

## PUSHING THE LIMITS OF AMS RADIOCARBON DATING WITH IMPROVED BAYESIAN DATA ANALYSIS

V Palonen<sup>1</sup> • P Tikkanen

Accelerator Laboratory, P O Box 43, FIN-00014 University of Helsinki, Finland.

**ABSTRACT.** We present an improved version of the continuous autoregressive (CAR) model, a Bayesian data analysis model for accelerator mass spectrometry (AMS). Measurement error is taken to be Poisson-distributed, improving the analysis for samples with only a few counts. This, in turn, enables pushing the limit of radiocarbon measurements to lower concentrations. On the computational side, machine drift is described with a vector of parameters, and hence the user can examine the probable shape of the trend. The model is compared to the conventional mean-based (MB) method, with simulated measurements representing a typical run of a modern AMS machine and a run with very old samples. In both comparisons, CAR has better precision, gives much more stable uncertainties, and is slightly more accurate. Finally, some results are given from Helsinki AMS measurements of background sample materials, with natural diamonds among them.

### INTRODUCTION

The novel continuous autoregressive (CAR) model for accelerator mass spectrometry (AMS) data analysis attains slightly better accuracy and more reliable uncertainties than the conventional mean-based (MB) method (Palonen and Tikkanen 2007a,b). The standard level is taken into account using a random walk process that adapts to the drift in the measurement. Correlations between the measurements due to drift, correlations due to the use of the same standard measurements, non-Gaussian probability density functions (PDFs) due to unknown variances, and non-Gaussian PDFs due to normalization with the standards are taken properly into account. The inferred type of the drift determines how the standards are used. Measurements of unknown samples can be used to estimate the machine drift when there is more than 1 measurement per sample, thereby improving the accuracy of the results.

Here, we develop the CAR model further by introducing Poisson-distributed measurement error. This will improve the results for very old samples. Also, in this work the variables representing the standard level are no longer integrated out (marginalized) but instead introduced as additional parameters, enabling one to see the shape of the trend. We present the tests of the model with simulated measurements representing a typical run of a modern AMS machine. We also compare the results of the present model to the result from a conventional mean-based (MB) method for very old samples. Lastly, we describe the CAR analysis of Helsinki AMS measurements of several background sample materials. We limit our discussion to radiocarbon measurements, but we note that the model can also be used for the measurement of other isotopes.

### THE CAR MODEL FOR AMS DATA ANALYSIS

An AMS measurement consists of sequential measurements of isotope ratios for unknown samples and standards. The standards are used to normalize out unknown effects or unpredictable changes in the machine throughput. Because the standards are not measured at the same time as the unknown samples, standard measurements are interpolated between the standards to get an overall standard level, to which the unknown measurements are then normalized.

In this section, we will develop a probabilistic model that relates the unknown isotope concentrations of the samples to the measurement data. We first obtain a predicting probability distribution, called the likelihood, for the measurements, given the concentrations and the throughput. (That is,

<sup>1</sup>Corresponding author. Email: vesa.palonen@helsinki.fi.

given that the concentrations were such-and-such and the standard level changed so-and-so, the probability distribution for the measurements would be such-and-such.) With Bayes' theorem, we will then invert that PDF to obtain the PDF we really want, namely the PDF for the concentrations and the standard level, given the measurements.

Some of the following equations we have introduced previously (Palonen and Tikkanen 2007a,b) but are also included here to help the reader get a complete picture of the model. We will start by introducing most of the symbols needed.

We will represent the measured  $^{14}\text{C}$ -count to  $^{13}\text{C}$ -current ratios with a size  $N$  vector  $\mathbf{R} = (R_1, \dots, R_N)$  measured at times  $\mathbf{T} = (t_1, \dots, t_N)$ . The vector of the ratios  $\mathbf{R}$  has been calculated from the measured  $^{14}\text{C}$  counts,  $\mathbf{c} = (c_1, \dots, c_N)$ , the corresponding  $^{14}\text{C}$  counting times,  $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)$ , and the  $^{13}\text{C}$  currents,  $\mathbf{I} = (I_1, \dots, I_N)$ , by using  $R_i = c_i / (I_i \tau_i)$ . (Note that we could just as well have used the  $^{12}\text{C}$  current instead of the  $^{13}\text{C}$  current.) In the following, the vectors  $\mathbf{I}$  and  $\boldsymbol{\tau}$  will be used as background information and will be dropped from the equations for brevity.

