Homogeneity Testing of Microanalytical Reference Materials by Electron Probe Microanalysis

Dennis Harries[1]

[1.] Department of Analytical Mineralogy, Institute of Geosciences, Friedrich Schiller University, Jena, Germany.

Well-characterized and suitable reference materials are key to accurate and reproducible chemical and isotopic microanalysis. The testing of homogeneity of potential microanalytical reference materials requires discriminating between the variations due to the instrumental precision and the variations due to true compositional differences. The theoretically easiest approach is to determine the instrumental precision by repeatedly analyzing the same spot on the sample and then compare different sites on the sample by an analysis of variance (ANOVA). However, in practical microanalysis this approach is expected to fail often, because repeated analysis of a single spot will often result in progressive degradation (or even loss) of the analyzed volume. In electron probe microanalysis (EPMA) this degradation is due to the build-up of surface contamination, induced diffusion, and structural damage. In consequence, the apparent instrumental variance is not applicable to sound statistical analysis.

In principle, perfect homogeneity in materials does not exist. It has to be defined what scale and relative contribution of heterogeneity can be accepted for a given purpose. Testing for homogeneity then tries to answer the question whether a material is 'fit for the purpose'. This question can be answered positively when significant heterogeneity cannot be detected with the method and its chosen conditions of operation. These are specifically the spatial and the 'compositional' resolution, the latter determined by the instrumental precision and the resulting uncertainty.

In the case that only one measurement per analysis spot is feasible, significance criteria for detectable heterogeneity can be derived in a statistically sound fashion if Poisson statistics of X-ray counts are the only significant contribution to the instrumental precision (and this precision is reasonably high). A traditional approach to this is the recently refined Homogeneity Index (H) [1, 2]. The Homogeneity Index compares the variance expected from counting statistics to the variance observed. If the latter is larger than the expected variance ($H > 1$), compositional heterogeneity may be significant. This can be tested based on F or chi-squared statistics. The answer is valid only for the specific set of measurement conditions that were used for the test.

The total uncertainty budget of a reference value is, in the simplest (accurate) case, composed of uncertainty due to instrumental precision and compositional heterogeneity. The Homogeneity Index can be used to derive a relative contribution of compositional heterogeneity ($s_{h,rel}$) to this total uncertainty budget. However, the approach is limited by the non-linear relationship between H and $s_{h,rel}$ (Fig. 1), resulting in a bias that yields apparently large contributions of compositional heterogeneity even if materials are almost perfectly homogeneous. Because the uncertainty of the Homogeneity Index is primarily determined by the number of measurements N, this bias and the danger of statistical type II error can be reduced by increasing N (Fig. 1). In consequence, the resolving power of a homogeneity test in terms of detectable heterogeneity is strongly related to the number of measurements. For example, in order to state that the contribution of compositional heterogeneity to the total uncertainty budget is *less than* 30%, a homogeneity test with $N = 577$ measurements has to be passed ($H < 1.048$).

For 20% this number increases to more than 3000 measurements. What upper levels of heterogeneity contribution and analytical expenses are acceptable has to be decided by the user.

The criteria discussed are intended for testing for homogeneity in terms of compositional resolution. These criteria are not suitable to assess homogeneity in terms spatial resolution. For example, it is not possible to address how the perception of homogeneity changes when the spot size or the average spacing between spots changes. This important question needs to be addresses by other means.

