Bayesian analysis of DO-climate events, version 3

Ditlevsen *et al.* has analyzed several models of the DO-climate events. These events has according to NGRIP data occured around the following years BP: 11700, 13130, 14680, 23340, 27780, 27920, 28900, 32500, 33740, 35480, 38220, 40160, 41460. The question is wether these events are triggered by some periodic external forcing.

In the paper several models are introduced, and their significance is tested by standard statistical methods. As always, the meaning of such methods are quite murky, while the problem can be analyzed by straightforward Bayesian logic. Further I shall base my analysis on a time series all the way down to 60000 years before present.

Assume we have two models, T_1 and T_2, who's probability given the data, D, we want to compare. Each model has a number of parameters, which I collectively will denote λ_1 and λ_2 respectively. Bayes' theorem plus a couple of standard rules of probability theory now gives

$$\begin{aligned} P(T_i|DI) &= \int P(T_i\lambda_i|DI)d\lambda_i \quad \text{(sum rule)} \\ &= \int \frac{P(D|T_i\lambda_i I)P(T_i\lambda_i|I)}{P(D|I)} d\lambda_i \quad \text{(Bayes)} \\ &= \frac{P(T_i|I)}{P(D|I)} \int P(D|T_i\lambda_i I)P(\lambda_i|T_i I)d\lambda_i \quad \text{(product rule).} \end{aligned} \tag{1}$$

If we want to compare the two models we simply take the ratio of their probabilities given data:

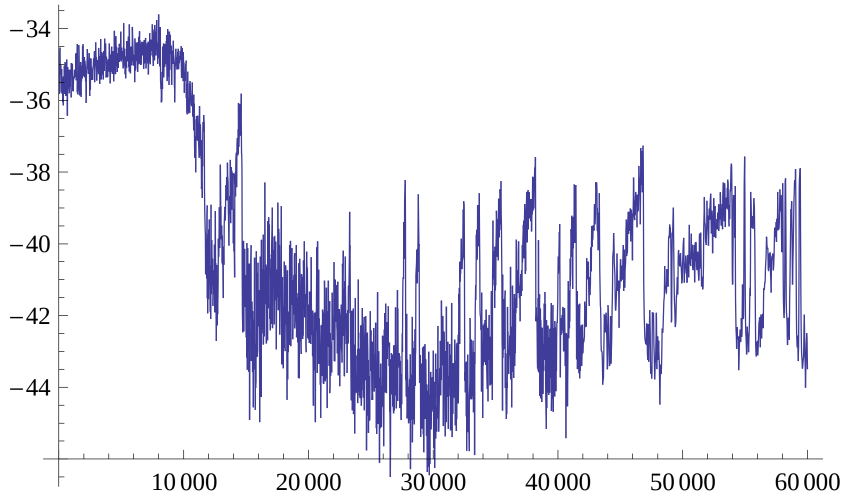
$$\frac{P(T_1|DI)}{P(T_2|DI)} = \frac{P(T_1|I)}{P(T_2|I)} \frac{\int P(D|T_1\lambda_1 I)P(\lambda_1|T_1 I)d\lambda_1}{\int P(D|T_2\lambda_2 I)P(\lambda_2|T_2 I)d\lambda_2}. \tag{2}$$

Here the first factor is the ratio of our beliefs in the models before we even have data. The next factor is essentially the ratio of the likelihoods of the two models. It can be easily calculated and tell us how we are to change our relative probabilities once we learn about the data. Note that these include the so-called *Occam* factors $P(\lambda_i|T_i I)$ which are 'penalties' for introducing too many parameters in a model.

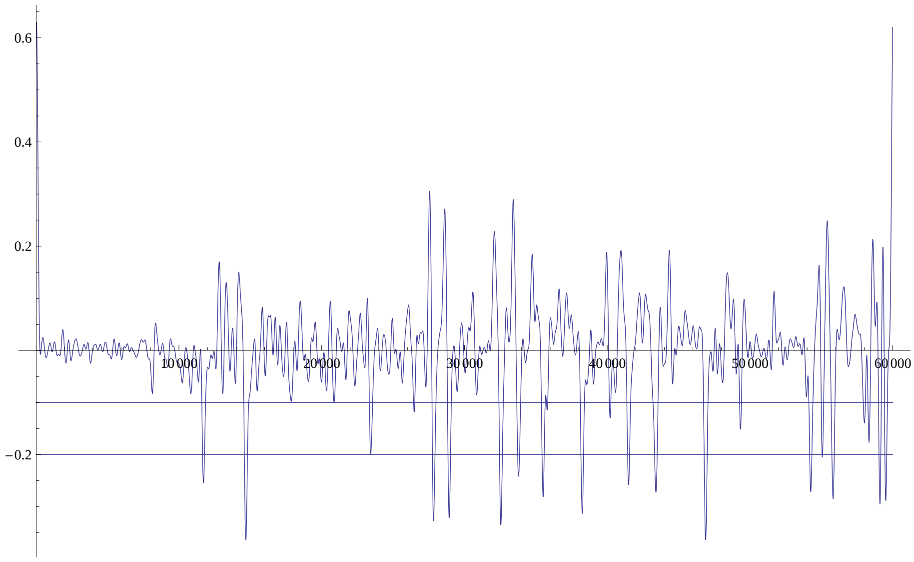
The two models we want to compare are: 1) Simple stochastic waiting times between events, and 2) Periodic occurrence of events with some probability of skipping periods.

