keyed or reformatted as needed. Appendix 1 of the HTTPS download site provides a list of the original data sources. Sometimes daily data from the GTS or other sources of digital data were used. The WMO codes for countries and stations, the station coordinates, altitudes and the English transcription of station names are from references sources. See Appendix 3 on the HTTPS site: https://noaadata.apps.nsidc.org/NOAA/G02174/ for the metadata sources. In some instances, topographical maps with a scale of 1:100,000 were used to verify coordinates.

The data set authors:

- Digitized (keyed) the data when needed.

- Performed quality control on the measurements. This included standardizing air temperature data and removing values more than 3.5 sigma outside the normalized mean.
- Applied standard Soviet bias corrections to precipitation data.
- Evaluated air temperature for spatial homogeneity using multifactor equations based on latitude, longitude, and elevation. See Equation 5 in the BD_INFOR document on the HTTPS site (Williams and Konovalov, 2005).
- Prepared metadata that are in Appendix 3.
- Assessed the spatial homogeneity of the data.

NSIDC published the data with only making a few minor changes to the data. The changes include substituting -999.0 for no data (original data had no data marked by a blank space colored in yellow), creating text versions of the Excel files, and removing equations from the Excel files, for example, Excel files containing only values. In Appendix_2.xls, rows that had MOVE in the INFO column followed by -999.0 in the year column were removed. Appendix_5.xls was linked to the original source data and this link was severed at NSIDC.

3.1 Consistency Checks

Checks were made for file truncations, formatting errors, unreadable records, date variables having reasonable values, each station is chronologically sorted, the number of lines and record length are consistent with the time span, variable storage locations and variable types (integer,real, character) are consistent and in the correct columns.

3.2 Evaluation of Station Time Series

For source data that passed the consistency checks, the individual station time series were evaluated. Information on station location (latitude, longitude, elevation) was checked against the WMO directory and local atlases for each country. The authors have local knowledge that gave an additional measure of rigor to the checks.

Other checks included checking for three or more months with values identical to a precision of one decimal place. Gross discontinuities such as those caused by inappropriate concatenation of two stations into one, or a change in the units, for example, from Celsius to Kelvin, were evaluated using two tests. The Cumulative Sum (CUSUM) developed by (van Dobben de Bruyn 1968) tests for a change in the mean of a time series and has been used to determine the homogeneity of climate stations (Rhoades and Salinger, 1993). This test is sensitive to both long-term trends and to outliers. Because mountain stations often have anomalous but valid climate measurements, each decade is tested separately. Changes in variance were evaluated using a variant of CUSUM called SCUSUM.

3.3 Air Temperature

The primary GHCN approach to homogeneity adjustments for air temperature is to create a homogeneous reference series for each station (Peterson and Easterling, 1994). With the reference series created, inhomogeneities are detected by examining the difference series between a station and its reference series (Easterling and Peterson, 1995). Homogeneity adjustments are than made on time series with significant discontinuities (Students t-test) by using the difference in the means of the difference series' (station minus reference) over a 12-yr window. An evaluation of monthly air temperature data from Central Asia showed that homogeneity adjustments were not needed.

3.4 Evaluation of Individual Air Temperature Points

Higher elevation stations may have much greater year to year variability than stations at lower elevations. A measure of variance, the biweight standard deviation (Lanzante, 1996), was used to normalize the data for each station for each month prior to evaluating the record for outliers. Great care was taken to not throw away good data that happens to be extreme, because extreme events represent very important aspects of climate.

The statistics used in the normalization are in Appendix 4. Normalization and computation of standard scores (z-scores) was carried out according to:

$z_i = (X_i - X_{mean})/stdX$

Outliers with a z score outside 3.5 standard deviations were identified and discarded using:

$-3.5 < z > 3.5$

Normalization and checking for outliers was carried out on a monthly basis for each station. Statistical values for each station are given according to month in Appendix 4. The minimum and maximum z scores (Utmin and Utmax) are given as well.

See the Control of Tolerance for Data on Air Temperature section in the document BD_INFOR on the HTTPS site (Williams and Konovalov, 2005) for an explanation of the graphical analysis. The final versions of these graphs are shown in Appendix 5 on the HTTPS site.

3.5 Precipitation

Groisman et al. (1991) have developed a protocol for homogeneity adjustments of precipitation measurements from the former Soviet Union. Essentially, three correction coefficients (K1, K2, and

K3) have been developed for unbiased and homogeneous precipitation records. Parallel testing of the Nipher shielded gauges and the new Tretyakov gauges made it possible to develop monthly adjusted coefficients for the change in gauges (Hydrometeorological Service of the USSR, 1964). Experience has shown that these monthly adjustment coefficients (K1) have a standard error of estimate of about 10 percent (Shver 1965, 1976) when used to estimate the monthly precipitation that would have been measured by a Tretyakov gauge:

TRET = (K1)NIPHER

The Reference Book on the Climate of the USSR (1966-1969) contains the estimates of the proportion of the total mean monthly precipitation that adhered to the bucket walls (K3) for every station in the FSU. The values of K3 were estimated based on two observations per day. So, to correct the time series for the unapplied wetting correction prior to 1936 when measurements were made only once a day, the following equation is appropriate:

PMOIST = PORIG + [(0.5K3)PORIG)]

Where PMOIST is the adjusted data for the wetting corrections and PORIG is the originally measured precipitation amount. For the period 1936-1965, use:

PMOIST = PORIG + [(K3)PORIG)]

More problematic is adjustments that need to be made because of undersampling of precipitation caused by blowing rain and snow. Scale correction multiplication coefficients (K2) were developed as a function of climatological wind speed, temperature, and precipitation from the work of Bogdanova (1966). The use of K2 coefficients at individual stations or individual months is not recommended. However, all three coefficients can be used to develop mean area estimates of the true precipitation at ground level (Groisman et al., 1991).

