

Erratum to Bailer-Jones (2011), MNRAS 416, 1163–1180

Bayesian time series analysis of terrestrial impact cratering

Coryn A.L. Bailer-Jones, 29 September 2011

Summary: There is an error in the write-up of the method in sections 3.2 and 3.4 of the article. It is essentially just an error in the notation used, but this may interfere with a complete understanding of the method. There is no mistake in the concept or in the code, so this does not effect the results, discussion or conclusions in any way.

The issue is with equation 5 (section 3.2), the equation for the measurement model. Roughly speaking the measurement model gives the probability of getting the measured data given the true value. More precisely, it is the probability of measuring the event at time s_j given both the true time of the event, t_j , and the uncertainty, σ_j , in the measurement s_j . In other words, rather than being $P(s_j, \sigma_j | t_j)$, the measurement model is

$$P(s_j | \sigma_j, t_j) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-(s_j - t_j)^2 / 2\sigma_j^2} \quad (5)$$

which is normalized with respect to s_j (and not t_j as was written in the article). Comparing the above to equation 5 in the article, the right-hand-sides are identical to what we had before (the order of s_j and t_j is irrelevant for a Gaussian). In fact, the left-hand-sides are also identical if we consider the data to be $D_j = s_j$ rather than $D_j = (s_j, \sigma_j)$, and implicitly condition on the measurement uncertainty, σ_j . The correct perspective, which was not clear in the article, is that we should think of the measurement model as a model of s_j conditioned not only on the (unknown) true time, t_j , but also on a fixed (measured/estimated) noise model parameter, σ_j . Yet as we neither infer nor marginalize over σ_j in this work, it has no practical consequence whether we regard this as measured data or a fixed parameter.

This changes the notation in section 3.4, but not the actual content of the equations or the calculations. One can just replace D_j with s_j and remember that there is an implicit conditioning on σ_j . Nonetheless, for completeness, the second, third and fourth paragraphs of this section with the corrected notation are now given.

The probability of observing data s_j from model M with parameters θ is $P(s_j | \sigma_j, \theta, M)$, the *likelihood* for one event. The time series model predicts the true age of an event, which is unknown. Applying the rules of probability we marginalize over this to get

$$\begin{aligned} P(s_j | \sigma_j, \theta, M) &= \int_{t_j} P(s_j, t_j | \sigma_j, \theta, M) dt_j \\ &= \int_{t_j} P(s_j | \sigma_j, t_j, \theta, M) P(t_j | \sigma_j, \theta, M) dt_j \\ &= \int_{t_j} P(s_j | \sigma_j, t_j) P(t_j | \theta, M) dt_j . \end{aligned} \quad (6)$$

The last step follows from conditional independence: in the first term – the measurement model – once t_j is specified s_j becomes conditionally independent of θ and M ; in the second term – the time series model

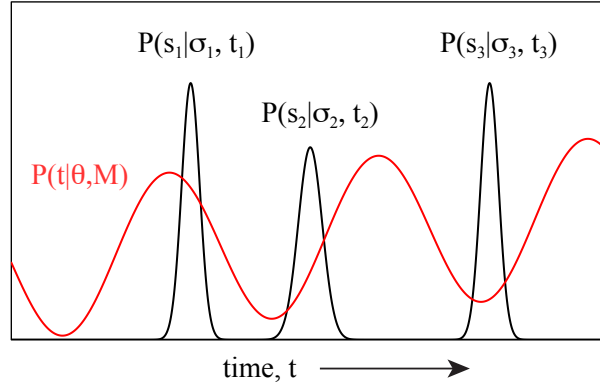


Figure 3: Principle of the likelihood calculation (equation 7)

– t_j is independent of σ_j . As the data are fixed, we consider both terms as functions of t_j . Both must be properly normalized probability density functions.

If we have a set of J events for which the ages and uncertainties have been estimated independently of one another, then the probability of observing these data $D = \{s_j\}$, the *likelihood*, is

$$\begin{aligned}
 P(D|\sigma, \theta, M) &= \prod_j P(s_j|\sigma_j, \theta, M) \\
 &= \prod_j \int_{t_j} P(s_j|\sigma_j, t_j) P(t_j|\theta, M) dt_j .
 \end{aligned} \tag{7}$$

where $\sigma = \{\sigma_j\}$. The principle of this calculation is illustrated in Fig. 3: the likelihood of an event for a given model is the integral of the probability distribution for the event (eqn. 5) over the time series model, $P(t_j|\theta, M)$. Specific cases for the latter are introduced in section 3.6 below.

The *evidence* is obtained by marginalizing the likelihood over the parameter prior probability distribution, $P(\theta|M)$,

$$\begin{aligned}
 P(D|\sigma, M) &= \int_{\theta} P(D, \theta|\sigma, M) d\theta \\
 &= \int_{\theta} P(D|\sigma, \theta, M) P(\theta|M) d\theta
 \end{aligned} \tag{8}$$

(where σ drops out of the second term due to conditional independence). For a given set of data (crater time series), we calculate this evidence for the different models we wish to compare, each parametrized by some parameters θ . The parameter prior $P(\theta|M)$ encapsulates our prior knowledge (i.e. independent of the data) of the probabilities of different parameters, normally established from the context of the problem (see section 3.7).

The consequence of this more precise notation is that it reminds us that the evidence, $P(D|\sigma, M)$, is actually conditioned on the fixed measurement uncertainties. This was not explicit in the original article, but the results, conclusions and discussion are entirely unaffected.