Graphically comparing the uncertainty of the original data and that obtained by the differencing method

Fig.1a and b shows more visual variability in the differencing series compared with the original data. Does the differencing method 'overestimate' the uncertainty? Formally, this question is easier to discuss for a non-transient signal. Let us compare the uncertainty obtained using formulae (9) and (14):

$$\frac{\sum_{i=1}^{n}(I_i - I_{mean})^2}{n(n-1)} \approx ? \approx \frac{\sum_{i=1}^{n/2}\left[(I_{2i-1} - I_{2i}) - (I_{2i-1} - I_{2i})_{mean}\right]^2}{n(n-2)}$$

For large n, denominators can be omitted in this relationship:

$$0 \approx ? \approx -2\sum_{i=1}^{n/2}(I_{2i-1} - I_{2i-1\ mean})(I_{2i} - I_{2i\ mean})$$

This relationship transforms to equality, if there is no covariance between the even and the odd sweep intensities, as is supposed when using the differencing method (see main text) and demonstrated by an example presented in Electronic Appendix 4. Thus, formulae (9) and (14) lead to the same result.

These formulae characterise the mean intensity uncertainty, while the graphical variability in the differencing series and in the original data concerns the uncertainty of the individual differences or individual sweep intensities: what is the size of an interval, into which an individual difference, or individual sweep intensity, falls with some probability? The individual uncertainties are given as follows:

$$\frac{\sum_{i=1}^{n}(I_i - I_{mean})^2}{n-1} \approx ? \approx \frac{\sum_{i=1}^{n/2}\left[(I_{2i-1} - I_{2i}) - (I_{2i-1} - I_{2i})_{mean}\right]^2}{\frac{n}{2}-1}$$

which is equivalent to the following relationship:

$$\frac{\sum_{i=1}^{n}(I_i - I_{mean})^2}{n(n-1)} \approx ? \approx 2\frac{\sum_{i=1}^{n/2}\left[(I_{2i-1} - I_{2i}) - (I_{2i-1} - I_{2i})_{mean}\right]^2}{n(n-2)}$$

Given the discussion above, we obtain:

$$\frac{\sum_{i=1}^{n}(I_i - I_{mean})^2}{n(n-1)} \leq 2\frac{\sum_{i=1}^{n/2}\left[(I_{2i-1} - I_{2i}) - (I_{2i-1} - I_{2i})_{mean}\right]^2}{n(n-2)}$$

Thus, the graphical variability in the series of differences is higher than in the original series.

The same conclusion could be obtained faster by using a more formal derivation: if x_1 and x_2 are independent random variables from the same population, $Var\ (x_1-x_2) = 2Var(x_1) \geq Var(x_1)$.

The differencing principle does not imply the same uncertainty for the differences and the original data. It derives the uncertainty of the original data from that of the differences.