Age Zircon model

Methods

Main model parameters

Let us define a dataset of N dated zircons from S samples. We indicate with $\mathbf{z} = \{z_1, \ldots, z_N\}$ the (unknown) ages of the zircons and with $\mathbf{x} = \{x_1, \ldots, x_S\}$ the ages of the samples. The samples are ordered according on their depth and we assume their age strictly follows that order such that $x_i > x_{i-1}$. We indicate the set of zircons found in a sample s as $\mathbf{z}^s$.

The exact age of each zircon (z_i) is assumed to be unknown but linked to its measured age, which is expressed as a mean μ_i and a standard deviation σ_i. The mean ages μ and a standard deviations σ of all zircons represent the input data of the model, which aims to estimate the vector $\mathbf{z}$ and the ages of all samples $mathbfx$ Since the age of zircons can be measured based on different methods (e.g. $m \in \{1, \ldots, M\}$, e.g. Zr-U-Pb, ZFT, Ar-Ar), the uncertainty around the true age of a zircon is assumed to be further affected by how it was measured. We compute the likelihood of the age of a zircon i based on a normal density (Fig. 1):

$$P(\mu_i | z_i, \sigma_i, \epsilon_m) \sim \mathcal{N}(z_i, \sigma_i + \epsilon_m) \tag{1}$$

where $\epsilon_m \in \mathbb{R}^+$ is the bias in the estimated error of the measurement method m. The parameters z_i and ϵ_m are considered as unknown and estimated from the data using a Bayesian algorithm.

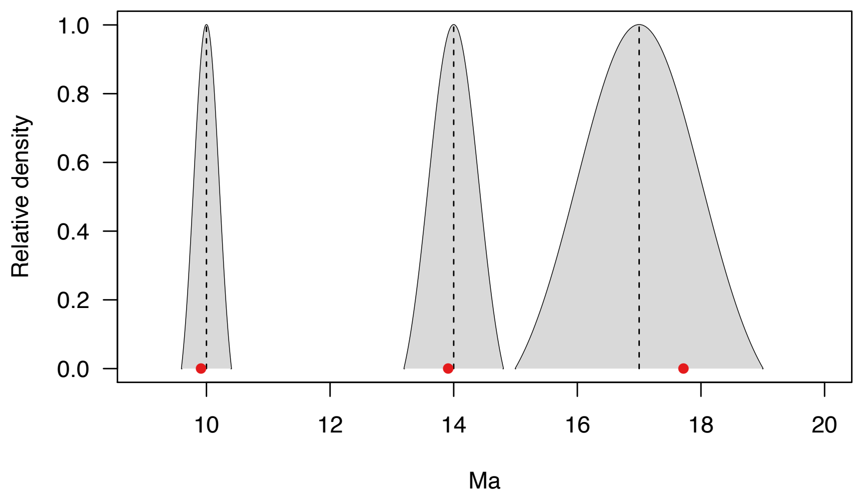

Figure 1: Example of three zircons with estimated ages of 10, 14, and 17 Ma (μ; indicated by the dashed vertical lines) and standard deviations ($\sigma + \epsilon$; shown by the gray shaded areas). The sampled true ages are indicated by the red circles.

The value of x_j, the age of the jth sample, is determined by two latent variables (r_j and $\mathbf{I}^j$) and constrained by sampled values of $\mathbf{z}$ such that $x_i > x_{i-1}$ for $i \in \{2, \ldots, S\}$. Specifically, we define as

$$\zeta_j = \min(\mathbf{z}^s), \text{ for } s \in \{j, \ldots, S\}, \tag{2}$$

the minimum age across all zircons included in sample j and in all older samples. Thus, ζ_j represents the upper (older) boundary for the age of sample j, which must be younger than all following samples (ordered by depth) and than its youngest zircon. Under this notation we define the age of a sample as:

$$x_j = \begin{cases} r_j(\zeta_j), & \text{if } j = 1 \\ x_{j-1} + r_j(\zeta_j - x_{j-1}), & \text{if } j > 1 \end{cases} \tag{3}$$

where $r_j \in (0, 1)$ is an estimated sample-specific latent parameter determining how close x_j is to its lower boundary x_{j-1} (or to 0 if $j = 1$). Thus, when $r_j \approx 0$ the age of the sample is close to its younger boundary and when $r_j \approx 1$ the age of the sample is close to its older boundary. We also account for the fact that a zircon can be younger than the sample it was found in, for instance due to a dating error or a later recrystallization of the zircon. To account for this possibility, we additionally estimate a vector of identifiers $\mathbf{I} = \{I_1, \ldots, I_N\}$ that define which zircons (identified by $I = 1$) are older than the sample and there fore used to determine its upper age boundary and which zircons (identified by $I = 0$) are younger than the age of the sample. Thus, the upper boundary of a sample age is:

$$\zeta_j = \min(\mathbf{z}^s \backslash \mathbf{I}_0^s), \text{ for } s \in \{j, \ldots, N\}, \tag{4}$$

where $\mathbf{z}^s \backslash \mathbf{I}_0^s$ indicates the subset of zircons in sample s with indicator equal to 1.

We can now define the prior probability of z_i^j, i.e. the ith zircon in sample j as a function of the age of the sample (x_j) and an estimated scale parameter s_j. Specifically we model the prior distribution of zircons in a sample using a compound-Uniform-Cauchy distribution, defined as:

$$P(z_{i,j}|x_j, s_j) = \begin{cases} z_{i,j} \sim \mathcal{U}(0, 2x_j), & \text{if } z_{i,j} < x_j \\ z_{i,j} \sim \mathcal{C}(x_j, s_j), & \text{if } z_{i,j} \geq x_j \end{cases} \tag{5}$$

where s_j is the sample-specific scale parameter of the Cauchy distribution. Under this parameterization, the age of the zircons identified as younger than the sample will have a prior uniform probability ranging from 0 to the age of the sample and rescaled to integrate to 0.5. The other zircons will instead follow a half-Cauchy distribution with mode equal to the age of the sample (Fig. 2).