Data

Since the first version of this note, I have received the full dataset of the ice core δO_{18}, which is a measure of the annual temperature. It looks as follows for the period ranging from the present to 60.000 years in the past. This includes almost 50.000 years belonging to the last ice age:



We see the famous DO-events, where abrupt changes in temperature occur. In order to find the precise times for these events, I first smooth the curve to iron out the large short time variations, and after that I differentiate the curve. Large negative peaks in the time series then marks the events with strong decreases in temperature, while large positive peaks are the upward jumps in temperature:



Using this procedure, I can identify the following times for large temperature drops: 11740, 14753, 17858, 20830, 23465, 26426, 27788, 28967, 30798, 32569, 33814, 35517, 38202, 40153, 41477, 43442, 46848, 54250, 55821, 59096, all years before present. There are in total 20 events.

Model 1

In this the waiting time, t, between events are taken from a simple Poisson waiting time distribution, with parameter τ_1:

$$p_1(t|T_1\tau_1 I) = \tau_1^{-1} e^{-t/\tau_1}. \tag{3}$$

The likelihood for the data are thus given by

$$P(D|T_1\tau_1 I) = \tau_1^{-n} \exp\left(-\sum_{i=1}^{n} t_i/\tau_1\right), \tag{4}$$

where n are the number of intervals between events in the data. This has a maximum for $\tau_1 = \langle t_i \rangle = \bar{t}$ and we can approximate the function very accurately by

$$P(D|T_1\tau_1 I) \approx \bar{t}^{-n} e^{-n} \exp\left(-n\frac{(\tau_1 - \bar{t})^2}{\bar{t}^2}\right), \tag{5}$$

i.e. a gaussian distribution in τ_1 with average $\bar{t}$ and variance $\sigma_1^2 = \bar{t}^2/(2n)$. Numerically we get using the data, that $\bar{t} = 2492$ years, $\sigma_1 = 404$ years.

The likelihood of the model is easy to calculate. First we have to set a prior for τ, given the model. I see two possibilities. Either τ can be anywhere in the interval from the resolution of the ice core data ($\sim$ 50 years) to the full duration of the ice age ($\sim$ 100.000 years). This is several decades hence the Jeffrey's prior is warranted. The other possibility is that we expect τ to be in some interval of width $\delta\tau$ around some characteristic time scale. I have no prior knowledge of the existence of such a scale, so let us go with the Jeffrey' prior. Finally, the model does not really give the probability of when the first event should occur. I am going to assume that it is occurring with equal probability in an interval of width $\bar{t}$. We then get

$$\begin{aligned}\int P(D|T_1\lambda_1 I)P(\lambda_1|T_1 I)d\lambda_1 &= \frac{1}{\bar{t}}\int_0^\infty \frac{d\tau}{\tau}\tau^{-n}e^{-T/\tau} \\ &= (n-1)!T^{-n}\frac{1}{\bar{t}}, \end{aligned} \tag{6}$$

where T is sum of all intervals, i.e. the distance between the start of the ice age and the last event. This is $T = 47356$ years and $n = 14$. Numerically we get

$$\int P(D|T_1\lambda_1 I)P(\lambda_1|T_1 I)d\lambda_1 = 3.78\times 10^{-77}\ \text{years}^{-20}. \tag{7}$$

Model 2

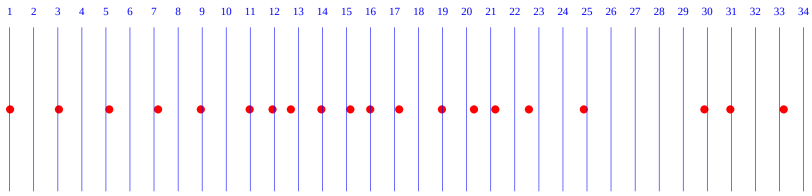
Like Ditlevsen *et al.* the events according to this model is supposed to occur at times

$$t_i = t_0 + n_i\tau_2 + s_i. \tag{8}$$

Here n_i is the period in which the i'th event occurred. The period length is τ_2. Not every period will trigger an event, so we will introduce a parameter, α, which is the probability that a period has *no event*. Finally, s_i is a stochastic offset, which we will take from a gaussian distribution with average 0 and variance σ_2^2. The model thus has 4 parameters, $\lambda_2 = (t_0, \alpha, \tau_2, \sigma_2)$.

