# National Climatic Data Center

# DATA DOCUMENTATION

# FOR

# DATASET 9813
# Daily and Sub-daily Precipitation for the Former USSR

# July 15, 2005

Version 1.0

National Climatic Data Center
151 Patton Ave.
Asheville, NC  28801-5001  USA

**Table of Contents**


| | Topic | Page Number |
|---|---|---|
| 1. | Abstract | 1 |
| 2. | Element Names and Definitions: | 3 |
| 3. | Start Date | 7 |
| 4. | Stop Date | 7 |
| 5. | Coverage | 7 |
| 6. | How to order data | 7 |
| 7. | Archiving Data Center | 7 |
| 8. | Technical Contact | 8 |
| 9. | Known Uncorrected Problems | 8 |
| 10. | Quality Statement | 12 |
| 11. | Essential Companion Data Sets | 13 |
| 12. | References | 14 |
| 13. | Appendices | 16 |

## 1. Abstract

These files are a compilation of in situ daily and hourly meteorological observations for the former USSR initially obtained within the framework of several joint projects of All-Russian Research Institute for Hydrometeorological Information-World Data Center of the Federal Service for Hydrometeorology and Environmental Monitoring, Obninsk, Russian Federation, and NOAA National Climatic Data Center, Asheville, North Carolina. The data in this archive have been compiled from several sources:

1. Raw daily precipitation records obtained from RIHMI-WDC in two swaths:
   - Former USSR stations of International Exchange [223 of them; a subset of this data has been initially distributed by Carbon Dioxide Information Data Center (Razuvaev et al. 1993)] but now is updated to 2001 for the Russian Federation;
   - Former USSR (thereafter, Russia) synoptic stations (totally 2187 stations with the last data records ending in mid 1996)
2. Derivative daily precipitation time series composed from the above but homogenized for wetting bias changes and rain gauge changes using the technique developed by Groisman and Rankova (2001);
3. Sub-daily precipitation accumulations for the 2095 former USSR synoptic stations compiled from several sources but thereafter verified using national data holdings of the Russian Federation (up to 2000) but for other NIS countries mostly up to 1991 (cf., NCDC 2005; Data Set-9290c).

There are a total of 2188 stations (Fig. 1) in the daily precipitation archive, and a few records of raw daily precipitation go back as far back as far as October 1, 1874. In this version of the archive, the time series starts from 1891, the year when standard rain gauges with the Nipher wind shield were introduced *en masse* across the meteorological network of Russia. Initially, many records in the daily archive ended in early or mid 1990's. These records (as well as all missing values) were updated, when possible, with the values of accumulated for the day 6- and 12-hourly precipitation from the synoptic archive.

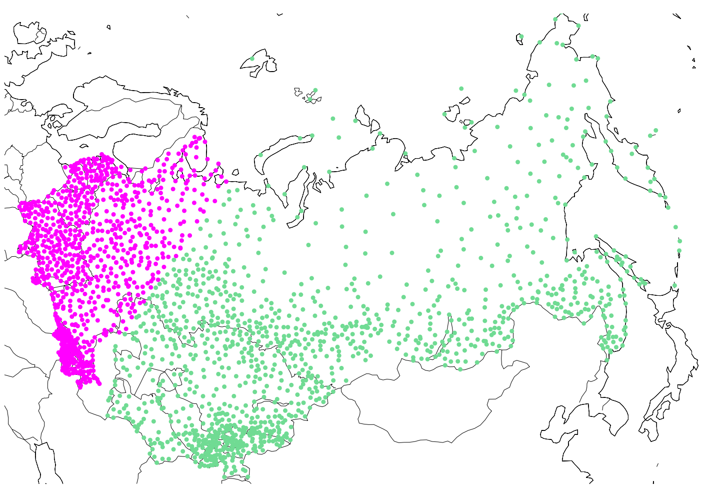

**Figure 1. Map of stations with data for the former USSR. File names in archive have 11-digit codes that distinguish Asian part of the former USSR (code 222000, green dots) and European part of the former USSR and Caucasus (code 638000, pink dots). Each station code has one of these two prefixes that follow thereafter with a standard 5-digit WMO code.**

Variables in the data set include (1) raw daily precipitation, (2) homogenized daily precipitation using the Groisman and Rankova, (2001) technique, (3) comprehensively corrected daily

precipitation using a technique developed by Bogdanova et al (2002a,b), and (4) raw sub-daily synoptic precipitation. Some temporal characteristics of archive data availability are as follows: 1381 stations have at least 83% of valid daily data during the 1961-1990 period. Among these stations, 749 have data (at least 50% of daily values) and 117 stations during the 1936-1945 and 1920-1929 decades respectively. Only 39 stations have daily data which include in the first decade of the 20th century, the 1920s and the 1961-1990 period. Mostly Russian data are present in the archive after 1991.

Homogenization and bias-correction are possible only when appropriate metadata and supplementary synoptic information are available. In particular, to homogenize the cold season precipitation time series for the instrument change in the early 1950s, we need at hand a coefficient K1 (Groisman et al. 1991; Groisman and Rankova 2001). There are 300 stations (listed in file "***ussr_stns_without_K1_but_with_data_prior_the_gage_change***" ), where we do not have this information and, therefore, had to replace the non-zero precipitation values in the homogenized archive with a special code (-99) that differs from missing code "-999" but indicates that on that day there was observed non-zero winter-type precipitation measured by the old rain gauge with the Nipher wind shield that we cannot homogenize with the following records of the Tretiakov rain gauge introduced in the 1950s. Also, methodology of bias correction suggested by Bogdanova et al., assumes the availability of wind information (at least wind climatology) and station exposure information or type of exposure. There are approximately 150 stations, where the exposure information was not available in the *Reference Book on Climate for the former USSR*. This could be due to two reasons: (1) the station is relatively new and at the time of the Reference Book preparation (mid-1970s) it was absent and (2) the station exposure dramatically changed during the station history and it was infeasible to assign a specific exposure class to the station. In both cases, there is no reason (in fact it could be detrimental) to implement sophisticated wind corrections developed by Bogdanova et al. to the data that can themselves introduce biases and/or inhomogeneities in the data. Stations where we do not have wind and/or exposure information (totally 433 stations) are listed in ***ussr_stations_without_wind_or_exposure*** file. Therefore, the column with bias-corrected precipitation in our archive contains more missing values than the column with raw daily data.

