

Exposé of pseudo-phenological nonsense

Dr. Henning Scholz

Project coordinator - BHL-Europe

Leibniz Institute for Research on Evolution and Biodiversity

Museum für Naturkunde, The Humboldt University Berlin

Dear Henning,

Pseudophenology and closed science at the Royal Botanic Garden Edinburgh

You ask for details of my firm decision to disassociate myself from *Sibbaldia* and the Royal Botanic Garden Edinburgh; and for any thoughts I may have on best practice guidelines for The Biodiversity Heritage Library's content acquisition process, and on what I like, and what I don't like about and BHL, and indeed on both the specific, and general, issues involved in this sad dispute with RBGE – its misleading of the general public - and especially with the low publishing standards of *Sibbaldia*. Quite a list - here goes ...

1. Peer Review - The Cornerstone of Modern Science

Any publication system has to have a way of pointing out errors by authors, judging whether they are indeed errors by the authors (and not by the reader), and correcting them. This is the crux of the matter. *Sibbaldia* has NOT been using a sensible peer-review process^a, and consequently RBGE has fallen into the trap of totally failing to recognise the gross scientific misconduct associated with its internal phenological reporting over the last half-decade. **[This is RBGE's key managerial failing].**

Of course peer review is by no means a perfect system—flaws ranging from outright fraud to subtle errors can easily slip past reviewers—but peer review can generally identify cases where a paper's conclusions aren't supported by the underlying data.

An equally important safety net, that has been lacking at *Sibbaldia*, and that should have been acting as a key barrier in preventing scientifically unsound ideas from filtering through RBGE, is the encouragement of outside readers to exchange information, ask questions of authors, or report conflicting data or alternative viewpoints via a 'letters to the editor' section^b. In all the scientific journals with which I have held an editorial role (*TRSE*, *SJG*, *GJRAS*, *PEPI*) the 'letters to the editor' section has always played a particularly valuable function in the drive to maintain standards. RBGE has unswervingly ignored the many attempts I have made over 7+ years to try to help them understand the pseudo-scientific phenological nonsense that has been, and continues to be, perpetrated under their name.

2. Bad Science

The UK government's chief scientific adviser, John Beddington, is currently challenging fellow scientists to reach up to the highest standards by stepping up the war on pseudo-science and to be "grossly intolerant" if science is misused. Beddington says "We are not grossly intolerant [enough] of pseudo-science, the

building up of what purports to be science by the cherry-picking of the facts and the failure to use scientific evidence and the failure to use scientific method." He urges that we should be "*completely intolerant of this nonsense*" and "*that we don't kind of shrug it off*".

Now as you point out "*Scientists usually are very much in favour of allowing digitisation and online display of their work, as it facilitates the search and retrieval of literature for them as well.*" You add "*Therefore, I'm surprised that you refused the upload of your article to BHL-Europe*". Why the refusal? My answer is that I strongly endorse Beddington's viewpoint - somebody needs to take a stand over the pseudo-scientific nonsense currently being ingested into the Biodiversity Heritage Library via *Sibbaldia*.

3. Sibbaldia and Pseudo-phenology

Harper's *Sibbaldia*^{1,2,3} claims represent the most flagrant abuse of the scientific method that I have encountered in over 40 years of research and of teaching applied data analysis.

Use of Misleading Language

Harper et al. create scientific-sounding terms (e.g. '*botanical regression*', '*thermal deceleration*') in order to add weight to their claims and thereby persuade non-experts to believe statements that are meaningless. Harper et al. use scientifically well-established terms (e.g. '*least-squares*', '*multiple regression*', '*stepwise*', '*r²*') in false, or idiosyncratic ways, thereby making their mindless conjectures sound respectable.

Absence of Progress

Apparent statistical significance is achieved just by mere repetition (of non-standard, eccentric assessments and of invalid procedures) until a success occurs due to chance variations. Also, despite the impressive methods that are claimed ('*multiple regression*' and so on), those claims as pointed out (pers. comm.) by R.M. Clark (one of the two top phenologists from the southern hemisphere) are outright lies, because they do NOT do what they assert.

Lack of Reproducibility - the "Touchstone" of Scientific Method

As highlighted above *Sibbaldia* and RBGE's major failing is to have fallen into the trap of lazily endorsing and promoting procedures which are scientifically irreproducible. If you go through RBGE's data-dredging procedures (Harper, Latta & Morter; Harper) you will find it bundles thousands of daily weather data into literally hundreds of millions of potential linear predictors^c. These embryonic predictors are then trawled through, until an aggregated correlation coefficient of say .95 is unearthed. Incredibly as few as eight data points are all that RBGE requires in order to identify the short/freak weather events that it alleges to have determined, for the first time ever, as the dominant controllers of the timings of plant development. RBGE's approach is the botanical equivalent of endlessly scanning photographic images of the back of the Moon for huge faces; or of forecasting psychic-experiences using planetary alignments; or in legal circles of the 'prosecutor's fallacy' where a well-

known, notorious, example is the covert statistical matching of DNA profiles, but where the large sizes of the databases used have elevated the likelihood of finding a match to pure chance alone. **[In short RBGE's major scientific gaff is to fail to recognise that their never-ending "cherry picking" leads to results which are neither reproducible nor repeatable, i.e fake-science.]**

Harper et al.'s boast of possessing privileged access to new botanical knowledge (which might go some way to exonerating their extreme data-dredging, but not their cherry picking) is utterly false. This is easily demonstrated by their inability to quote a single reference which contains any well-defined, explicit, biological information rather than worthless generalities.

