Researchers develop a new ocean data quality control system

January 30 2023, by Li Yuan

Credit: CC0 Public Domain

Over the past century, more than 16 million ocean temperature profiles had been collected by various instruments. However, each instrumentation provides data of different accuracy, different quality, and different completion of the metadata.

Before using this raw data to do the scientific research, the quality control (QC) process is compulsory to ensure data accuracy and availability. In early years, the QC was usually performed manually by

the experts. However, the manual QC of large datasets is not feasible due to the manpower and time cost.

Researchers from the Institute of Atmospheric Physics (IAP) of the Chinese Academy of Sciences (CAS) and their collaborators provide a new climatological range-based automatic quality control system for ocean temperature in-situ profiles. The system is called CAS Ocean Data Center—Quality Control system, or CODC-QC for short, and it includes 14 distinct quality checks to identify outliers.

The study was published in *Deep Sea Research Part I: Oceanographic Research Papers*.

"We developed this new QC system to provide a quality-homogenous database, with reduced human-workload and time cost on manual QC," said Tan Zhetao from IAP, first author of the study.

In CODC-QC, the 0.5% and 99.5% quantiles are used as thresholds to define local climatological ranges. These thresholds are time-varying, which aims at erroneously excluding real data during "extreme events." The above strategies are used in local climatological range check for both temperature and vertical temperature gradient, in which the anisotropic feature of water properties is accounted for, and the topography barriers adjustment of water mass are made.

In addition, the performance of CODC-QC system was evaluated using two expert/manual QC-ed benchmark datasets. This evaluation demonstrated the effectiveness of the proposed scheme in removing spurious data and minimizing the percentage of mistakenly flagged good data.

The CODC-QC was also applied to global World Ocean Database (WOD18) including 16, 804, 361 temperature profiles from 1940 to

2021. Based on the statistics of temperature outliers, 7.97% of measurements were rejected, in which XBT data took the highest rejection rate (15.44%) whereas the Argo profiling float took the lowest rejection rate (2.39%). "We suggest a dependency of the quality of temperature observations on the instrumentation type," said Viktor Gouretski, researcher from IAP and co-author of the study.

The researchers also applied the CODC-QC system to the study of monitoring global ocean warming. "We found that the application of the CODC-QC system leads to a 15% difference for linear trend of the global 0–2000m ocean heat content changes within 1991–2021, compared with the application of WOD-QC (NOAA/NCEI), implying a non-negligible source of error in ocean heat content estimate," said Prof. Cheng Lijing from IAP, corresponding author of the study.

The quality-controlled (by CODC-QC) and bias-corrected ocean in-situ profile data of CAS-Ocean Data Center, Global Ocean Science Database (CODC-GOSD) are now freely accessible at http://www.ocean.iap.ac.cn/ and http://english.casodc.com/.

More information: Zhetao Tan et al, A new automatic quality control system for ocean profile observations and impact on ocean warming estimate, *Deep Sea Research Part I: Oceanographic Research Papers* (2022). DOI: 10.1016/j.dsr.2022.103961

Provided by Chinese Academy of Sciences