References:
[1] F.R. Boyd *et al.*, Carnegie Institution of Washington Year Book **67** (1967), p. 210.
[2] D. Harries, Chemie der Erde **74** (2014), p. 375 (doi: 10.1016/j.chemer.2014.01.001).
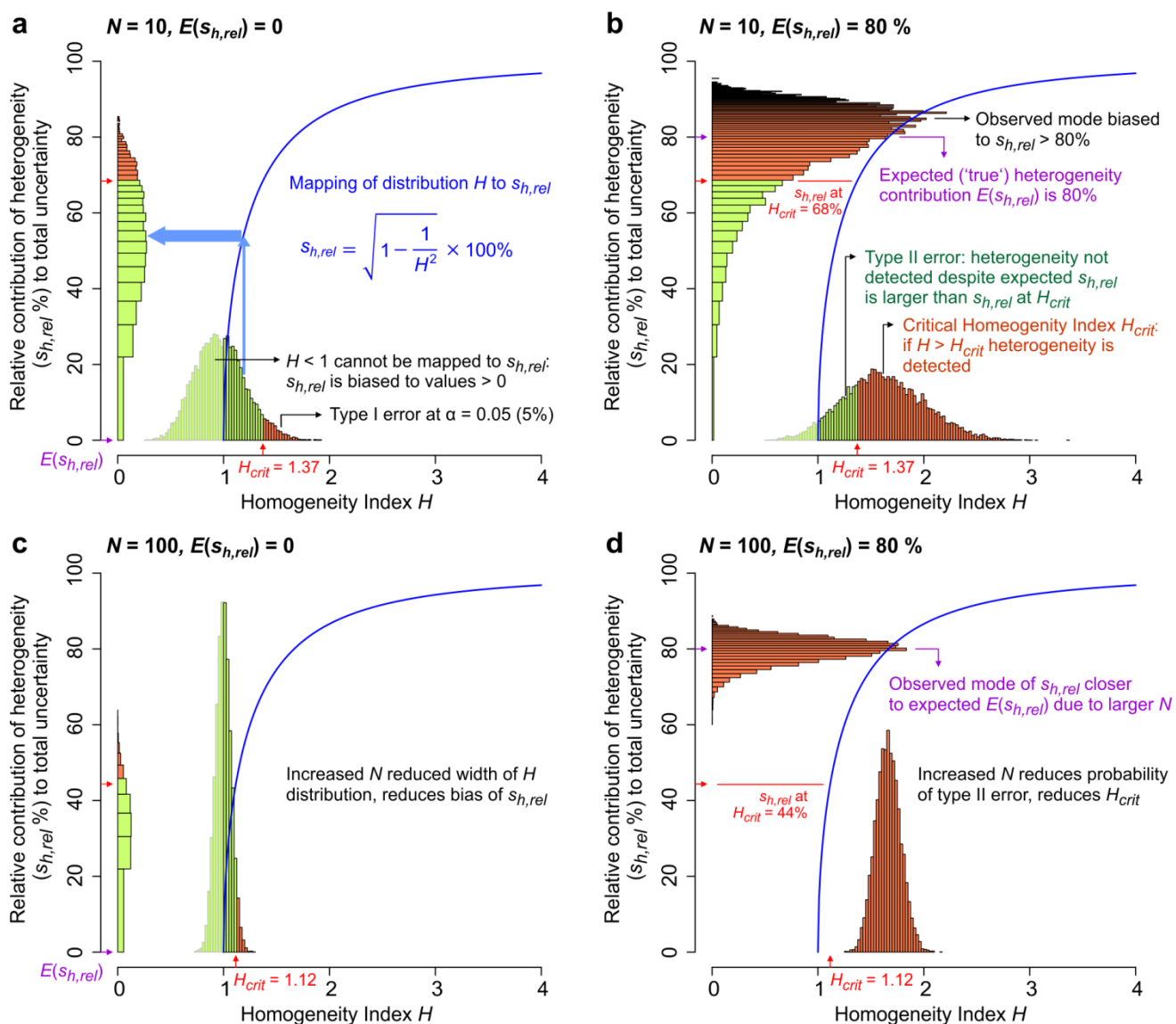
Figure 1. The important role of the number of measurements N in homogeneity testing. Shown are histograms of 10,000 simulated homogeneity tests with $N = 10$ (a, b) and $N = 100$ (c, d) measurements per test. The (simulated and therefore known) expected relative contribution of heterogeneity $E(s_{h,rel})$ to total uncertain is 0 (a, c; case of perfect homogeneity) and 80% (b, d; case of detectable heterogeneity).