A quantitative approach to optimize the quality control system for surface data at MeteoSwiss

Valentin Knechtl, Deborah van Geijtenbeek and Christian Sigg

Swiss Federal Office of Meteorology and Climatology MeteoSwiss
Operation Center 1, P.O. Box 257, CH-8058 Zurich-Airport
valentin.knechtl@meteoswiss.ch
Tel: +41 58 460 93 90

Abstract
MeteoSwiss operates a network of over 130 automated meteorological surface stations which measure every 10 minutes an extensive set of meteorological parameters. An automated quality control system is in place to detect erroneous measurements. It is based on a set of logical rules and corresponding limits. Measurements are flagged if a rule exceeds its defined limits. The flagged values are subsequently inspected by an expert. Currently, the quality control system at MeteoSwiss consists of more than 200 different types of rules, and the specific limits for each rule instance can depend on the measurement parameter, location and acquisition time of the measurement. Rules’ limits are set and customized based on operative experience.

One of the main issues in the day-to-day business is the high false positive (FP) to true positive (TP) ratio, i.e. the high number of flagged but correct measurements relative to the number of flagged and erroneous measurements. The negative consequence is that the time spent dealing with FPs takes away time from the inspection and correction of TPs.

In order to optimize the overall performance of the quality control system, we estimated the FP / TP ratio for all rules based on 18 months of measurement data. In a second step, we tested the poorest performing rules with a set of limits including the default limit to find out whether the FP / TP ratio can be improved or not. Based on these results, we selected a new limit if we observed a performance improvement or we deactivated a rule if we obtained no performance improvement.

With this approach a significant reduction of the overall FP / TP ratio was achieved. Firstly, it could be shown that some of the rules with a high FP / TP ratio cannot be improved. Consequently, these rules are deactivated. Secondly, the FP / TP ratio can be substantially improved for some rules by choosing a new limit. In summary, the amount of false positives is reduced considerably whereas the amount of true positives is kept at virtually the same level.