

EGU24-1788, updated on 20 May 2024

<https://doi.org/10.5194/egusphere-egu24-1788>

EGU General Assembly 2024

© Author(s) 2024. This work is distributed under the Creative Commons Attribution 4.0 License.



A Global Database of quality-controlled Crowd Weather Station Data

Benjamin Bechtel¹, **Jonas Kittner**¹, Daniel Fenner², and Matthias Demuzere³

¹Bochum Urban Climate Lab, Ruhr-Uni Bochum, Bochum, Germany (climate@rub.de)

²Chair of Environmental Meteorology, Faculty of Environment and Natural Resources, University of Freiburg

³B-Kode, Ghent, Belgium

Privately-owned weather stations, Crowd Weather Stations (CWS), offer high spatial and temporal density in many urban regions across the globe, and therefore have been used in a variety of urban climate studies, mostly focusing on single cities. One challenge in crowdsourcing CWS data lies in the fact that the link between measured atmospheric data and (historic-) metadata is often lost due to the limited metadata available from popular CWS networks. This poses challenges in retrieving and analyzing data, as, e.g., past changes in CWS location remain undetected, introducing incorrect data, thus reducing data integrity.

We developed an end-to-end workflow for consistently collecting and checking CWS (meta-)data in 257 areas worldwide, covering over 500 urban regions since 2019. The workflow automatically adds newly set-up CWS to the database, as well as consistently handling changes in CWS location. Until now, the database includes over 310,000 CWS with 7 Billion hourly observations of air temperature and relative humidity (mean, maximum, minimum). Over 65,000 changes in CWS location have been detected since 2019. This highlights the importance of continuous metadata updates for this dynamic data source, further enabling the use of the measurements for different applications. Within the database, CWS are linked to additional metadata, including a global digital elevation model, a global Local Climate Zones map, and the Global Human Settlement Layer Urban Center Database.

The database was developed using open data and open-source software, combining PostgreSQL, PostGIS, and Timescale, which allows us to manage billions of measurements efficiently. All air-temperature measurements are consistently and continuously quality controlled using the state-of-the-art open R-Package CrowdQC+. The result is a dataset of consistently-processed metadata and measurements with potential for global-scale (intra-)urban climate studies and in-depth city analyses.[MD1] [DF2]