

CADBV2 Technical notices for changes

Changes to data will impact data produced on the day that is listed in the change unless noted (ie. data will not be reprocessed for past dates unless noted in the notice).

9/5/24 - v2.55

- The station meta-data file used by the CADB suite has been updated. There are improvements in filling in information for some stations that were originally missing some meta-data, especially country information. The updated file is available at https://ftp.cpc.ncep.noaa.gov/cadb_v2/library/cpc_station_list_master.csv, a copy of https://ftp.cpc.ncep.noaa.gov/cadb_v2/library/cpc_station_list_master_08072024.csv

8/5/24 - v2.54 daily/weekly/monthly summaries

- Fixed bug introduced in v2.53 that incorrectly introduced the problem of not summing report precipitation values in some cases for the final precipitation estimate.
- Change to final estimate methodology: When there is at least one positive report precipitation value and one positive weather code estimate in the input hourly data within the bounded period, for U.S. stations, only the report precipitation values are summed for the final estimate. For all other international stations, weather code values supplement missing hourly report precipitation data if available and included in the summed total. The final precipitation flag assigned is 1.
- Daily and monthly data from May 24, 2024 to present is reprocessed and reflects these updates. Weekly and monthly data for this range of data will be reprocessed and updated online by Aug 9, 2024.

5/23/24 - v2.52 (temperature fix) and v2.51(precipitation fix)

- v2.52: Improve and do bug fix for method applying quality control check of zero air temperature hourly reports with neighboring time values.
 - Bug: Erroneous minimum temperature values of zero over Brazil (and other locations)
Fix: Fix the way the function of checking zero air temperature reports with neighboring (temporally) hourly values (v2.50) was being called. This function was not calculating the difference correctly for some reason when applied to multiple stations and therefore was not throwing out zeros as expected. Updated functionality so that all times are evaluated in the bounding time for comparison (iterating from nearest to furthest times until a valid non-zero air temp is encountered). Beforehand, only the closest hours were evaluated regardless if it was missing or zero. Added setting all hourly air temperatures to missing if all values are zero.
- v2.51: Redefine precipitation quality flags. Trust GTS reports of zero precipitation less, utilizing weather codes more for precipitation estimation

- Bug: Sometimes there are precipitation estimates that are too high over parts of Asia and Africa, also suspect incorrect reports of zero precipitation.
Fix: Increase checks of 6-hr reports and utilize weather code more in the calculation of the final estimate. Adjust precipitation flag definitions based on the fact that zero reported precipitation cannot be trusted.
Flag definitions:
- 0: (Remains the same) 24-hr reported value set as the daily estimate
- 1: There is at least one positive 6hr report precip AND one positive wxcode estimate. The positive wxcode estimate and 6hr report precip can occur at any time including the same time to be designated with this flag. Final est is the sum of 6hr report precip and wxcode estimates.
- 5: There is at least one or more positive 6hr report precip values but NO positive wxcode estimates. Final est is the sum of the 6hr report precip values.
- 7: All report precip values are zeros and/or missing but there was at least one positive wxcode est value. Final est = sum of ONLY wxcode estimates.
- 9: All report precip values were zeros and/or missing and there were no positive wxcode estimates. Final est = sum of 6hr report precip. This would either be zero if there is at least one zero or missing if all of the 6hr report precip were missing.
- -9: There was no data present in the bounded precip period to process. Final est is set to missing.

4/18/23 - v2.50 daily/weekly/monthly summaries

- Add 3 quality control (QC) features to improve values based on user reported suspicious values:
 - Bug: Spurious high precipitation bullseyes, especially over Africa. Fix: Set input GTS precipitation reports with values of 990 to missing, as this is likely an incorrect coding of a missing value.
 - Bug: Erroneous temperature related daily summary values of zero Deg C. Fix: Input GTS air temperatures sometimes have reports of zero Deg C, which is likely erroneous. Air temperature input reports with a value of zero are now thrown out if the absolute difference with the previous or next time exceeds 10 Deg C.
 - Bug: Maximum temperature daily summary value sometimes too low for Africa stations. Fix: For Africa stations increased the allowed threshold comparing GTS input reports of max and min temperatures to the standard deviation of the air temperatures. Over Africa, it was determined that reported values of min and max temperature may have a more significant difference compared to the range of air temperatures compared to other global stations.
 - Bug: Missing check for consecutive values of zero Deg F air temperatures before calculating wind chill. Fix: Include this check for consecutive zeros to improve wind chill estimates.

- All daily/weekly/monthly summary data has been backdated to include these bugfixes from January 1, 2021 to present. The backend database for the timeseries webapp will also have these fixes for the same dates.
- Some of the past archive data in the previous version may have been created with erroneous input, so the suite of data including anomalies have the fix with correct inputs in this version.

10/12/22 - v2.40 Updated normals/station meta-data/daily/weekly/monthly/wwcb summaries/timeseries webapp and minor format changes to some files.

