

Question and Answer (Q&A) Guide

Homogenization Presentation

WMO/SERCOM/SC-CS/ET-DDS November 2025

1. Challenging Questions

Q1: Does homogenization methods consider the rapid climate change?

A: Absolute homogenization would detect rapid changes in climate series as inhomogeneities. That is why relative homogenization should be applied, where only changes alien to the climate signal will be detected and corrected.

Q2: Is Artificial Intelligence (AI) used for homogenization?

A: There are current efforts to explore the application of AI to homogenization, but this is a field that is in its early stages.

Q3: Can homogenization introduce bias?

A: When done incorrectly, yes. That's why strict methodology, transparency, and documentation are crucial. Using multiple statistical methods and reference series reduces risk.

Q4: How do you justify adjusting historical data to skeptical audiences?

A: Explain that adjustments remove **artifacts**, not real climate signals. Show examples where unadjusted data would misrepresent trends, and demonstrate validation against reference series.

Q5: How do you handle stations with very short or inconsistent records?

A: Short or inconsistent records can still be useful in regional analyses. We document uncertainty clearly and use them cautiously in homogenization, often cross-checking with neighboring stations.

Q6: What is the WMO Data Policy for homogenized data?

A: There is no specific rule for homogenized data in the WMO data policy. Homogenized (climate) data fall under climate data as such and the WMO data policy defines historic time series data needed to describe the climate as core data (basically focusing on ECVs), which shall be exchanged on a free and unrestricted basis (with a tentative maximum delay of one year).

Q7: Do we need the same homogenized process (method) for a regional approach? That means for a RCC to ask each RCC's Members the same homogenized method to be able to gather all datasets together?

A: It would add the plus of consistent methodology, but it is not strictly necessary.

2. General / Conceptual Questions

Q1: What exactly is homogenization?

A: Homogenization is the process of adjusting climate data to remove biases caused by changes in instruments, station locations, or observation practices. It ensures that observed trends reflect **true climate variations**, not artificial shifts.

Q2: Why is it necessary?

A: Without homogenization, we risk misinterpreting climate trends. It's essential for research, climate services, and decision-making in sectors like energy, health, and agriculture.

Q3: How often should homogenization be done?

A: Homogenization is an ongoing process. Datasets should be updated regularly as new observations come in or as old data are rescued or re-evaluated. While the first homogenization of a dataset may be time consuming, due to the election of the method, preparation of the input data and the execution of exploratory runs until optimal results are obtained, subsequent homogenizations of the same dataset will be much easier, especially if the process is implemented in an automatic script.

3. Technical / Implementation Questions

Q1: What timescales are typically homogenized?

A: Monthly, daily, and sub-daily timescales. The choice depends on data availability and the intended applications.

Q2: How do you handle gaps in historical data?

A: Through **data rescue**, interpolation, statistical methods and proxy series (reanalysis, etc) used as references in homogenization procedures. Metadata and documentation are crucial to ensure transparency.

Q3: Are there standard software tools for homogenization?

A: Yes. For example, ACMANT, Climatol, MASH, RHtests and other WMO-recommended packages. The choice depends on national capacity and dataset specifics.

Q4: How do you assess the quality of homogenized datasets?

A: By assigning **quality levels**, validating against reference series, and documenting all adjustments made. Continuous feedback from users is also important.

4. Organizational / Strategic Questions

Q1: What are the key prerequisites before starting homogenization?

A: Strong governance, compliance with **WMO standards**, knowledge of datasets (metadata, QC, rescued data), and a clear vision aligned with national and global priorities.

Q2: How much resources are typically needed?

A: It depends on the dataset size and scope. You need trained staff, technical infrastructure (hardware/software), and sustainable funding. Even small teams can start and gradually scale up.

Q3: How do you ensure the process is sustainable?

A: Through a structured plan, clear responsibilities, staff training, secure data storage, and regular updates. Collaboration at regional and global levels also helps.

5. User / Impact Questions

Q1: How does homogenization benefit sectors like agriculture or energy?

