

WORLD METEOROLOGICAL ORGANIZATION

Expert Team on Data Requirements
for Climate Services (ET DRC),
Topic meeting on data gaps,
15 February 2023, 13:00 – 15:00 Geneva time

Original: English

Participants:

Mr Axel Andersson	Mr Ali Eddenjal	Mr Atsushi Goto
Mr Rachid Sebbari	Ms Xiaolan Wang	Mr David Inglis Berry
Mr Uli Looser	Mr Jose Guijarro	Ms Elke Roskamp
Ms Christina Lief	Mr Denis Stuber	Ms Caterina Tassone
Ms Ge Peng	Mr William Wright	Ms Belen Martin Miguez
Mr Robert Dunn	Ms Manola Brunet	Mr Peer Hechler
Mr Markus Ziese	Ms Tina Leiding	

Notes (actions and key agreements highlighted in **bold**):

Peer opened the meeting, explained the brainstorming nature of the meeting and reflected on Members' requests for assistance on how to deal with data gaps (e.g. related to the recent calculation of new Climatological Standard Normals). Some guidance is provided in the WMO Guide to Climatological Practices, in the WMO Guidelines on Climate Normals and, implicitly, in WMO Guidelines on Generating a Defined Set of National Climate Monitoring Products. More specific guidance may be developed as part of guidance material on basic climate statistics (tbc).

Atsushi introduced a joint GCF/WMO initiative on data gaps with the aim of drafting a scientific paper on modern methods on infilling gaps.

Atsushi and Peer noted the benefits of collaboration between the two initiatives.

The following points were made during the brainstorming on data gaps (all reflected in the attached mind map drafted, presented and updated by Denis):

Upcoming observational networks including GBON and GRSN are supposed to help minimising data gaps in the future

Canada recently used a spatial modelling approach to fill in data gaps. A subsequent homogenisation exercise helped adjusting the estimated data accordingly.

A literature review will be needed – there are many papers on data gaps and mitigation measures available, wjw: Will has sought advice from Australian climate change experts as to the best way to address this. A literature review by a PhD student may be one way, but perhaps a more detailed academic collaboration will be required.

It is important to flag estimated data. Data centres may consider removing estimated data from basic time-series data sets to avoid misinterpretation (estimated data should not appear as original observations). wjw: Agree, however suitably-flagged data may well still have value. I would not recommend restricting access to interpolated data, but such data need to be correctly flagged as "estimated/interpolated", and access restrictions may apply. A suitable flagging process should be factored into future guidance on data quality control flagging.

Some tools for spatial analyses require rather high correlation coefficients, which may rule out the few neighbouring stations available.

Most homogenisation packages include data infilling functionality. These should be spelt out.

Acknowledge limits of data estimation.

Validate implications of data gaps.

It is important to discover the reason for the data gap – different reasons may require different infilling methods. Need also to distinguish between missing data infilling, and disaggregation, where the total amount (in terms of precipitation) is known.

Hydrologists use various standard techniques for data infilling including interpolation techniques and modelling techniques (modelling run-off data and estimate run-off at certain points). Need to reach out to hydrologists for more details.

Marine climatologists are facing the specific issue of moving observational sites. Standard techniques are used for data infilling including statistics, reanalyses and ensemble methods.

Data rescue is an important source for discovering and digitising missing data. This includes data held in various possible archives, such as NCEI and other global data centres, but also data held by other Member NMHSs (which is particularly relevant for countries with a colonial past).

Acknowledge and document potential methodological differences when using different data flows for infilling, e.g. CLIMAT and SYNOP.

Consider implications of data estimation on derived data, statistics and products.

It is important to document infilling processes and estimation methods. An international flagging system may help recognising estimated and disaggregated data. Yes.

Will GCF provide some funds for research related to data infilling, including data rescue and data fusion concepts? Previous experience suggests this would need to be done via a bid for GCF funding that includes a Climate Rationale.

International data centres may have copies of missing national data. NCEI responds to related requests.

Conclusion:

While this brainstorming just marks a starting point for addressing Members' requests for guidance and assistance in closing data gaps, there is value in documenting broadly the different steps suggested to fill in data (identify the gaps, explore the reasons for the gaps, lessons learnt to minimize data gaps in the future, search for data copies in international data archives, search for data on outdated media to be rescued, use of auxiliary data including remote sensing and reanalysis data, statistical data estimation including spatial and temporal interpolation etc., documentation of data gaps and infilling methods and flagging of infilled data).

Peer will organise a follow-up topic meeting in April 2023. Hands-on assistance from ET members to start drafting guidance material based on the above brainstorming elements and the attached mind map will be appreciated.