Another specific example of the RBGE nonsense is given in Fig. 10 of Harper & Morris, where the ludicrous claim of climate change driving botanical behaviour at rates of 17.2 days/year is 'dressed up' by the use of apparently technical jargon in an effort to give their assertion the superficial trappings of science, and in order to lay claim to a newly identified "*plant functional group*". Here, by focusing only on the most extreme plant behaviour available to them (one that conveniently happens to confirm their prior notions or to validate their hunches), Harper & Morris are merely indulging in naive cherry-picking.

A final, third, example is Harper's statement in *Sibbaldia* that at RBGE "*with such a small data set for each plant, formal tests of significance are not appropriate*". A more startling lack of critical self-appraisal can scarcely be imagined.

In science replication means the repetition of a scientific experiment, or trial, to obtain a consistent result. Ideally, as a check on the validity of Harper et al.'s claims, one would want to independently re-analyse the original data. However, this entails the need to know the integrity of the data, i.e., whether the alleged original data was actually original. Nevertheless it is an internal test that RBGE could carry out if it so wished.

Alternatively one can run simulations (as I have done) with equivalent data but of known properties. Or more simply one can consult (as I have done) the relevant statistical tables. Both these tests demonstrate the irreproducibility and hence the utter nonsense underlying the botanical breakthroughs that *Sibbaldia* claims. Once the lack of reproducibility in the RBGE work is recognised, the whole phenological edifice being erected through Harper's ongoing swath of *Sibbaldia* articles totally collapses.

4. Best Practice - Signal-to-noise-ratio

As Goldacre points out "*the signal-to-noise ratio in the scientific literature is getting ever lower*". Sadly, when BHL ingests *Sibbaldia* 5, 7, 8... the phenological signal-to-noise ratio will significantly drop. In the same vein Kirby⁴ notes "*The internet provides an unparalleled distribution medium but the central issues of quality assurance for science publishing remain. The active researcher who wishes to keep abreast of developments in their field simply cannot read everything that is of possible relevance. We rely on others to raise published work to a level that is worth spending the time to read.*"

Bauer⁴ addresses the question of best practice amongst journals. He notes “*The ‘quality’ of a scientific publication is not an absolute but must be assessed in relation to a journal’s mission. It should be judged primarily by its disciplined intellectual rigour, bearing in mind what course the publication aims to take, as between the Scylla of hidebound conservatism and the Charybdis of mindless speculation*”. Best practice must relate to where BHL is aiming to position itself along Bauer’s continuum from intellectual rigour through to mindless speculation.

In my estimation uncritical ingestion of journals like *Sibbaldia* into BHL will not be doing the scientific world a service. And that is the reason why I wish to be disassociated with *Sibbaldia*, with RBGE, and hence with BHL. In this instance the general public, non-experts, and indeed many phenologists will undoubtedly be taken in, and misled, by RBGE’s former reputation and by the BHL badge.

As RBGE is continuing to actively solicit, for its own, closed, internal publications, further phenological articles BHL can anticipate receiving many more nonsensical articles from *Sibbaldia* as Harper cherry-picks his way, by hand, through hundreds of further RBGE plants.

5. Lessons That Can Be Learnt

BHL-Europe appears to be run professionally. So the main lessons, from this debacle, fall squarely on RBGE. You ask for my views. I see three major problems at *Sibbaldia* and RBGE:

First, at RBGE, I have no doubt the major failing lies with the management. They have consistently ignored many experts’ advice (including professors of ecology, members of highly-regarded academic institutes, eminent statisticians, and the professional statistical consultants - BIOSS). The whole problem could, and should, have been addressed and solved years ago, rather than in Beddington’s words of just perniciously “*shrugging it off*”.

As more and more institutes come to rely on amateur volunteers, they inevitably will have to deal with cranks, charlatans, stubborn deluded amateur hobbyists, and worse. They will need robust, effective ways of identifying and weeding out the worst culprits.

Secondly, *Sibbaldia* has had an ineffective review system^a.

A statement by BHL of its own assessment of the quality of the peer-review process, as operated by each of its consortium members, could greatly assist BHL’s readership understand where individual journals fall along Bauer’s continuum. Otherwise BHL is effectively providing a means of the evasion of peer review before the publicizing of results.

Thirdly, and most importantly, a genuine ‘letters to the editor’ section at *Sibbaldia* and an equivalent e-version at BHL, would go a long way in helping raise the intellectual rigour of *Sibbaldia* and of the Biodiversity Heritage Library project.

References

1. Harper and Morris 'Flowering and Climate Change – Part II', *Sibbaldia* **5**:25-42.
2. Harper, Latta & Morter 'A correlation and regression approach to phenology' *Sibbaldia* **7**:139-158.
3. Harper 'Lessons from phenology: an interim report' *Sibbaldia* **8**:149-164.
4. Quality in science publishing. <http://www.int-res.com/articles/theme/m270p265.pdf>

Notes

- a. As a specific example A. Roberts (pers. comm. 2011) describes how his referees' comments, for *Sibbaldia*, have been totally ignored.
- b. *Sibbaldia* has no viable 'Letters to the Editor' section.
- c. - Potential weather parameters? Max, min, mean temperature, rain, frost days, radiation, sun hours, etc (say a minimum of 6).
 - Potential time periods? Monthly, 10-day, weekly, 5-day. (Hence (12 + 36 + 52 + 72), all using the Christian calendar, or solar calendar etc; say a minimum of 150).
 - Potential years? 2, or more, previous years (say minimum of 2).
 - Number of individual predictors? When merely adding together correlation coefficients, as here, there is no upper mathematical limit. (Let's conservatively say the number of data points minus 1, or even 5).
 - Hence total number of available predictors = (6 x 150 x 2) all raised to the power 5 = 10^8