Atmos. Meas. Tech. Discuss., 3, C2221–C2222, 2010
www.atmos-meas-tech-discuss.net/3/C2221/2010/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Assessment of BSRN radiation records for the computation of monthly means" *by* A. Roesch et al.

Anonymous Referee #2

Received and published: 9 December 2010

The paper under consideration addresses a simple question, namely how to build robust means of time series which have pronounced gaps. While the paper treats monthly averages for radiation fluxes from the Baseline Surface Radiation Network (BSRN), the same ideas can be applied to various data sets with gaps, as the authors point out. The scientific significance is therefore present, and the paper should be considered for publication.

There are, however, some significant deficiencies in the paper. First and foremost are its far-reaching statements, e.g., "it is impossible to accurately 'manufacture' values corresponding to the missing data." This may seem intuitive, but it is by no means necessarily true, and no mathematical discussion is offered to support these statements.

The authors should be encouraged to take a look at the field of Robust Statistics, e.g., the review of Huber (1972) or Hampel (1986, et al.). Robust estimators have been derived in subsequent works, and have been applied to incomplete data (Frahm 2008) and to non-stationary processes (Horenko, 2010). The paper needs to have a much stronger mathematical foundation if the authors wish to make the bold statements which appear at the end of section 2.3.

Secondly, the authors state vaguely that "there are many possibilities for methods of forming monthly averages from incomplete data" and present seven methods for calculating monthly means on BSRN data. While it is mentioned that these methods serve a demonstrative purpose, it is not clear whether any of these methods are actually used by practitioners (if so, references need to be added). Some of the methods are clearly inappropriate, e.g., M1, which uses all data, including those flagged by the quality control. M3 and M4 include interpolations within the gaps, a technique which is not viewed as a optimal (Horenko, 2010). M5, M6 and M7 are heuristic methods which include apparently arbitrary thresholds, for instance: "a single bin requires at least 20% valid data", "there must be at least 60% of the possible data available", "there must be at least 1300 minutes of the possible 1440 minutes overall data available", or "at least 95% of the expected hourly means must be present". These choices are neither justified with mathematical/statistical arguments nor with intuitive ones.

The authors are encouraged to provide a much stronger mathematical foundation for the paper, or, if this is not possible, to thoroughly explain the selection of the seven methods; for example, which methods are actually used in scientific evaluations, which were chosen to demonstrate technical problems (and in this case, which technical problems they accentuate), and why thresholds are chosen as they are.

Interactive comment on Atmos. Meas. Tech. Discuss., 3, 4423, 2010.