Let also the number of samples in the measurement be  $M$ , and  $\mathbf{O} = (O_1, \dots, O_M)$  be the true unknown fraction modern values for each sample. Inferring the  $\mathbf{O}$  from the data  $\mathbf{R}(\mathbf{c}, \boldsymbol{\tau}, \mathbf{I})$  is then the goal of our analysis. Further, let the function  $n(i)$  give the corresponding sample number for each measurement  $i$ , so we can map each measurement  $i$  to the corresponding fraction modern value  $O_{n(i)}$  of the sample. We will take the uncertainty of the  $^{13}\text{C}$  current to be negligible. The error of  $R_i$  will then result from the  $^{14}\text{C}$  counting statistics, which is a Poisson distribution.

Let  $\mathbf{L} = (L_1, \dots, L_N)$  be the standard level, including a sufficient constant for unit conversion to go from  $^{14}\text{C}$  concentrations to measured ratios. Now, given the standard level and the  $^{14}\text{C}$  concentration of a sample, the expected number of counts in a single measurement of the sample is  $\lambda_i = O_{n(i)} L_i I_i \tau_i$ , which is also the mean of the Poisson distribution for the  $^{14}\text{C}$  counts. On the other hand, the measured number of counts is  $c_i = R_i I_i \tau_i$ . The predicting probability for the measured isotope ratio is, from the Poisson distribution for the  $^{14}\text{C}$  counts,

$$p(R_i | O_{n(i)}, L_i) = I_i \tau_i \frac{e^{-\lambda_i} \lambda_i^{c_i}}{c_i!} \quad (1)$$

The standard level or throughput  $\mathbf{L}$  is assumed to follow a CAR(1) process  $x(t)$  around a mean  $m$ ,

$$L_i = x(t_i) + m \quad (2)$$

The CAR(1) process is a generalization of the discrete-time AR(1) process to the continuous time domain. It is a solution to the stochastic differential equation (Jones 1993; Broemeling and Cook 1997):

$$dx(t) + \alpha x(t) dt = \sigma dW(t) \quad (3)$$

where  $W(t)$  is a continuous random walk process (Wiener process) with  $W(t_1) - W(t_2) \sim N(0, |t_1 - t_2|)$ . The correlation coefficient  $\alpha \in [0, \infty]$  is small for a highly correlated process. The parameter  $\sigma \in [0, \infty]$  describes roughly the variance of the CAR(1) process. For a CAR(1) process, a useful difference equation (Broemeling and Cook 1997) is:

$$x(t_i) = x(t_{i-1}) e^{-\alpha \Delta t_i} + v_i \quad (4)$$

where  $\Delta t_i = t_i - t_{i-1}$ ,  $v_i \sim N(0, \frac{1 - e^{-2\alpha\Delta t_i}}{2\alpha} \sigma^2)$ . From Equation 4, we get the probability density for  $x(t)$ . Denoting  $\mathbf{x} = (x_1, \dots, x_N)$ , we have,

$$p(\mathbf{x}|\alpha, \sigma) = \prod_{i=1}^N \frac{\sqrt{\alpha}}{\sqrt{\pi(1 - e^{-2\alpha\Delta t_i})}\sigma} \exp\left(-\frac{\alpha(x_i - e^{-\alpha\Delta t_i} x_{i-1})^2}{(1 - e^{-2\alpha\Delta t_i})\sigma^2}\right) \tag{5}$$

where we take the limit  $t_0 \rightarrow -\infty$ , resulting in  $x_1 \sim N(0, \sigma^2/(2\alpha))$  for the first data point.

Figure 1 gives the hierarchy of the different parameters in the model.

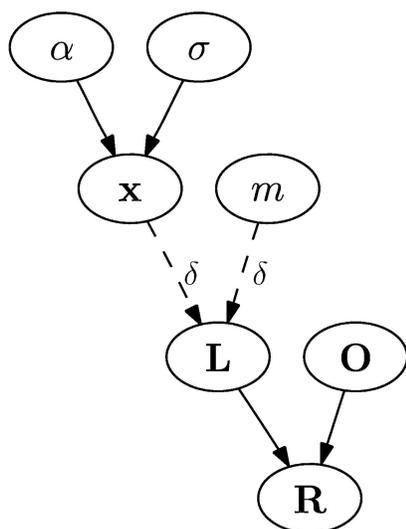


Figure 1 The hierarchy of the different parameters in the model. Solid arrows represent directional probabilistic dependencies between the parameters. A deterministic dependence is represented with a dashed line and the symbol  $\delta$ , which denotes Dirac’s delta function. In general, when writing down the terms in the joint PDF, a parameter at the arrow’s source (parent) is needed as background information in the PDF for the arrow’s target parameter (child) (Jensen 1996).

