Clim. Past Discuss., 3, S100–S102, 2007
www.clim-past-discuss.net/3/S100/2007/


Climate of the Past Discussions



# *Interactive comment on* "On the verification of climate reconstructions" *by* G. Bürger

**G. Bürger**

Received and published: 23 March 2007

The comments of the reviewer are welcomed and will be considered accordingly. The following points, however, call for a response.

1. ("skill"): - For the introductory statements, "skill" was intentionally used in a more colloquial and general sense. For *any* skill score measuring the correspondence of a modeled quantity to observations the following is almost axiomatic: If a model, calibrated by whatever method, is applied to independent data the skill is expected to shrink. I should be very surprised if the rev disagrees on this point. (If he/she is able to provide a counterexample it will of course be considered accordingly.) The actual law of shrinkage certainly depends on the chosen model and score. Eq. (4), for example, applies to cross validity, $R_c$, for model estimates based on least-squares. The scores mentioned by the rev, such as odds ratio (ORSS) and Peirce (PSS), are binary and as such not really appropriate for climate reconstructions.

2. ("model imperfections"): - It appears that the notion of model imperfection has not come across as intended. Any model of the form $y = b * x +$ "noise" will show, when estimated from a finite sample, errors in the parameters $b$ (which I shortly call the "model") that render it imperfect.

3. ("independence"): - "Independent" is not meant here in an absolute sense, as for time-series or as in "iid". In a calibration/validation context it is common to consider the validation data *independent of* the calibration data. As an example, let $(x_c, y_c)$ be data pairs used for calibrating a model (e.g. that of 2. above). Here $y_c$ is not independent of $x_c$ (one hopes), but the validating pairs $(x_v, y_v)$ should be independent of $(x_c, y_c)$.

4. I will reread Thorne and Stephenson and look for pedagogical advice when revising abstract & introduction.

5. Two assumptions were made in section 2: That the entire population is used for validating, and that the observed validation mean is zero. Only the latter is needed for Eqs. (2) and (3), and both are needed for relating those to Eq. (4). But they are really not essential and only serve notational ease. It is not difficult to see that all equations remain valid without these assumptions, using the following, slightly more general expressions for the three biases $\alpha$, $\beta$, and $\gamma$:

$$\alpha = \frac{\mu_c - \mu_v}{\sigma_v}$$
$$\beta = \frac{\mu_v + \hat{\mu}_v}{\sigma_v}$$
$$\gamma = \frac{\hat{\sigma}_v}{\sigma_v}$$

where $\mu$ and $\sigma$, with and without hat, denote mean and standard deviation of simulated and observed series, respectively (the subscript indicating calibration and validation). The revision will have this general version.

6. It is certainly interesting, and was considered as an option, to investigate the performance of alternative methods and compare them with the existing ones. But since, according to the rev, the paper already handles a lot of complex statistical concepts it



CPD

3, S100–S102, 2007

Interactive Comment

was probably a good idea to discard that plan.

---

Interactive comment on Clim. Past Discuss., 3, 249, 2007.