Geosci. Instrum. Method. Data Syst. Discuss., 2, C90–C94, 2012
www.geosci-instrum-method-data-syst-discuss.net/2/C90/2012/

Geoscientific Instrumentation Methods and Data Systems
Discussions

Interactive comment on "Innovations and applications of the VERA quality control" *by* D. Mayer et al.

D. Mayer et al.

dieter.mayer@univie.ac.at

Received and published: 2 August 2012

Thank you for reviewing our paper and giving valuable remarks to improve the article.

Answers to general comments:

We agree with the statement concerning the "benefit related to the combination of the human and automatic quality control" - especially in the creation of processed climatological data sets. Also with the quality of precipitation fields we see a high demand of more research to be carried out. Significant improvements could be expected by inclusion of additional data e.g. remotely sensed fields (radar or satellite).

Answers to the specific comments:

1) The reviewer addresses the problem of rejecting extreme weather related observations and points out that unusual atmospheric phenomena might be rejected by our gross error recognition relying on station statistics by mistake. At this point we would like to abstract the process of the gross error recognition based on station statistics: For every station, deviations are computed and stored. After some time period we use these deviations to characterize the station specific performance and calculate thresholds to indicate remarkably high deviations for each station, respectively. These thresholds may vary from station to station depending on their individual performance. If new measurements shall be controlled, VERA-QC first computes deviations for all these observations. Subsequently, not the measurements themselves but the deviations are compared to each station specific threshold. Compared to a station's observation history, a rare extreme weather event surely results in unusual measurement readings at this station. But if the extreme phenomenon is of such a horizontal extent that also neighboring stations are affected (e.g. consider a strong cold air-outbreak), the deviations computed by VERA-QC (by comparing neighbored stations) will not be extreme. Consequently, these deviations do not necessarily exceed the station specific thresholds. Of course, if the extreme event only affects one single station (e.g. local convective precipitation or downbursts), the buddy check of VERA-QC identifies this observation to be non-representative for the measuring network and rejects it as gross error.

Concerning the reviewer's comment regarding the "unavoidable tradeoff between the necessity of accounting for unusual atmospheric phenomena without using exceedingly permissive thresholds", we totally agree that the threshold based rejection of extreme weather related observations might be challenging and problematic - especially if decisions depend on measurements describing the atmospheric state as accurately as possible (e.g. severe weather warnings). The setting of such thresholds is strongly constrained by the intended usage of observations and the relationship between hits, misses, false alarms and correct rejections has to be optimized individually. However, the intention of VERA-QC is to provide representative measurements regarding the

available observation network. This means, VERA-QC only accepts atmospheric phenomena of a scale that is large as compared to the mean station distance.

Regarding the reviewer's second statement in point 1 (the chosen threshold values for the station selection algorithm (page 211)): In the text we tried to explain why we chose these three criteria for selecting representative stations (avoiding reduction errors, excluding stations that are most probably not representative). The process of optimizing the chosen threshold values can be considered as an iterative process taking quite some time. Experienced meteorologists supervise and evaluate the results achieved by a certain set of thresholds. According to their suggestions, the thresholds are altered. This procedure is being repeated constantly with the goal to optimize the proportion of rejected observations and that of accepted not representative observations for our available observational network (thus, for the available station network density). There are no objective configurations valid for an arbitrary observational network and the human component is essential at this point. The thresholds for the station specific gross error recognition are chosen subjectively as well (page 219 & 220, section 3.4.3). We would like to add that the +/- 4°C threshold (page 220, line 4) is only effective in the last instance of the station specific gross error recognition. It only prevents to reject observations based on corrections smaller than +/- 4°C even if the station statistics suggest to do so.

Thank you for pointing out the absence of the threshold concerning the cost function reduction for gross error recognition. Operationally, we chose a cost function reduction of 80% to indicate a potential gross error. Additionally, the absolute value of the involved unweighted deviation has to exceed the median of all absolute deviations within the whole domain by a certain factor (in our setting 30 times).

2.) The reviewer requests to comment the decision to exclude stations from the VERA-QC even if they deliver accurate information.

Operationally, we optimized VERA-QC (and especially the station selection algorithm)

to deliver the best input data for VERA analyses. The object of VERA is to analyze two dimensional surface observations and to give adequate meteorological information in valleys and plains (Minimum Topography). VERA analyses are used for nowcasting and are a valuable tool for distinguishing air masses as well as for identifying fronts. Thus we are primarily interested in the horizontal synoptic- and mesoscale and tend to exclude observations representing vertical small scale variability in complex topography which is done by the station selection algorithm. Naturally, this station selection algorithm can be skipped (cf. section 4.3. application of VERA-QC to COPS data) or altered depending on individual requirements. We intend to compute VERA analyses also for different pressure levels and consequently we will have to adapt the station selection algorithm and VERA-QC's thresholds accordingly. Depending on the pressure level, it will be necessary to accept mountain stations and therefore exclude observations from valley floors.

Concerning the usage of stations above 1500 m (operational limit) in figure 4: In order to investigate the height dependency of the bias correction in more detail, we raised the threshold for this case study to 1800 m. This shift resulted in the additional acceptance of three stations which fulfill the gradient- and Minimum Topography criteria.

3.) The reviewer addresses VERA-QC's ability to automatically adapt to varying station densities and describes it as a spatial filter. In areas of a denser station network, VERA-QC can - with a higher reliability - control, confirm, or criticize measurements that were caused by phenomena of a wider range of meteorological scales. The probability of gross error detection for an error of the same magnitude decreases with increasing median station distance. This can be explained by the fact that an error of a certain magnitude leads in a dense observation network to a higher curvature of the observation field than it does in areas of lower station density. In order to quantify this relationship, we investigated the dependency of the absolute value of a deviation (identified as a gross error) on the median of station distances (considering only primary neighbors). Expressing this relationship in terms of a linear regression, we found

an dependence of approximately +3.5 K/1000 km regarding potential temperature as example (based on results from July 2011 to June 2012).

4.) Concerning the statement "one would see nothing but the influence of the topography‘s height": We totally agree that the three dimensional temperature field is a valuable information (e.g. for estimation the 0°C isoline height). Naturally, one would use observations representing three dimensional gradients if a three dimensional analysis model is available. As long as VERA is carried out for two dimensions, we are restricted to consider observations that are representative for the Minimum Topography. We will reformulate this statement.

5.) The reviewer correctly noticed, that we did not discuss the VERA-QC application for wind (as we originally intended to). In the course of writing the paper, we decided to omit this content in order to keep the number of pages reasonable. Thank you for your precise review and this advice. We will rewrite the statement.

We agree with the technical corrections and we will apply them in the final paper.

Interactive comment on Geosci. Instrum. Method. Data Syst. Discuss., 2, 205, 2012.