- Updated station meta-data list, which is based on various station meta-data inputs.
- Updated normals to 1991-2020, associated with the updated station meta-data list.
- The daily, weekly/monthly/wwcb (weekly and monthly crop bulletin files) summaries are backrun using the updated station meta-data and normals (for files that apply) from January 1, 2021 to present.
- Normals and anomalies in the timeseries web application (<https://www.cpc.ncep.noaa.gov/products/timeseries/>) reflect the updated station list and normals from Jan 1, 2021 to present. All data prior to Jan 1, 2021 is removed and not available to ensure consistency in the data presented.
- Weekly and monthly city temperature and precipitation files (https://ftp.cpc.ncep.noaa.gov/cadb_v2/monthly/yyyy/mm/mctyprcp_yyyymm_v2.txt and https://ftp.cpc.ncep.noaa.gov/cadb_v2/weekly/yyyy/mm/wctyprcp_yyyymmdd_v2.txt) will have updated headers, previously "CLIMATE ANALYSIS CENTER-NMC-NWS-NOAA" to now "CLIMATE PREDICTION CENTER-NCEP-NWS-NOAA".
- Fix the ending date of data included in the monthly summaries, which should be the first day of the next month.
- Fix issue of the weekly summary regarding the incorrect month being indicated for the start of the season during December.

6/3/21 - v2.39 daily/weekly/monthly summaries

- Set final precipitation estimate values (column 'final_p') to missing that are greater than 400mm, except for stations that range from 41850 to 43599 which include Bangladesh, India, and Sri Lanka which can have extremely high precipitation values during the wet season. The associated precipitation quality flag (column 'p_flag') is then set to -9. The report precipitation value (column 'report_p') does not get changed. This QC-ing tries to eliminate some erroneous large precipitation values due to bad station reports.

12/18/20 - v2.38.2 daily/weekly/monthly summaries

- For selecting 6hr report values from duplicate times, change from selecting highest to the lowest value to avoid taking an erroneous bulls eye.
- Improve QC of 6hr and 24 hr precipitation reports, which compares summed 6hr reports to 24 hr report precipitation to decide whether to throw out any of the reports. Previously,

just threw out 24hr value if $\text{sum } 6\text{hr} - 24\text{hr} > \text{threshold}$. Now there is more QC-ing of both 6hr and 24 hr reports.

- Below is the methodology:

First clean the 6hr reports assigned to 3hr estimates:

1) If the summed 6hr report precip or 24hr value is zero, do not change any values

2) Else if $6\text{hrSum} - 24\text{hr} > \text{threshold}$, set the max 6hr report value and any duplicates of it to missing

(e.g. same value as the max 6hr value but at another time).

Only the max is removed (or dups of that value). Other values that are high may not end up being removed

3) Else if the $24\text{hr} - 6\text{hrSum} > \text{threshold}$, set the 24hr precip value to missing

4) Else do not change any values.

Next assess values for setting the precip flag (-9 means further processing needed, 0 means 24hr value should be the final summary value:

1) If the 24hr precip is negative or zero or missing, set flag to -9

2) Else if the $24\text{hr} - \text{sum } 6\text{hr precip} < \text{threshold}$ OR 6hr sum is zero (24hr would be positive, if it passed previous statement), set flag to 0

3) Else if $\text{sum } 6\text{hr precip} - 24\text{hr} > \text{threshold}$, set flag to -9

- Changed a coefficient for latitudinal adjustment for calculating estimated precipitation from weather codes (ie. reported precipitation type) from 50 to 0. It is suspected that this value in old documentation is likely incorrect and causing precipitation estimates from weather codes to be large in some cases.
- Replace report precip values that have a 999.x value with subtracting 999 from it, since it is likely that these are incorrectly reported with CLIMAT format.
- Fixed logic to not throw out 24 hr report precip values based on the difference from summed 6hr report precip if the sum is zero. This may have been previously throwing out good 24hr reports when all 6hr reports were zero - and it is possible that zero reported precip from a station is actually a missing.
- Fixed logic that could potentially lead to incorrect pairing of 6hr report precip values to precipitation period and start valid time.

All of November 2020 to present of CADB daily, weekly and monthly data is replaced in the current archived data. All new data as of valid December 18, 2020 moving forward will contain these latest updates.

9/25/20 - v2.34 daily/weekly/monthly summaries

- Implemented bug fix for QC-ing 24 hour precipitation. Changes in weekly and monthly data would be picked up by this change after this date.

8/7/2020 - v2.33 daily/weekly/monthly summaries

- Remove air temperatures (dewpoints) input report values where the value is zero and there is a duplicate report at that time/date with a value of -273.2 (less than -200).
- Set daily summary minimum and/or maximum temperature daily summary value to missing if there is only one input temperature report in the bounded temperature report data (evaluates reported air temperature, max and min temperature reports). Since the max and min temperatures have different bounding times it is possible for a station to have a valid min or max temperature and not the other.

7/17/20 - v2.31 daily/weekly/monthly summaries and station list

- Logic in daily code removes consecutive zeros in air temperature reports (in bounded temperature data to calculate tmin and tmax) if it exceeds the threshold (3 consecutive zeros).
- Station list - a few Africa stations were updated with missing elevation values, which were incorrectly set to 9999.