Each data value is characterized by four flags: data measurement flag, quality control flag, confidence level/status flag, and data source flag.

Archive contains two types of data files:

1. Major data files with daily and sub-daily precipitation **STN_ID.dat** (2188 files, 1 per station), and
2. Several supplementary files including:
   - **USSR_precip_stns.inv** (a file with essential station metadata such as station identifier, latitude, longitude, elevation, date of the first and the last records, and station name),
   - **ussr_stations_without_K1_but_with_data_prior_the_gage_change** (mentioned above and several more files with information about the archive metadata and data quality/availability)
   - **cases_with_flag_O** this file is described in **9.5.**
   - **ussr_stations_without_wind_or_exposure.** This file is described in **9.6.**
   - **percent_diff.final.** This file is described in **9.9.**

## 2. Element Names and Definitions

**2.1**. File **USSR_precip_stns.inv** has 2188 lines**.** Each line contains information about one station and can be read with the help of the following FORTRAN statement:

```
      REAL lat, lon, elev
      INTEGER time_beg, time_end
      CHARACTER*11 STN_ID
      CHARACTER *30 stn_name
      CHARACTER *3 repub
       READ(*,1) STN_id,lat,lon,elev,
     *   time_beg,time_end,stn_name,repub
1     format(a11,6x,f5.2,1x,f7.2, 1x,f4.0,2(1x,i8),1x,a30,3x,a3)
```

Where

**STN_ID** is the station ID composed from "222" and "638" – characters that identify former USSR in the coding system initially accepted in the Global Daily Climatology Network (DSI-9101), three reserved characters "000" and 5-character expanded WMO code for station accepted by the USSR Meteorological Service. In this WMO coding system the first two-digits are associated with specific latitude/longitude boxes (Figure 2).

**lat** is the station latitude (degrees, N). Range from 35.28 to 80.60 (degrees and hundredths).

**lon** is the station longitude (degrees, E). Range from 54.10 to 190.17 (degrees and hundredths).

**elev** is the station elevation in meters. Range from -28 to 4169. Missing code -999.

**time_beg** and **time_end** provide year, month, day, and hour with the first and the last data record respectively available for the station in the archive. Note that these dates may not indicate the starting/ending points of hourly observations at the station. Many digital records start in January 1936 and/or 1966 which were the beginning of the new observational practices at the network that participated in the data records but not the actual start of the observations at a given location.

**stn_name** provides the name of the station used in the national Catalogue (as of year 1990) and

**repub** is a three letter abbreviation used for stations of the republics of the former USSR (all these republics currently are independent states; Appendix 1).

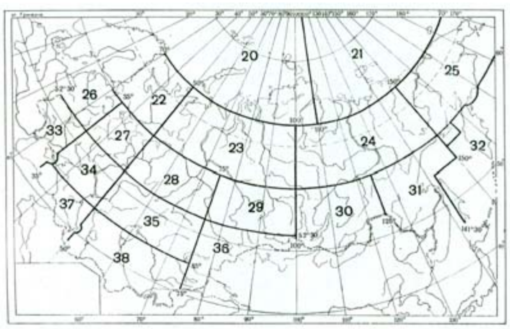


**Figure 2. First two-digits allocation for the WMO station IDs.**

---------------------------------------------------------------------------------------------------

**2. 2. Each file `STN_ID`.prcp contains the entire data available for STN_ID station.**

**These files were created using the following FORTRAN write statement:**

```
      INTEGER precip_raw, precip_adjusted, precip_corrected, precip_synoptic(8),
     *         precip_type, year, month, day ,hour
      CHARACTER*4 flag_precip_raw, flag_precip_adjusted, flag_precip_corrected,
     *         flag_precip_synoptic(8)
      WRITE (*,80) year, month, day, precip_raw, flag_precip_raw, precip_adjusted,
     *         flag_precip_adjusted, precip_corrected, flag_precip_corrected,
     *         num_synoptic, (hour(j),precip_synoptic(j), flag_precip_synoptic(j),
     *         precip_type(j), j=1,num_synoptic)
   80 FORMAT (i4.4, 2i2.2, 3(i4,a4),i2,1x,4(a4,i4,a4,i1,1x))
```

The same record is described below in detail:

```
Position    Variable
   1-  4  Year
   5-  6  Month
   7-  8  Day
       9  Blank
  10- 13  Daily raw data value (tenth of mm)
      14  Data measurement flag daily data value
      15  Quality control flag daily data value
      16  Confidence level or status daily data value
      17  Data source flag daily data value
  18- 21  Daily adjusted data value (Groisman & Rankova 2001 algorithm)
      22  Data measurement flag daily data value
      23  Quality control flag daily data value  (set to “H”)
      24  Confidence level or status daily data value
      25  Data source flag daily data value
  26- 29  Daily adjusted data value (Bogdanova et al 2002a,b algorithm)
      30  Data measurement flag daily data value
      31  Quality control flag daily data value (set to ‘H’)
      32  Confidence level or status daily data value
      33  Data source flag daily data value
  34- 35  Number of measurements per day (-1 missing; -9 unknown)
      36  Blank
  37- 40  Hour minutes of the first precipitation observation
  41- 44  First measurement, raw data value
      45  Data measurement flag
      46  Quality control flag
      47  Confidence level or status
      48  Data source flag
      49  Precipitation type (0–liquid,1–solid,2– mixed,9–can not be defined)
      50  Blank
  51- 54  Hour minutes of the second time of precipitation observation
  55- 58  Second measurement, raw data value
      59  Data measurement flag
      60  Quality control flag
      61  Confidence level or status
      62  Data source flag
      63  Precipitation type (0–liquid,1–solid,2– mixed,9–can not be defined)
      64  Blank
```

```
 65- 68  Hour minutes of the third time of precipitation observation
 69- 72  Third measurement, raw data value
     73  Data measurement flag
     74  Quality control flag
     75  Confidence level or status
     76  Data source flag
     77  Precipitation type (0–liquid,1–solid,2– mixed,9–can not be defined)
     78  Blank
 79- 82  Hour minutes of the fourth time of precipitation observation
 83- 86  Forth measurement, raw data value
     87  Data measurement flag
     88  Quality control flag
     89  Confidence level or status
     90  Data source flag
     91  Precipitation type (0–liquid,1–solid,2– mixed,9–can not be defined)
     92  Blank
```

-999 data value represents missing, -99 data value represents impossibility to adjust

The file flagging system resembles that for the entire Global Daily Climatology Network (cf., DSI 9101) but is more sophisticated.