4 REFERENCES AND RELATED PUBLICATIONS

Easterling, David R., Thomas C. Peterson, and Thomas R. Karl, 1996: On the development and use of homogenized climate data sets. Journal of Climate, 9: 1429-1434.

Groisman, P.Y. and E.Y. Rankova. 2001. Precipitation trends over the Russian permafrost-free zone: removing the artifacts of pre-processing. International Journal of Climate 21: 657-678.

Groisman, P.Y., V.V. Koknaeva, T.A. Belokrylova, and T.R. Karl. 1991. Overcoming biases of precipitation measurement: A history of the USSR experience. Bulletin of American Meteorological Society pp. 1725-1733.

Jones, P.D. and A. Moberg. 2003. Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. Journal of Climate 16 (2): 206-223.

Legates, D.R., and C.J. Willmott, Mean seasonal and spatial variability in gauge-corrected, global precipitation. 1990. International Journal of Climatology, 10 (1): 111-1270.

Nastavlenie gidrometeorologicheskim stantsyyam i postam. Vypusk 3, chast' 1. (Manual for hydrometeorological stations and posts. Iss.3, part 1). 1985. Gidrometeoizdat, Leningrad. 300 pp.

Peterson, T.C., and R.S. Vose, 1997: An overview of the Global Historical Climatology Network temperature database. Bulletin of the American Meteorological Society, 78 (12): 2837-2849. (PDF Version available at http://www.ncdc.noaa.gov/)

Peterson, T.C., R. Vose, R. Schmoyer, and V. Razuvaev, 1998: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. International Journal of Climatology, 18 (11): 1169-1179.

Williams, M. W. and V. G. Konovalov, 2005. GHCN: Regional Data Base on Climate of Central Asia v.1.0. Unpublished Microsoft Word Document. ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02174/BD_INFOR_E-1.doc

Vose, R. S., Richard L. Schmoyer, Peter M. Steurer, Thomas C. Peterson, Richard Heim, Thomas R. Karl, and J. Eischeid, 1992: The Global Historical Climatology Network: long-term monthly temperature, precipitation, sea level pressure, and station pressure data. ORNL/CDIAC-53, NDP-041. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

In addition, the following related document is available on NSIDC's Web site:

Document	Description	URL
NOAA at NSIDC's Precipitation Data Set Station Lists and Overlap Analysis Web page	Provides an analysis of the overlap in station coverage for various precipitation data sets.	https://nsidc.org/noaa/search/

4.1 Related NSIDC Data Collections

- Arctic Climatology Project. 2000. Environmental Working Group Arctic Meteorology and Climate Atlas. Edited by F. Fetterer and V. Radionov. Boulder, CO: National Snow and Ice Data Center. CD-ROM.
- National Snow and Ice Data Center. 2003. Meteorological Data from the Russian Arctic, 1961-2000. V. Radionov, compiler. Boulder, CO: National Snow and Ice Data Center. Digital media.

- National Snow and Ice Data Center, compiler. 2006. Monthly mean precipitation sums at Russian Arctic stations, 1966-1990. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.
- V.F. Radionov, Ye. I. Aleksandrov, P.N. Svyashchennikov, and F. Fetterer. 2004. Daily precipitation sums at coastal and island Russian Arctic stations, 1940-1990. Boulder, CO: National Snow and Ice Data Center. Digital media.

4.2 Other Related Data Collections

- USSR Monthly Precipitation For 622 Stations 1891-1999, NCDC DSI-3720. Available from the NOAA National Climatic Data Center.
- Global Daily Climatology Network, V1.0 and V2.0. Available from the NOAA National Climatic Data Center.
- Air temperature and daily precipitation data from 223 former USSR stations. Available from the All-Russian Research Institute of Hydrometeorological Information World Data Center (RIHMI-WDC).
- Global precipitation analyses for monitoring and research from the Global Precipitation Climatology Centre (GPCC).

5 CONTACTS AND ACKNOWLEDGMENTS

M. W. Williams
Institute of Arctic and Alpine Research
Boulder, USA

V. G. Konovalov
Institute of Arctic and Alpine Research
Boulder, USA

Acknowledgements:

The National Science Foundation funded the creation of this data set through Award No. ATM-0118384: Regional updating and expansion of the Global Historical Climate Network (GHCN) database: High-Mountain areas of Central Asia

Publication of the data set by NSIDC is supported by funding from NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) and the National Geophysical Data Center (NGDC).

6 DOCUMENT INFORMATION

6.1 Document Authors

F. Fetterer and L. Ballagh prepared this document based on correspondence with M. Williams. An unpublished document supplied by M. Williams and Konovalov (2005) provided a majority of the content along with a presentation by Williams, Konovalov, and Dyurgerov given in 2004, and other material provided by M. Williams.

6.2 Publication Date

February 2008

6.3 Date Last Updated

December 2020