## Reporting 40Ar/39Ar uncertainties

### Assessing goodness of fit: pitfalls and opportunities

The random scatter of the data around an isochron or weighted mean fit can be assessed using the Mean Square of the Weighted Deviates (MSWD, McIntyre et al., 1960). This statistic is more generally known as the ‘reduced Chi-square statistic’ outside geology. The MSWD is defined as the sum of the squared differences between the observed and the expected values, normalised by the analytical uncertainties and divided by the degrees of freedom of the fit. In the context of the weighted mean age, the MSWD of values is given by:

where is the th (out of ) dates, is the corresponding analytical uncertainty, is the number of degrees of freedom, and is the weighted mean of all dates. The definition for the MSWD of an isochron is similar but has one fewer degree of freedom () and involves a few more terms to account for correlated uncertainties between the x- and y-variable. The following are general MSWD considerations:

1. If the analytical uncertainties are the only source of scatter between the aliquots, and is reasonably large (for , say) then MSWD1 (Fig. 5A,B). For smaller sample sizes, the MSWD has a much wider distribution with an expected value of less than one (Wendt and Carl, 1991; Mahon 1996). Figure 1 of the Supplementary Information shows the probability distribution of the MSWD for different sample sizes. The remainder of this section will assume that .
2. MSWD values that approach zero indicate that analytical uncertainties have been overestimated or have not been propagated correctly (Fig. 5B,C). Assigning ages to samples based on such *underdispersed* data should be done with caution.
3. MSWD-values considerably greater than one indicate that there is some excess scatter in the data, which cannot be explained by the analytical uncertainties alone. This may reflect underestimation of analytical uncertainties, but usually reflects the presence of some geological *(over)dispersion* affecting the dataset and/or neutron fluence gradients. Possible causes of such dispersion may include the protracted crystallization history of a sample, variable degrees of inheritance, or partial loss of radiogenic 40Ar by retrograde reactions, thermally activated volume diffusion, deformation, or chemical alteration (Fig. D,E).

It would be wrong to believe that only datasets with MSWD1 are suitable for publication. Trimming an overdispersed dataset by selectively rejecting outliers until achieving an MSWD1 is likely ill-advised as this risks the loss of geologically valuable information and biasing the results. Outlier identification and rejection must always be accompanied by full disclosure of the specific criteria used for such evaluation, and not simply to improve the statistics of a dataset. High MSWD values do not necessarily indicate poor data. In fact, with modern-day high-precision mass spectrometery, a higher MSWD may simply reflect the achieved higher analytical precision of the data. Increasingly dispersed datasets are likely to become even more prevalent in the future, as a result of the ever-increasing improvements of mass spectrometers. In this case, the excess dispersion can be formally assessed with a Chi-square test for homogeneity, and its associated p-value, in the case the experiments were carried out with large number of heating steps or total fusions of single crystals. However, unpowered (sensu Cohen, 1992) statistical hypothesis tests have come under criticism in recent years, and scientists are increasingly advised not use them (Wasserstein and Lazar, 2016; Amrhein et al., 2019).  
Dispersed datasets should always be evaluated carefully on a case by case basis, and any conclusions based on dispersed data should be done with caution. It is important to attempt to consider the potential causes (geologic, analytical, mineralogic, etc.) of the data dispersion (e.g., Verati and Jourdan, 2014; Phillips et al., 2017). In some cases, a subset of a dispersed dataset can be used to assign an age for a sample given sufficient geologic context. For example, single crystal fusion dates from a volcaniclastic layer intercalated within a fluvio-lacustrine succession along the Tiber River, Italy show significant dispersion (MSWD = 603; Marra et al., 2019). The volcaniclastic layer has lithologic and mineralogic characteristics that are nearly identical to another volcaniclastic layer located   6 km to the northwest that was dated at 327.5 3.5 ka. The youngest six 40Ar/39Ar dates of the dispersed dataset give a weighted mean age of 328.7 1.6 ka, which led Marra et al. 2019 to conclude that the two dated volcaniclastic layers are indeed identical and have been tectonically displaced by 50 meters.  
When no potential sources of data dispersion can be confidently identified, it can be assumed that the excess dispersion is multiplicative and scales in proportion to the analytical uncertainty. In this case, the standard error of the weighted mean or isochron intercept may be augmented by multiplying it with the square root of the MSWD (Ludwig, 2003). A second option is to parameterize the overdispersion as an additive term and estimate it as a separate parameter (Galbraith and Laslett, 1993; Vermeesch, 2018).  
**Additional references:**  
Cohen, J., 1992. A power primer. Psychological bulletin, 112(1), p.155.  
Galbraith, R.F. and Laslett, G.M., 1993. Statistical models for mixed fission track ages. Nuclear tracks and radiation measurements, 21(4), pp.459-470.  
Ludwig, K.R., 2003. Mathematical–statistical treatment of data and errors for 230Th/U geochronology. Reviews in Mineralogy and Geochemistry, 52(1), pp.631-656.  
**Supplementary Figure:**

![Probability distribution for the Mean Square of the Weighted Deviates (MSWD) for different degrees of freedom. Small samples (df<20) are characterised by skewed and broadly distributed MSWD distributions whose mean, median and mode are all less than 1. With increasing sample size (df>20), the probability distribution of the MSWD becomes more symmetric converges to an expected value of MSWD\approx1. ](data:application/pdf;base64,)

Probability distribution for the Mean Square of the Weighted Deviates (MSWD) for different degrees of freedom. Small samples () are characterised by skewed and broadly distributed MSWD distributions whose mean, median and mode are all less than 1. With increasing sample size (), the probability distribution of the MSWD becomes more symmetric converges to an expected value of MSWD1.