**Data measurement flag**

blank = measured value
A = accumulated value
E = estimated value
F = estimated accumulation value
S = accumulation period, next non missing value is accumulated value
T = trace of precipitation
U = non-zero precipitation value (absolute value is unknown; in these rare cases the values itself can be coded -99);
X = value exceeds field width, add 1000.0 to value
Y = value exceeds field width, add 2000.0 to value
Z = value exceeds field width, add 5000.0 to value

**Quality control flag**

blank = unknown
N = normal QC by provider of data
1 = data were subjected to GDCN version 1.0 QC procedures (replaces “N” flag but does not replace the “H” flag)
H = instrumental homogeneity adjusted value

**Confidence level/status flag**

blank = unknown
A = accumulated value
B = value failed QC checks
C = scale corrected
D = derived value, when present at position 32 to characterize status daily data value for homogenized precipitation according to Bogdanova et al. routine, this flag value indicates that the precipitation value, P, was derived as a product of mean monthly precipitation intensity ($J_p$) and the monthly rainfall duration ($\tau_p$) for that month: $[P = (J_p)x(\tau_p)]$. These situations emerged when (a) we did not have temperature and/or wind information to apply the Bogdanova bias correction to observed precipitation values and/or (b) strong blizzard or blowing snow events with wind speeds

```
   above 10 m sec-1 made the Tretiakov rain gauge readings completely unreliable
   due to snow blow-in from the ground (cf., Bogdanova et al. 2002a,b).
 E = estimated value (e.g., when daily precipitation
                      was estimated as a sum of 6-hourly or 12-hourly totals)
 F = estimated accumulation value
 H = homologous value, rigorously tested
 I = interpolated value, not verified
 M = missing value
 0 = a sum of synoptic data differs strongly from raw data, could be scale
     error
 N = not tested but within observed climatological boundaries
 Q = questionable (actually wrong)
 R = record-breaking value
 S = Suspect value (outside climatological boundaries, not verified)
 T = tested value, manually checked but not perfectly homologous
 U = value suspect
THESE FLAGS are a result of GDCN version 1.0 QC and replace "blank" and "N"
flags when needed
X - exceeds known world extreme or impossible value
```

**Data source flag**

```
 blank = unknown
 G = GTS
 N = national meteorological service
 P = preliminary national data (e.g. operational update)
 ***************************************
 Note:  Combination of the precipitation data value  -99 and
 the second flag = H means that the homogeneity adjustment was impossible
and the valid non-zero precipitation value was replaced by  code '-99'.
```

**2.3. File "Cases_with_flag_O".**

**This file was created with FORTRAN write statement:**

```
      Integer, year, day
      Character*11 stnID
      Character*1    fl11, fl12, fl13, fl14
     Character*4 ar, a1
      i=year-1890
      write(19, 3) stnID, year, month ,day, ar, a1, fl11, fl12, fl13, fl14
 3      format(a11, i4, 2i2.2, 2(1x, a4), 4a1)
```

Each line has station ID (stnID), year, month, and day, where the flag "O" was inserted into the raw precipitation record, the value of this record (**ar**), replacement (**a1**, only in calculations of following homogenized precipitation values at that date) and group of flags associated with this record.

**2.4. File "ussr_stations_without_K1_but_with_data_prior_the_gage_change"** contains the list of 300 stations (StnID, Latitute, Longitude, Elevation (meters), and Station name) and was created by the following FORTRAN program

```
      Character*11 stnID
      Character*31 name
      Integer elev
      Real lat, lon
```

```
      write(19, 3) stnID, Lat, Lon, Elev, name
 3    format(a11, 2f8.2, i5,1x, a31)
```

**2.5. File "ussr_stations_without_wind_or_exposure"** contains the subset of file **USSR_precip_stns.inv** for which we do not have wind information and exposure. Record structure is the same as **2.1.** Contains list of 433 stations (of 2188).

**3. Start Date: 18910101**

**4. Stop Date: 20011231**

**5. Coverage:**

- a. Southernmost Latitude: 35.28 N
- b. Northernmost Latitude: 81.80 N
- c. Westernmost Longitude: 14.30 E
- d. Easternmost Longitude: 190.17 E

**6. How to Order Data:**

The data are distributed free of charge for use with a non-commercial purposes according to regulations outlined in WMO resolution 40. Contact NCDC's Climate Services about costs of copying of the off-line version of this dataset.

Phone 828-271-4800
Fax 828-271-4876
e-mail NCDC.Orders@noaa.gov

**7. Archiving Data Center:**

Name: National Climatic Data Center/NCDC
Address: Federal Building
151 Patton Ave.
Asheville, NC 28801-5001

Phone: 828-271-4800
Fax 828-271-4876
e-mail NCDC.Orders@noaa.gov

## 8. Technical Contact:

Name: Pavel Ya. Groisman
Address: National Climatic Data Center
Federal Building
151 Patton Ave.
Asheville, NC 28801-5001

Voice Telephone: 828-271-4347
Facsimile Telephone: 828-271-4328
Electronic Mail Address:Pasha.Groisman@noaa.gov