Let us first calculate the probability of the data given the model. The function n_i of course will depend on the period τ_2. For models where $\tau \approx 1400 - 1500$ years the data suggest the following sequence of n_i's (1, 3, 5, 7, 9, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 23, 25, 30, 31, 33).

The following figure shows graphically the quality of the fit. The vertical lines are spaced 1470 years, while the red dots show the actual time of the events.



This means that in total 19 periods are without an event. This adds a factor $\alpha^{14}(1-\alpha)^{20}$ to the likelihood. Other choices of τ_2 will off course have other such factors. We now have, that the probability of the data are given by

$$P(D|T_2\lambda_2 I) = \frac{1}{(\sqrt{2\pi}\sigma_2)^n} \exp\left(-\sum_i \frac{(t_i - t_0 - n_i\tau_2)^2}{2\sigma_2^2}\right) \tag{9}$$

This is a gaussian in t_0 and τ_2 with means $\bar{t_0} = 10334$ years and $\bar{\tau_2} = 1468$ years. The maximum likelihood for σ_2 can now be found analytically. With the data we get $\sigma_2 = 257$ years. Likewise the most likely value for α is $19/34 \approx 0.41$.

In order to calculate the likelihood of the model, we need to determine the prior probabilities. The parameters t_0, the period τ_2 and σ_2 are clearly correlated. t_0 need to be earlier than the onset of the ice age, and be in an interval of width τ_2 around that onset. The prior for τ_2 is taken to be Jeffrey's for the same reasons as in model 1. It does not make much sense to have $\sigma_2 > \tau_2$, so we choose the prior for σ_2 to be uniform in an interval of width τ_2. Finally, the prior for α is taken to be $1/\alpha/(1-\alpha)$. This is discussed by Jaynes, and it favors values $\alpha = 0$ or 1.

The σ_2 integral is of the following form

$$\frac{1}{(2\pi)^{n/2}} \int_0^\infty \sigma_2^{-n} \exp\left(-\frac{na^2}{2\sigma_2^2}\right) P(\sigma_2|T_2 I)d\sigma_2 = \frac{1}{2}\left(\frac{2}{na^2}\right)^{(n-1)/2} \frac{\Gamma((n-1)/2)}{\tau_2(2\pi)^{n/2}},$$

where a^2 is a quadratic polynomial in t_0 and τ_2.

This quadratic polynomial is written as

$$a^2 = a_0^2 + (\delta t_0 \delta\tau_2)\begin{pmatrix} 1 & \langle n\rangle \\ \langle n\rangle & \langle n^2\rangle \end{pmatrix}\begin{pmatrix} \delta t_0 \\ \delta\tau_2 \end{pmatrix},$$

where δt_0 and $\delta\tau_2$ are the deviations of these parameters from their most likely, given above, and a_0 is the most likely value for σ_2, i.e. 256. The integrals over t_0 and τ_2 now become

$$\int \frac{d\tau_2}{\tau_2}\int \frac{dt_0}{\tau_2}\frac{1}{a^{(n-1)}} = \frac{2\pi}{n-3}\cdot\frac{1}{\bar{\tau_2}^2 a_0^{n-3}\sqrt{\langle n^2\rangle - \langle n\rangle^2}}. \tag{10}$$

The final contribution to the likelihood of model 2 is the integral over α. Here we have

$$\int_0^1 \alpha^{m-1}(1-\alpha)^{n-1}d\alpha = \frac{(m-1)!(n-1)!}{(n+m-1)!}, \tag{11}$$

where $n = 20$ and $m = 14$. Collecting everything we get the following likelihood for model 2:

$$\begin{aligned} &\int P(D|T_2\lambda_2 I)P(\lambda_2|T_2 I)d\lambda_2 \\ &= \frac{1}{2(n-3)}\left(\frac{2}{n}\right)^{(n-1)/2}\frac{\Gamma\left(\frac{n-1}{2}\right)}{(2\pi)^{n/2-1}}\frac{1}{\sqrt{\langle n^2\rangle - \langle n\rangle^2}}\frac{n!m!}{(n+m+1)!}\frac{1}{a_0^{n-3}\bar{\tau_2}^3} \\ &= 5.86\times 10^{-70}\,\text{years}^{-20}. \end{aligned} \tag{12}$$

Comparison and comments

The ratio of the likelihoods of the two models is thus given by 6.45×10^{-8}. The data thus tell us to change our initial relative odds by a factor of 10^7 in favor of model 2. An earlier analysis, based only on data up to 40.000 before present actually gave a likelihood ratio very close to 1[1], so the new data from 40.000 to 60.000 have changed the conclusion dramatically.

Per Hedegård, 24. april 2008

[1] This is in contrast to the conclusion reached by Ditlevsen et al. They reject model 2, but as far as I am concerned this is wrong, since they use the uncertainty in the time determination of a given event, which is ~ 50 years, as the parameter, σ_2 in the model. The latter is a parameter belonging to a climate model, while the former is a parameter associated with the experimental technique, and hence irrelevant, since it is much smaller than σ_2.