Contribution ID: **45** Type: **Poster Presentation**

Automated quality control for data from ASTI automatic weather stations

Meteorological observations are vital for scientific research, for weather monitoring and weather-related decision making. Such data sources act as support for several applications relating to agriculture, aviation, early warning, weather prediction, and disaster mitigation. In the last four years, the Advanced Science and Technology Institute (ASTI) has deployed over 1600 ground stations (automatic weather stations and rain gauges) over the Philippines in response to the government's effort in disaster risks management. The growing demand for these data requires that the quality control be performed to ensure their reliability and integrity. Notably, this tightly associates with the validation of high-quality data against \emph{ground truth} coming from manned synoptic stations of PAGASA. The ultimate goal is not only to provide guidance to the end-users but also aid operations in identifying and potentially predicting problems that may require immediate attention.

This work presents the development of quality control procedures applied onto data obtained from ASTI ground stations. These procedures include verification of geolocation information, timestamp, range, step, persistence and internal consistency checks. In the proposed system, data are not altered; rather, assigned a flag according to the result of each check. In the initial tests performed on data (air pressure, air temperature, relative humidity, rainfall and wind speed) collected between 2012 and 2015 from 15 stations, 7.97% of the records were detected to have invalid timestamps, 0.11% were flagged in the range check, 0.26% in the step check, 6.35% in the persistence check and 0.87% in the internal consistency check.

By comparing AWS temperature and rainfall data with that of PAGASA synoptic stations, we found that eliminating dubious observations improved the agreement of AWS measurements with ground truth. Validation statistics for temperature and rainfall were (\emph{r}=0.94; bias=0.37°C; RMSE=1.01°C) and (r_s=0.80; bias=-0.79mm/d; RMSE= 12.04mm/d), respectively, as opposed to the baseline statistics using raw data given by (\emph{r}=0.84; bias=0.47°C; RMSE=1.77°C) for temperature and (r_s=0.80; bias=-1.17mm/d; RMSE= 14.36mm/d) for rainfall.

While the system was able to identify various errors in the measurements, continuous improvements to the algorithm are necessary to achieve more robust quality control. A few examples are spatial consistency checks, application of site-specific quality control parameters, long-term time series analytics, and data interpolation. Nevertheless, this development of quality control system provides considerable value to ASTI ground station data that is expected to promote further end-user utilization for research and various applications.

Primary author: Mr COMBINIDO, Jay Samuel (Advanced Science and Technology Institute)

Co-author: Mr MENDOZA, John Robert (Advanced Science and Technology Institute)

Presenter: Mr MENDOZA, John Robert (Advanced Science and Technology Institute)

Track Classification: Earth & Environmental Sciences & Biodiversity Applications