## 9. Known Uncorrected Problems

**9.1.** The data represent a transformation of original USSR observational data[1] at synoptic stations into a format accepted for GDCN. Some bias corrections (for example, the use of scale factors K1 to small amounts of daily precipitation totals) were inaccurate and generated random errors.
**9.2.** There are several changes in wind speed measurement techniques in Russia (cf., TD 9290c). Wind measurements are a part of the observing routine at 2095 stations (i.e., for most of this archive). Observations are always made at dedicated meteorological sites [26 m by 26 m; Nastavlenie…, 1958, 1985] at 10 m above the ground. However, most of these sites are not located at airports. The stations' surrounding has often gradually (and uncontrollably) changed due to urbanization and/or reforestation. At the same time, the wind adjustment that uses wind speed information at a given location to correct precipitation gauge measurements is unhampered. It should be clearly understood, however, that some of the changes revealed in such studies carry in a systematic land use component and the absence of wind correction (e.g., when we were not able to apply Bogdanova et al. routine) itself may introduce an "increasing" trend in the measured component of precipitation totals. For example, decreasing exposure may systematically reduce the wind speed above the gauge orifice and "help" the rain gauge to catch more precipitation then previously.
**9.3.** Several variables (sea level pressure, humidity, cloud types and present weather information) were unavailable in the original data set prior to 1936 and thus affected the accuracy of our corrections. Throughout the entire period of record, we used mean daily wind values during the day with precipitation to apply Bogdanova et al. bias-adjustments to avoid inhomogeneities due changes in number of wind measurements per day (3 measurements prior to 1936, 4 measurements during the 1936-1965 period, and 3-hourly measurements thereafter).
**9.4.** The WMO-A list of stations and existing GIS boundary files were used to assign the **repub** parameter to each of stations in **ussr_hly_stn.list.txt** file. While all efforts were made to avoid erroneous assignment, we cannot guaranty absolute accuracy of these assignments near the borders. For example, manual inspection revealed several cases, when some oases and enclaves in Central Asia had the **repub** parameter assigned incorrectly by GIS due to restricted precision (to hundredths of degrees) of the stations' coordinates. The revealed errors were fixed but a few others may still exist.

[1] or archived data with special restrictions to the data accuracy.

**9.5.** When we compared daily raw precipitation sums with sum of sub-daily precipitation at the same day, we sometimes observe small discrepancies of the order of 0.1 to 0.2 mm. These differences were attributed to rounding problems and were not corrected. In fact, we trusted the daily totals more than sub-daily totals (10.3). However, we also revealed a systematic error that could indicate a scale error in daily precipitation records (File "**cases_with_flag_O"**). In these cases (totally 9234 of them throughout he entire data set), the sum of sub-daily precipitation was approximately 10 times less than daily total and this total was always "rounded" to full mm. This means that these daily records could have, for example, a value of 1.0 mm but never 1.3 mm. We assumed that these records had scale error, flagged them in raw data set with flag "O" and used in the following homogenized records the corrected "raw" values reduced by factor of 10 as shown in File "**cases_with_flag_O"**.

**9.6.** We restricted the development of Bogdanova et al. corrections to measured precipitation at 433 stations (from 2188) when (a) we did not have at all wind measurements and/or (b) we did not have station exposure information at the station. Wind induced biases cannot be evaluated when no wind information is at hand. They will be also misleading when the site exposure is absent. There are approximately 150 stations for which we were not able to find exposure type information in Reference Books on Climate and the latest updates. Frequently, this means that the station exposure significantly changed with time. Exposure information is used in Bogdanova et al. routine twice: when the wind speed at the gauge orifice is derived from a standard wind measurements at 10 m above the ground and in estimation of the fetch length of blowing snow for corrections for "false" precipitation that blow-into the gauge during snow storms. When some of synoptic variables used in Bogdanova et al. routine were missing (relative humidity, number of rain events during the day, atmospheric pressure) we used their mean monthly values estimated for the period of record. When type of precipitation was not reported in present and/or past weather reports, temperature was used to partition daily precipitation into liquid, frozen, or mixed precipitation using the site or regional monthly climatologies. When wind was not available on a given day (but we had a plenty of wind measurements on other days with precipitation in a given month at a given station), we used climatological estimates of wind speed during precipitation events at this station to correct for wind-bias. However, if the station historical records show that for this month we can have more than one snow storm (on average), the wind climatology cannot be used for an estimate of missing daily wind values at the station in a given month. Frequencies of blowing snow events for the former USSR by region and cold season months are shown in Figure 3.

**9.7.** We encountered in some synoptic files "precipitation" reports that should not be there (e.g., additional precipitation measurements at the station that were supposed to make a smaller number of measurements per day). Most frequently, we encountered eight precipitation records per day since 1977 while only four measurements were ordered by Observers' Manual (Nastavlenie…1985). In all these cases, "additional" precipitation measurements had zero values. This situation emerges in the 3-hourly synoptic records, when the absence of precipitation measurement was coded by zero (or blank that was later converted into zero). For each station, we conducted statistical analyses of these false observations that during the entire period of record usually report only zeros, located and eliminated them altogether from the final data set.