Table 1 gives the priors used in the model. We are using uniform priors for the most important parameters  $O_i$ . For a normally distributed measurement error, the uniform prior is a Jeffreys’ rule prior for the  $O_i$ , calculated by holding other parameters fixed. Jeffrey’s priors have been noted to perform well in most situations (Kass and Wasserman 1994). For typical AMS measurements, the choice of priors is justified by the accurate results from CAR for simulated data shown in the next section. In general, when the amount of data increases, the contribution from the prior diminishes. The uniform prior for the  $O_i$  values gives zero bias for very old samples, for which the contribution of the prior is most significant. We have also made tests with different priors for the other parameters. For typical AMS measurements, reasonable priors seem to give effectively the same results.

Table 1 Priors for the model parameters.

Parameter	Symbol	Prior type	Prior range
CAR correlation coefficient	$\alpha$	uniform	$[0, \infty]$
CAR standard deviation	$\sigma$	$\sigma^{-1}$	$[10^{-20}, \infty]$
Standard level mean	$\mu$	uniform	$[0, \infty]$
Sample fraction modern	$O_i$	uniform	$[0, \infty]$
Standard fraction modern	$O_i$	$N(F_s, \sigma_s^2)$ <sup>a</sup>	

<sup>a</sup>The values of  $F_s$  and  $\sigma_s$  are known for a standard, e.g.  $F_s = 1.3407$  fraction modern and  $\sigma_s \approx 10^{-4}$  for a OX-II standard when using <sup>13</sup>C currents. To the extent that <sup>14</sup>C calibration has been established using the above value, one might take the limit  $\sigma_s \rightarrow 0$ .

We now have both the probabilistic relations between the entities in Figure 1 (the arrows) and the prior probabilities for the parameters without parents. This enables us to write down the joint PDF for all the parameters and data  $p(\alpha, \sigma, m, \mathbf{O}, \mathbf{x}, \mathbf{R}) = p(\alpha, \sigma, m, \mathbf{O})p(\mathbf{x}|\alpha, \sigma)p(\mathbf{R}|m, \mathbf{O}, \mathbf{x})$  (see e.g. the chain rule for Bayes nets [Jensen 1996]). We can now write down the full posterior, the probability for all the parameters given the data, by using the definition of conditional probability,  $p(A|B) = p(A, B)/p(B)$ , in a way similar to the Bayes' theorem (Sivia 1996; Gelman et al. 2003; Jaynes 2003):

$$\begin{aligned}
 p(\alpha, \sigma, m, \mathbf{O}, \mathbf{x}|\mathbf{R}) &= \frac{p(\alpha, \sigma, m, \mathbf{O}, \mathbf{x}, \mathbf{R})}{p(\mathbf{R})} \\
 &= \frac{p(\alpha, \sigma, m, \mathbf{O})p(\mathbf{x}|\alpha, \sigma)p(\mathbf{R}|m, \mathbf{O}, \mathbf{x})}{p(\mathbf{R})} \\
 &\propto \frac{1}{\sigma} \prod_s \exp\left(-\frac{(O_s - F_s)^2}{2\sigma_s^2}\right) \\
 &\times \prod_{i=1}^N \frac{\sqrt{\alpha}}{\sqrt{(1 - e^{-2\alpha\Delta t_i})}\sigma} \exp\left(-\frac{\alpha(x_i - e^{-\alpha\Delta t_i}x_{i-1})^2}{(1 - e^{-2\alpha\Delta t_i})\sigma^2}\right) \\
 &\times \prod_{j=1}^M e^{-\lambda_j} \lambda_j^{c_j}
 \end{aligned} \tag{6}$$

In the third line, we dropped out the marginalization constant  $p(\mathbf{R})$ . The first product in Equation 6 is the product of the priors of the standard samples (see Table 1).  $F_s$  is the known fraction modern value of the standard and  $\sigma_s$  is the uncertainty of the standard value. This product effectively fixes the value of the CAR process at the points of standard measurements. With the posterior at hand, we are now ready to move on to the more practical things.

## COMPUTATION

Inferences from the posterior are made with an MCMC code, an algorithm that gives parameter space points distributed according to the posterior given in Equation 6 (Gilks et al. 1996). An adaptive Metropolis-Gibbs hybrid sampler was used in this work. The Metropolis guess for most parameters is adaptive Gaussian, the exception being that the guessing-distributions for the trend components are drawn from approximated full conditional distributions (hence the Gibbs part).