7/10/20 - Updated station list

- Some fixes and upgrades to the station list. No format change. Some stations have been adjusted, e.g. 72511 is correctly associated with KCXY. Some Nigeria station elevations have been fixed as well. This will result in some changes for station WMO ID number and station call, but not a significant number of stations. All data from June 1, 2020 to real-time present has been created using this updated station list.

6/25/20 - v2.30 daily/weekly/monthly summaries

- **Format changes: 1)** New column (2nd col) inserted after stn id for station call (5 character). This is to help users explicitly see what hourly input synoptic (WMO ID) data has been merged with metar (station call) data to produce the summary values. This column is now in the daily, weekly, and monthly summary csv files. This meta-data station reference may help users and the CADB team in the future of keeping track of changing WMO ID/station call assignments when combining daily data to produce 30 year normals. Also users then can search by either the WMO ID or station call in the datasets. NOTE: The column of stn_id REMAINS formatted as '99'+call letters for stations that only have call letters and no WMO ID. **2)** The weekly and monthly summary csv files no longer have a precip flag at the end of the files. These are no longer needed. : *weekly_summary_YYYY\$mm\$dd_v2.csv and monthly_summary_YYYY\$mm_v2.csv*
- Newly created and formatted CPC station reference list file is incorporated in the CADB summary suite of products. This was mainly done to improve the quality of matched WMO ID and station call letters, which represents two streams of input data (WMO ID-synoptic, call letters - metar). Previously some hourly input data may have been combining incorrect station data. This new station list has some manual adjustments on CPC's side, as well as utilizes the NCEI mshr_enhanced station reference file and aviation weather station list as input sources. NOTE: CADB normals have also been recreated with this new station list so that the stations should match properly between real-time data and normals. This new version of the file is in the library dir.
- Change in QC-ing tmin/tmax input hourly values: Lower the number of required valid hourly air temperatures to apply a standard deviation QC step to tmin and tmax hourly

values. Only 4 valid hourly air temps are required now, previously 6. This allows more stations to go through QC-ing since users were reporting stations that often have one or more bad tmin/tmax reports that were not getting QC-ed.

- Change in QC-ing report precip: If the reported 24hr precip value minus the summed 6hr report precip exceeds a determined threshold, then set the 24hr report precip value to missing. Thresholds are subjective and based on the number of 3-hour increments of estimated report precip (in the 6 hour report precip column). This adjustment attempts to fix some stations that have hourly reports that have likely incorrect very high precip totals.
- Change in QC-ing report precip: If the 6 hour report precip exceeds 572mm, set to missing, if 12 hour report precip exceeds 1144mm (global extreme for 12 hours), set value to missing.

6/25/20 - v1.7 normals

- The normals have been updated using the new station list which uses.
- Logic is adjusted to account for trace values in the NCEI monthly and daily normals that get merged into the normals dataset. Previously these values were thrown out, including the station if one month of data had a -7777 trace flag. Now these are handled, setting it to zero precipitation, allowing more stations to exist in the normals dataset now.

6/25/20 - station list dated 06112020 (cpc_station_list_master_06112020.csv).

Note: cpc_station_list_master.csv is a copy of this new version.

- Station list file name has changed to **cpc_station_list_master_\$\$mddYYYY**, where \$\$mddYYYY is the date that the station list was created. cpc_station_list_master.csv is a copy of it, which is done so that codes would not have to change file names with new versions.
- The station list is now a csv format instead of fixed width. Details of format changes are in the station library documentation in the FTP /docs folder.
- New station list created by Melissa Ou, containing new inputs from NCEI and aviation weather, as well as manual adjustments by CPC for station fixing. Daily, weekly, monthly data and normals are using the new station list. The normals use the newly, improved matching of NCEI normals based on CPC's WMO ID/station call convention and NCEI's GHCND ID convention. This correction should improve the data that is being combined to obtain a summary. Previously some incorrect synoptic versus metar stations were combined from different stations that were not matched correctly.
- The last column of precip flag is no longer existent.
- There is a new last column, with header name 'source'. This denotes the source of information for that row of meta-data. 'mixed' indicates that a blend of sources information was used.

3/5/2020 - v2.24

- Errors in the station list meta-data resulted in some stations combining incorrect station data together during processing. A CADB team member has manually evaluated and produced a new draft version of the station list data to fix this issue. Additionally, the software that produces the daily data has increased quality control to handle issues associated with duplicate station numbers being assigned to station call letters. The station list will continue to evolve over time with potentially improved input and include more stations reflected in the summary files (not all stations in the summary data files are included in the station list reference file).
- It was found that some Australia stations are reporting their 24-hour precipitation values earlier than the CADBv2 software anticipates, so the window of time to look for these reports are expanded. Now 3 hours prior to the end bound of the precipitation period to the end of the precipitation period, locally, is used instead of just one hour prior to the end of the period. The majority of stations with this issue is now no longer a problem.