A: Homogenized data provide reliable long-term climate records, enabling **better planning** — for example, irrigation schedules, crop selection, energy demand forecasting, or flood risk assessment.

Q2: How do you communicate with users?

A: Through webinars, workshops, forums, or direct feedback channels. This ensures datasets meet users' needs and builds trust.

6. Questions put in the chat of the consultations

6.1. Homogenization Tools and Methods

Q1: Except MASH and MISH, what homogenization tools are available?

A: Currently, the most commonly used **publicly available** homogenization software packages are, in alphabetical order: ACMANT, Climatol, HOMER, MASH, PHA and Rhtests.

Q2: Is there an option for the Climatol algorithm to detect changes in standard deviation as a criterion?

A: Yes, using the parameter test='cuct', changes in the mean and/or in the standard deviation will be detected and corrected. However, it can only be used with complete datasets, since changes in the distance to the reference stations would change the standard deviation of the spatial anomalies, hence producing false positives.

Q3: Have any NMHSs been using 2–3 automatic tools simultaneously to identify breakpoints, considering lack of good metadata?

A: There are several papers published by University and NMHSs researchers making use of 2 or 3 homogenization tools. They acknowledge that all results improve the series, although it is often difficult to tell which method performs best.

Q4: Different homogenization methods may come from different expert teams, producing different dataset versions. How to share these and help users select the appropriate dataset?

A: As in real cases it is impossible to know which results, either produced by different methods or with different parameterizations of the same method, are the best, an interesting possibility is to calculate climatic normals, trends, etc., with all of them. The differences will offer an idea of the uncertainty of the results, and the user can choose those adapting better to the desired product or use a mean or a median of all of them.

6.2. Metadata and Data Validation

Q1: How can we validate breakpoints detected through homogenization techniques if historical metadata is scarce?

A: A high value of the inhomogeneity test is a strong evidence of a real breakpoint. When the test gives a value near the detection threshold, the acceptance of the breakpoint depends on the tolerance to have false positives.

Q2: I have heard about homogenization by Members. How about homogenization across and between Members?

A: There have been regional homogenization projects spanning several countries. Also there have been national projects including cross-border data provided by neighboring countries.

Q3: How can we verify that homogenized data are correct? What data and methods can be used to validate homogenization results themselves—ensuring accuracy and avoiding over-correction?

A: Validation is difficult without knowing the truth. However, a better temporal and spatial consistency of the homogenized series and of any statistical parameters calculated from them (normals, trends, etc.) gives confidence on the validity of the results.

6.3. Training, Tutorials and Capacity Development

Q1: Any suggestions on tutorials/training materials or workshops for beginners?

A: Homogenization software packages include manuals on their use. In <https://climatol.eu> there are also some videos on the use of Climatol. Some workshops are organized from time to time by different institutions, depending of fund availability.

Q2a: Can WMO organize a training course focused on climate observation data homogenization? We have limited human resources and technical expertise.

Q2b: What is the status of the training on the homogenization component?

A: WMO takes note of the training requests. Members are strongly encouraged to add a work package on homogenisation training to national or international projects. WMO will be happy to help drafting an agenda for homogenisation training and identify international experts/trainers.

6.4. Technical Challenges / Software

Q1: What impact will the ban on mercury thermometers have on homogenization processes? Digital thermometers have additional issues with technology, maintenance, and costs.

A: Yes, those changes will introduce biases in the series, similar to those derived from changes to automatic measurements. A couple of years of parallel observations is always recommended, although not always possible.

Q2: Is it recommended to use the CDT (Climate Data Tools) software to complete and homogenize the data?

A: CDT seems to include a homogenization module, but it is unclear in the documentation which algorithm is implemented.

6.5. Information Management / Policy

Q1: What possibilities are there for making this information publicly available and providing training to observers, who are the first filter of information?

A: Training to observers is provided at NMHSs level following the guidelines published by WMO. Homogenization is more a task to be undertaken at the climatology departments.