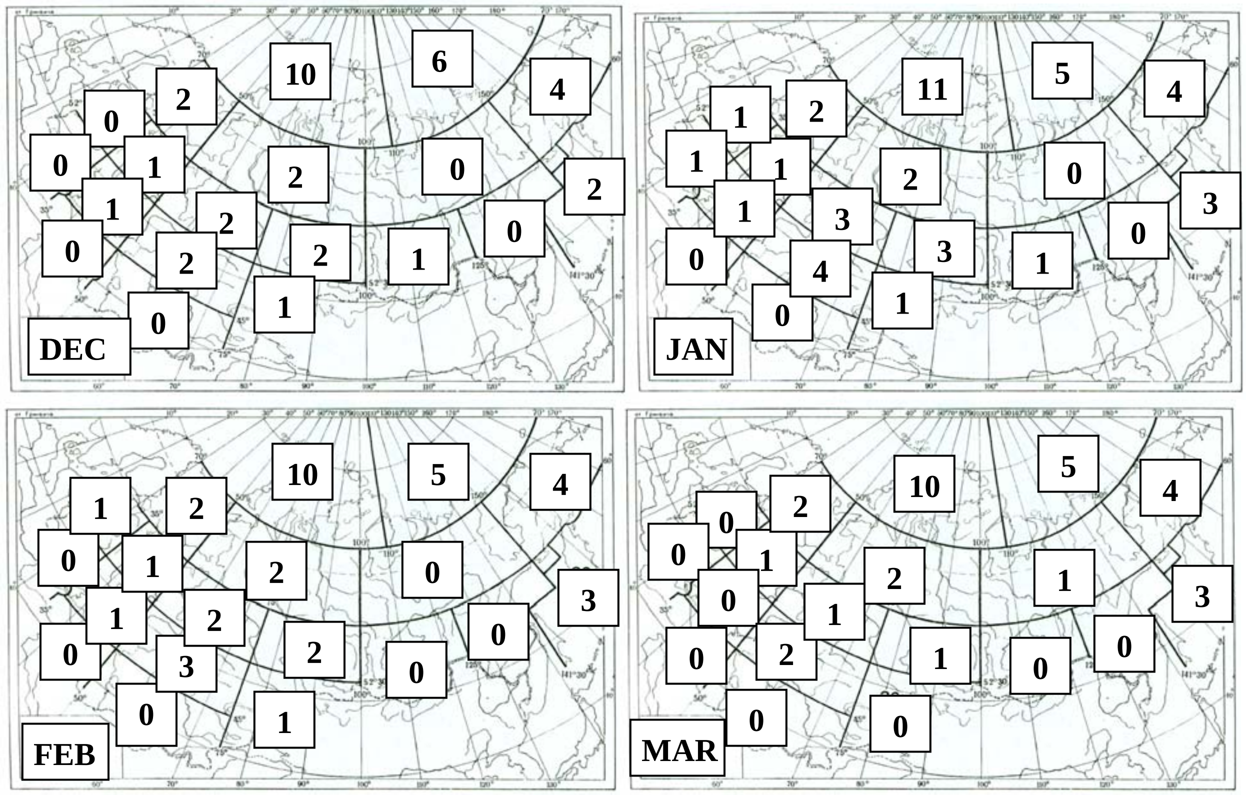


**Figure 3. Mean monthly number of days with blowing snow for period from December to March. Estimates have been made using the TD-9290c archive. In the Arctic the entire cold season from October through mid-May is prone to frequent snow storms.**

**9.8.** General rules of number of daily precipitation observations are as follows. Prior to 1936, there was one measurement per day. Usually, these were morning observations. At station Tomsk (ID 22200029430), the time of observation varied between morning, daytime, and evening hours. During the 1936-1965 period, there were two observations per day. Measurements were made at astronomical times at 7AM and 7PM. At station Auge (ID 63800026419) observations at 9PM were made instead of 7PM during part of this period. Since 1966, meteorological observations at the USSR stations became 3-hourly, while precipitation observations were made only four times a day. However, intervals between precipitation observations were not 6 hours in each time zone but varied. Two measurements that correspond to 03Z and 15Z (i.e., 06 Moscow standard time and 18 Moscow standard time) were always present in the record. In 1984-1985, four measurements per day were gradually eliminated everywhere, except the westernmost time zone (UTC + 3 that encompasses western Russia and, at that time, the Baltic States, Belarus, Ukraine, and Moldova). Still, in 1986, 108 stations beyond the westernmost time zone (62 of them in Kyrgyz Republic) continued for a while recording four precipitation measurements per day. Kyrgyzstan has preserved 4-time per day precipitation measurements up to the end of our data records for this country in 1991. We converted all times of observation into local standard time as of 1990 (see more on this issue in TD-9290c documentation). Stations with unstable times of observation during the post 1966 period are listed in Appendix 2.

**9.9.** We performed the data intercomparison between daily and sub-daily precipitation records at each station and day when it was possible. We noticed that the sum of sub-daily values may differ from a daily total by a few tenths of mm (usually by less than 0.3) without any particular pattern. After 1966, these differences could be associated with wetting corrections that could be inserted (or not inserted) in sub-daily precipitation totals transmitted immediately to the regional meteorological center while the daily totals were accumulated at the station and always contain these corrections. However, at some locations the differences were quite large indicating that for a while, the sub-daily daily precipitation could be reported from different locations. Therefore, a large-scale intercomparison has been conducted and some of its results are provided in the file **percent_diff.final**. In this file we list 649 stations (only slightly less than a third of the entire data set), where at least in one year we encountered more than 5% of valid daily precipitation values that differ from the sum of sub-daily totals by more than 0.3 mm. The file structure is as follows: Information line from file **USSR_precip_stns.inv (stnID,** and **line** with coordinates, station name, etc**.)** was appended with the number of years, **nyear**, in which we found a sufficient number (here more than 5%) discrepancies above the selected threshold (here 0.3 mm). Thereafter, the next line (or lines) provides these years (**ybad)** and percent of days with mishaps (**nbad**) per year. The station information was created with the following FORTRAN code:

```
        character*11 stnID
       character*80 line
        integer nyear, ybad(110),nbad(110)
........
         if(nyear.ne.0) then
             write(20,201) stnID,line,nyear
             write(20,205) (ybad(i),nbad(i),i=1,nyear)
         endif
201        format(a11,a80,i2)
205       format(10(i4,1x,i2,1x))
```

It should be noted that slight changes of thresholds (e.g., 0.3 mm to 0.5 mm) did not substantially affect the amount of information in the file **percent_diff.final**. At 500 stations (of those 649 that were found “guilty”) only one year (360 stations), two (98 stations), or 3 years (42 stations) have more than 5% discrepancies between daily and sub-daily precipitation records. At 58 stations, ten or more years have more than 5% discrepancies and among them, 17 stations have from 20 to 32 years with more than 5% discrepancies between daily and sub-daily precipitation records. In this last group of 58 stations, the years with differences are grouped within a particular period and percent of discrepancies can reach 50% per year clearly indicating the station’s mismatch during this period. Finding the solution to correct these mismatches is currently under thorough investigation at RIHMI.
**9.10.** Sub-daily precipitation data (retrieved from national and international synoptic data streams) did not have information about trace precipitation. This affected to some extent our ability to perform the full size corrections using Bogdanova et al. (2002a,b) routine.

## 10. <u>Quality Statement</u>

**10.1. Control using the source data set.** The data in this archive have been compiled from several sources (listed in **11**) that include information delivered via the Global Telecommunication System, international exchange, and random data acquisitions. This information was verified versus the original data holdings at the national hydrometeorological archive of the Russian Federation and appended when it was possible. The discrepancies when they became apparent, were always treated to the benefit of the national archive data source. This was the first level of the quality control.

**10.2. Statistical Control**. Small radii of correlation of daily precipitation totals prevented us from statistical control at this stage.