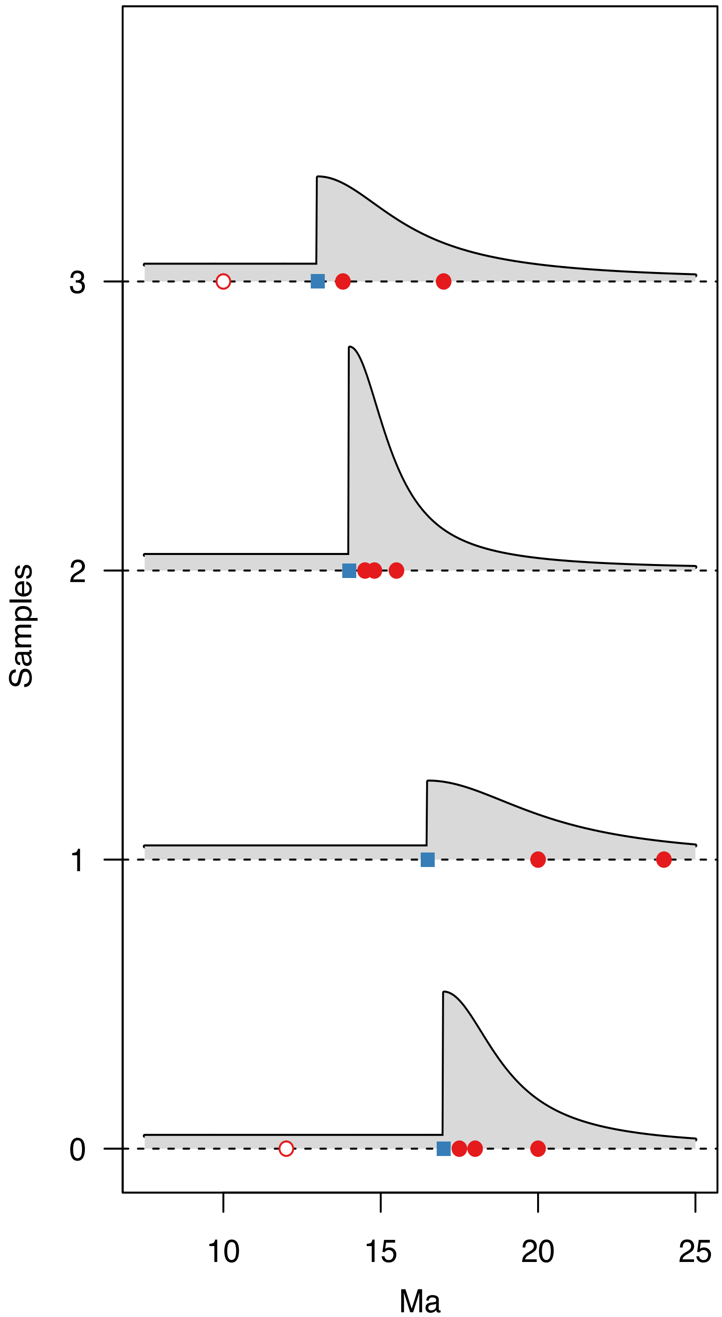

Figure 2: Probability distributions of zircon ages in four samples sorted from oldest (sample 0) to most recent (sample 3). Red filled circles indicate zircons identified as older that the age of the sample, shown as a blue square. Red empty circles indicate zircons identified as younger than the sample. The gray shaded areas display the relative probability distribution of the zircon ages multiplied by the probability of the respective indicator (here set to $P(I = 0) = 0.1$). We note that the maximum age of sample 1 (x_1) is not determined by its youngest zircon (of age 20 Ma), but by the age of the previous sample x_0.

Priors and hyper-priors

We use a half-Cauchy prior for the vector of scale parameters of the compound Uniform-Cauchy distributions $\{s_1, \ldots, s_S\} \sim \mathcal{C}^+(0, \beta)$, where the scale parameter β is assumed to be unknown and estimated through MCMC, with a gamma hyper-prior $\beta \sim \Gamma(10, 0.5)$. We set an exponential prior on the vector of parameters $\{\epsilon_1, \ldots, \epsilon_M\} \sim \text{Exp}(0.1)$. We use a beta distribution as prior on the vector of parameters $\{r_1 \ldots, r_S\} \sim \mathcal{B}(a, b)$ and consider the shape parameters a and b themselves as unknown parameters and assign them exponential hyper-priors, $a, b \sim \text{Exp}(0.1)$. Finally we use a Bernoulli distribution as prior on the indicators $\{I_1, \ldots, I_N\}$ with parameter $p = 0.99$, thus assigning a 0.01 prior probability for a zircon to be younger than the sample it is assigned to. This informative prior assumes that only a small fraction of the zircons might have re-crystallized or is otherwise erroneously dated.

Parameter estimation

The model includes $2N + 2M + 2S + 3$ parameters (z and I, ϵ, s and r, a and b and β, respectively). All parameters are estimated through Metropolis-Hastings Markov chain Monte Carlo (MCMC). We use a sliding window proposal with reflection at the boundaries for r, normal kernel proposals for z, binomial proposals for I. We use multiplier proposals on all other parameters since they only span the positive range.