**10.3. Data sets intercomparison**. In addition to the GTS data, we had three major sources of daily precipitation from national data sources of The Russian Federation and Uzbekistan. First, we had 223 stations of International Exchange (Razuvaev et al. 1993) and its updates. Second, we had a national daily precipitation data set (2188 WMO stations) and sub-daily precipitation data (GTS data, updated and corrected by national Russian synoptic data set when possible). The data mostly coincide when compared to each other. However, some small (and large, cf., **9.5** and **9.9**) discrepancies were revealed. The following rules of preference were followed when we were forced to choose among the different data sets. Stations of International Exchange (as most rigorously verified at the site) have preference compared to the daily data set and daily totals have a preference compared to sums of sub-daily precipitation values (see however, 9.5).

**10.4. Significant changes in seasonal cycle and total annual precipitation** were found after Bogdanova et al. corrections were applied to the data. Figure 4 shows (a) increase in mean annual precipitation totals (in percent) throughout the former USSR during the 1954-1990 period compared to measured precipitation (prior to wetting corrections were introduced into the data) and (b) seasonal cycle in mean monthly precipitation over the western half of the Russian Arctic. It should be noted that in the Russian Arctic, the most prominent changes in corrected mean precipitation (according to the Bogdanova et al. algorithm) were not in the mid-winter months (December through March) when wind-induced and wetting biases were to some extent mitigated by the blow-into-the-gauge biases of opposite sign, but in the May-November period.

**10.5.** When wind speed during snow storms or blowing snow events is higher than 10 m $sec^{-1}$, the amount of "precipitation" caught by Tretiakov rain gauges could be very high due to the blow-in effect (Bogdanova and Golubev 1996; Bogdanova et al. 2002a,b). To alleviate this mishap, the bias-correction algorithm replaces the gauge-measured snowfall with a derived estimate equal to product of the mean precipitation intensity and monthly precipitation duration. The data status flag "D" is assigned to this derived value. While preserving the monthly averages of precipitation, this approach introduces random errors at the daily time scale that cannot be avoided.

**10.6.** We incorporated bias-corrections into the precipitation measurements made by rain gauge with Nipher wind shield (which was replaced by Tretiakov precipitation gauge in the early 1950s). These corrections were developed by Russian climatologists using detailed multi-decadal observations at Valdai experimental site in Central Russia (Golubev et al. 1999; Bogdanova et al. 2003; E.G. Bogdanova 2005, Personal Communication). Thus, while all efforts were made to secure homogeneity of the Nipher-shielded rain gauge readings with the follow-up readings of Tretiakov gauge, the initial incomparability of these two gauges in strong wind environment (e.g., at coastal Arctic and Pacific stations) may not be fully reconciled. User must be aware of this deficiency (cf., Groisman et al. 1991).

**Disclaimer:** While every effort has been made to ensure that these data are accurate and reliable within the limits of the current state of the art, NOAA and the Federal Service for Hydrometeorology and Environmental Monitoring cannot assume liability for any damages caused by any errors or omissions in the data, nor as a result of the failure of the data to function on a particular system. NOAA and the Federal Service for Hydrometeorology and Environmental Monitoring make no warranty, expressed or implied, nor does the fact of distribution constitute such a warranty.

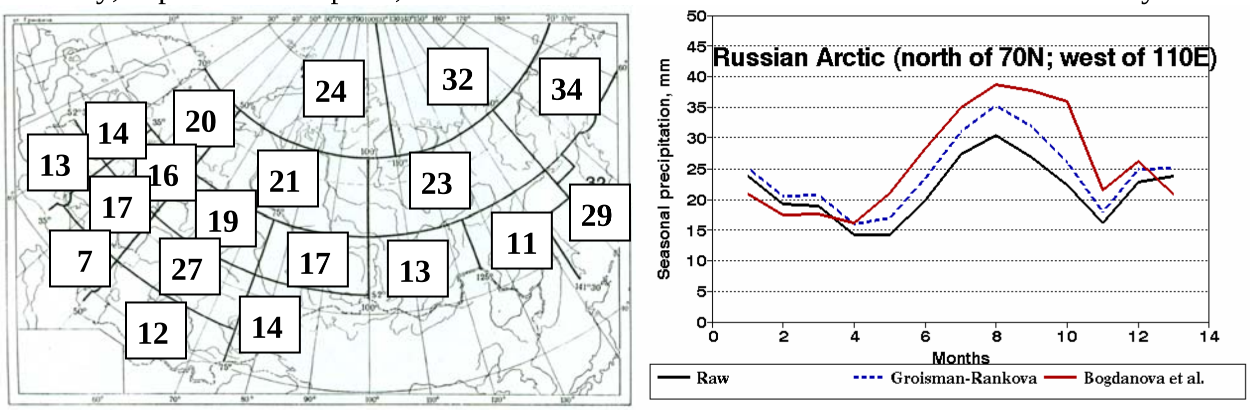


**Figure 4. Left. Increase (%) in mean annual precipitation totals throughout the former USSR during the 1961-1990 period compared to measured precipitation prior to when wetting corrections were introduced into these measurement data. Right. Mean monthly precipitation over the western half of the Russian Arctic (WMO Region 20 in Figure 2). The plot presents arithmetic averages of raw monthly precipitation measurements (without wetting corrections), and their homogenized values constructed using the Groisman and Rankova (2001) and Bogdanova et al. (2002b) routines during the 1954-1990 period. Data from nine synoptic stations in the region that had wind, precipitation, present weather, and temperature information sufficiently long during this period were used to construct the plot. Appendix 3 gives two more examples of non-trivial changes in the seasonal cycle of the bias-free precipitation.**

## 11. Essential Companion Datasets

Data description and additional information contained in these data sets can be helpful when working with this archive:

- TD-3505 Integrated Surface Hourly Database (ISH); Data Set 3505, DSI-3505 available at http://www4.ncdc.noaa.gov/ol/documentlibrary/datasets.html
- TD-9956 DATASAV3 Global Surface Hourly Data; Data Set 9956, DSI-9956 available at http://www4.ncdc.noaa.gov/ol/documentlibrary/datasets.html
- TD-1300 Synoptic Surface Observations – USSR Teletype. Data Set 1300, DSI-1300 available at http://www4.ncdc.noaa.gov/ol/documentlibrary/datasets.html
- Razuvaev, V.N., Apasova, E.G. and Martuganov, R.A. 1993. 'Daily Temperature and Precipitation Data for 223 U.S.S.R. Stations, ORNL/CDIAC-66, NDP-040, Carbon

Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee. 64 pp. + Appendices. Numerical Data Package, NDP-040 is available at http://cdiac.esd.ornl.gov/epubs/ndp/ndp040/ndp040.html

- Razuvaev, V.N., Apasova, E.G. and Martuganov, R.A. 1995. 'Six- and Three- Hourly Meteorological Observation from 223 USSR Stations', ORNL/CDIAC-66, NDP-048, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee. p. 68 + Appendices. Numerical Data Package, NDP-048 is available at http://cdiac.esd.ornl.gov/epubs/ndp/ndp048/ndp048.html
- TD-9101. Global Daily Climatology Network. Data Set 9101. Data set available at http://www4.ncdc.noaa.gov/ol/documentlibrary/datasets.html
- TD-9290c. Global Synoptic Climatology Network. C. The former USSR. Data Set 9290c. Data set available at http://www4.ncdc.noaa.gov/ol/documentlibrary/datasets.html

## 12. References

a. Adam, J.C and D. P Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias, *J. Geophys. Res*., 108 1-14

b. Berngardt, R.P. and M.V. Zavarina 1974: Climatological estimates of accuracy of measurements of strong winds. *Trans. Main Geophys. Observ*., 333, 121-129 (in Russian).

c. Bogdanova, E.G, Ilyin, B.M., and Dragomilova, I.V. 2002a: Application of a comprehensive bias correction model to precipitation measured at Russian North Pole drifting stations. *J. Hydrometeorol*., 3, 700-713

d. Bogdanova, E.G., V.S. Golubev, B.M. Ilyin, and I.V. Dragomilova, 2002b: A new model for bias correction of precipitation measurements, and its application to polar regions of Russia. *Russian Meteorol. and Hydrol.,* No.10, 68-94

e. Bogdanova, E.G., P.Ya. Groisman, and B. M. Ilyin 2003: The features of methods for correcting solid precipitation in the polar regions of Russia, Alaska and Canada. CD ROM with extended abstracts of presentations at the Final ACSYS Science Conference, St. Petersburg, Russia, Nov, 11-14, 2003

f. Braun,L., W. Grabs, and B. Rana, 1993: Application of conceptual precipitation - runoff model in the Langtang Khola basin, Nepal Himalayas. *IAHS Publ*. 218, 221-237.

g. Fedorova, E. G., 1966: Calculation of wind speed at the level of the precipitation gauge considering the degree of station protection. *Trans. Main Geophys. Observ*., 195, 63-68 (in Russian).

h. Gandin, L.S. 1988: Complex Quality Control of Meteorological Observations. *Mon. Wea. Rev*., 116, 1137-1156.

i. Golubev V.S., Bogdanova E.G., Simonenko A. Yu. 1997: The effect of blizzards on the correction method of precipitation data in Arctic region. In Proc. of the Workshop on the implementation of the Arctic precipitation data archive of the Global Precipitation Climatology Center. — Offenbach, Germany, July, 1996. WMO/TD No. 804, 1997, p.p.74-79.

j. Golubev, V.S. and Bogdanova, E.G., 1996: Accounting of blowing snow events in precipitation measurement in Russia. *Proc. of the ACSYS Solid Precipitation Climatology Project Workshop*, Reston, VA, USA, 12-15 September 1995, World Meteorol. Organ., WMO/TD No. 739.

k. Golubev, V.S., E.G.Bogdanova, A.Yu. Simonenko, 1997: The effect of blizzards on the precipitation gauge readings and proposals on the experimental check of the correction method of precipitation data in Arctic region. In Proc. Workshop on the implementation of the Arctic precipitation data archive of the Global Precipitation Climatology Center. Offenbach, Germany, July, 1996. WMO/TD No.804, 74-79.

l. Golubev, V.S., Konovalov, D.A., Simonenko, A.Yu., and Tovmach, Yu. V., 1999: Precipitation measurement correction and quality of corrected data. *Russian Meteorol. and Hydrol.*, No.1, 103-113

m. Goodison, B.E., Louie, P.Y.T., and Yang, D. 1998: WMO solid precipitation intercomparison. Final Report.

World Meteorol. Organ., Instruments and Observing Methods Rep. 67, WMO/TD 872, 87 pp. + Annexes.

n. Groisman, P.Ya., Koknaeva, V.V., Belokrylova, T.A., Karl, T.R. 1991: Overcoming biases of precipitation measurement: a history of the USSR experience. *Bull. Amer. Meteorol. Soc*., 72, 1725-1733

o. Groisman, P.Ya. and E. Ya. Rankova, 2001: Precipitation trends over the Russian permafrost-free zone: removing the artifacts of pre-processing. *Internat. J. Climatol.* 21, 657-678

p. Groisman, P.Ya., 2002: "Homogeneity Issues in the Global Daily Climatology Network: Precipitation in Cold Climate Regions", Extended Abstract, Proceedings of the WCRP Workshop on Determination of Solid Precipitation in Cold Climate Regions, Fairbanks, Alaska, June 9-14, 2002, 10 pp. CD ROM.

q. Groisman, P.Ya., and D.R. Legates, 1995: Documenting and detecting long-term precipitation trends: where we are and what should be done. *Climatic Change*, 31, 601-622

r. Groisman, P.Ya., E.G.Bogdanova, and B.M. Ilyin, 2004b: Adjusted precipitation for the United States and Russia. CD ROM of presentations at the International Workshop on Climate Data Homogenization (Toronto, Canada, 19-21 April, 2004), Meteorological Service of Canada

s. Hydrometeorological Service of the USSR, 1967-1970. *'Reference Book on the USSR Climate'*, Issues 1 through 34, part IV, Leningrad, Gidrometeoizdat (in Russian).

t. Korzun, et al., 1978: *World Water Balance and Water Resources of the Earth*: 1978, UNESCO Publ

u. Meshcherskaya A.V., I.F. Getman, M.M. Borisenko and E.I. Shevkunova, 2004: Wind speed monitoring on watersheds of the Volga and Ural rivers in the twentieth century. – *Russian Meteorol. Hydrol*., 2004, No. 3, 83-97

v. Nastavlenie meteorologicheskim stantsiyam i postam. Instructions to meteorological stations and posts. Issue. 3, Part I. Leningrad, Gidrometeoizdat, 1958, 223 pp. (in Russian).

w. Nastavlenie meteorologicheskim stantsiyam i postam. Instructions to meteorological stations and posts. Issue. 3, Part I. Leningrad, Gidrometeoizdat, 1985, 300 pp. (in Russian).

x. National Climatic Data Center, NCDC, 2002: Data Documentation For Data Set 9101, Global Daily Climatology Network, Version 1.0, 26 pp. [Available: at http://nndc.noaa.gov/?http://ols.nndc.noaa.gov].

y. National Climatic Data Center (NCDC), 2005: Data Documentation for dataset 9290c. Global Synoptic Climatology Network. C. The former USSR. 15 pp. [Available http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9813.pdf].

z. Razuvaev, V.N., Apasova, E.G. and Martuganov, R.A. 1993. 'Daily Temperature and Precipitation Data for 223 USSR Stations', ORNL/CDIAC-56, NDP-040, ESD Publ. No. 4194, Carbon Dioxide Information Data Center, Oak Ridge National Lab., Oak Ridge, TN. 47 pp. + Appendices

aa. Razuvaev, V.N., Apasova, E.G. and Martuganov, R.A. 1995. 'Six- and Three- Hourly Meteorological Observation from 223 USSR Stations', ORNL/CDIAC-66, NDP-048, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee. p. 68 + Appendices.

bb. Sevruk, B., 1982: Methods of correction for systematic error in point precipitation measurement for operational use. Oper. Hydrol. Rep., 21, Publ. 589, Geneva, Switzerland: World Meteorol. Organ. p. 91

cc. USAFETAC, 1986: DATSAV2 Surface USAFETAC Climatic Database. User Handbook No. 4, December 1986. USAF Environmental Technical Application Center, Operating Location A, Federal Building, Asheville, NC. 6 pp. + Appendices.

dd. Yang, D., and T. Ohata, 2001: A Bias-Corrected Siberian Regional Precipitation Climatology. *J. Hydrometeorol*., 2, No.2, 122-139

ee. Yang, D., 1999: An Improved Precipitation Climatology for the Arctic Ocean. *Geophys. Res. Lett*., 26, No.11, 1625-1628

## Appendix 1.

Abbreviations used for **repub** parameter (see also **9.9**).

| | | |
|---|---|---|
| **ARM** Armenia | **AZE** Azerbaijan | **BYE** Belarus |
| **EST** Estonia | **GEO** Georgia | **KAZ** Kazakhstan |
| **KYR** Kyrgyzstan | **LAT** Latvia | **LIT** Lithuania |
| **MOL** Moldova | **RUS** Russia | **TAJ** Tajikistan |
| **UKR** Ukraine | **UZB** Uzbekistan | **TUR** Turkmenistan |

## Appendix 2.

**Stations with variable times of precipitation observations since 1966.**

Period: January-June 1966:
**22200025538 22200035396 22200035401 22200035777 22200038049 22200038338 22200038383 22200038527 63800037328 63800037608 63800037626 63800037787 63800037884**

Period: July 1966 - December 1969: **63800027539 63800027729**

Period: January 1970- December 1976: **63800037947 63800027217**

Period: January 1981- December 1983:
**22200025551 63800034122 63800026518 63800027848 63800027928 63800027930 63800027935 63800027944 63800027947 63800027957 63800034047 63800034139 63800034146 63800034231 63800034238 63800034247 63800034336**

Period: January 1986- December 1991:
**63800026144 In 1991, station was upgraded to four measurements per day**
**63800027532 In January 31 1986, switch of observational times occurred to twice a day routine at 06 AM and 18 PM.**

**Appendix 3.**

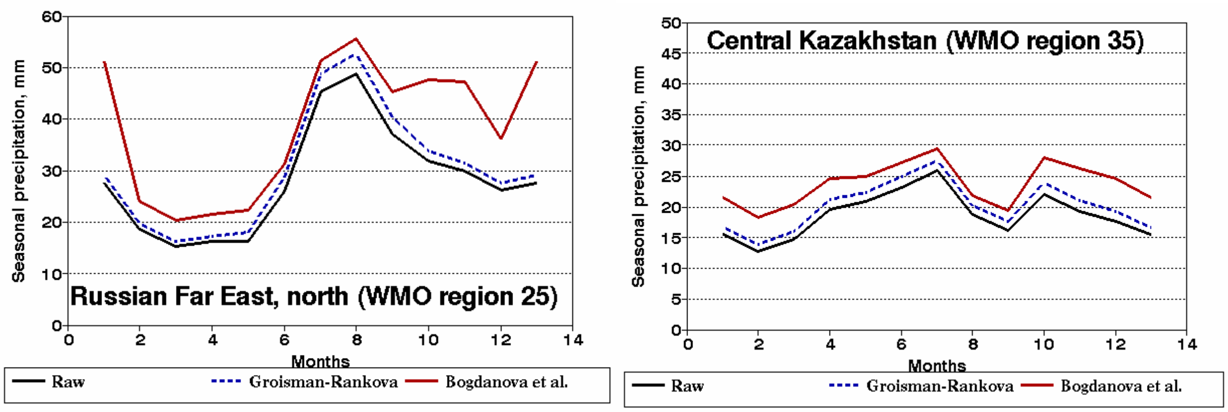


**Figure 5. Same as Figure 4, right panel, but mean monthly precipitation over the northeastern Russian Far East and Central Kazakhstan (in Figure 2, WMO regions 25 and 35 respectively). The plots present arithmetic averages of raw monthly precipitation measurements (without wetting corrections), and their homogenized values constructed using the Groisman and Rankova (2001) and Bogdanova et al. (2002b) routines during the 1954-1990 period. Data from 19 (region 25) and 100 (region 35) synoptic stations in these regions were used to construct the plots.**