The convergence of the MCMC chain to the target PDF was checked both by looking visually at the trace plots and by using the PSRF and MPRSF convergence diagnostics developed by Gelman and Brooks (Gelman and Rubin 1992; Brooks and Gelman 1998). Packages ready for implementing the methods are mcmcdiag (Särkkä and Vehtari 1999) for MATLAB® and BOA (Smith 2005) for R. In this method, multiple independent chains with random starting values are simulated and total posterior variance and the within-chain variance are compared. When their ratio  $\hat{R}$  is near unity, it is taken that convergence has been attained. The common constraint  $\hat{R} < 1.2$  is used as a requirement for convergence. A convergence plot using 4 chains for the analysis of background sample measurements (described below) is shown in Figure 2. With the present code, convergence is usually attained after a burn-in of  $10^8$  MCMC points, taking a couple of hours per chain on a modern 3 GHz PC.

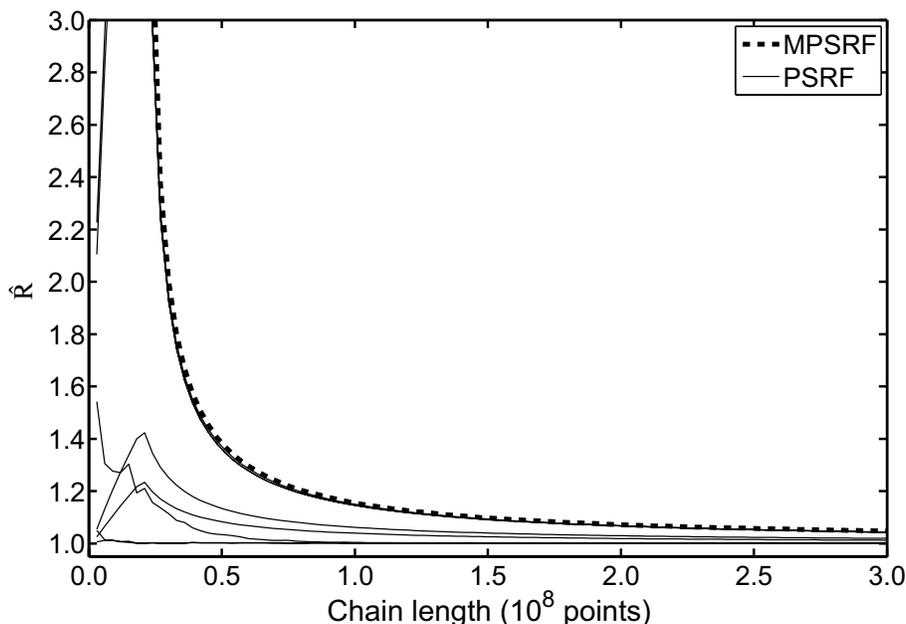


Figure 2 MPSRF (the thick dashed line) and PSRF (thin lines) convergence diagnostics for a typical measurement analysis run. PSRF lines are drawn for the parameters  $\alpha$ ,  $\sigma$ ,  $m$ , and for a couple of  $O_i$  and  $x_i$  values. Convergence is taken to be attained when  $\hat{R} < 1.2$ .

**RESULTS**

**Typical AMS Data**

Measurement data was simulated in order to compare the current Bayesian CAR model with a conventional mean-based (MB) method. In the MB method used here, the measurements are normalized to 4 nearby standards and the final results are the means of the individual normalized measurements for each sample. The standard errors of the means

$$\sigma_{\langle R \rangle} = \sqrt{\frac{1}{n(n-1)} \sum_i (R_i - \langle R \rangle)^2}$$

are used as uncertainties. The standard level at the point of a sample was taken to be a mean of the 4 nearby standard measurements, weighted inversely according to the distance between the measurements. We will also use the MB method with the assumption of a constant standard level (MBCSL method). In this case, the standard level is taken to be the mean of all the standard measurements. Each simulated measurement was drawn from a Poisson distribution whose expectation value was a product of the known <sup>14</sup>C concentration and Gaussian AR<sub>α=1</sub> (random walk) noise representing the drift. Six measurements were simulated per sample and 9 unknown samples were measured once between standard samples. So, every tenth measurement was done on a standard. Altogether, 40 × 6 × 4000 measurements were simulated with MATLAB. Because the analysis time of CAR is long, only 40 × 6 × 30 measurements with CAR were analyzed for each trend strength.

Figure 3 shows the means of true errors of the estimated <sup>14</sup>C concentration as a function of increasing machine drift for both methods. Note that the simulated data represents runs of a modern high-precision AMS machine with the drift varying from zero to being clearly noticeable. Because a sam-

ple was measured only once between standards, the correlations between the measurements were minimized. Had this choice not been made, the performance of the MB model would have been worse (Palonen and Tikkanen 2007a,b). It is seen that for the no-drift case, the CAR model is more accurate than MB and as accurate as MBCSL. With increasing drift, the errors made by MBCSL increase heavily, while CAR continues to give accurate results.

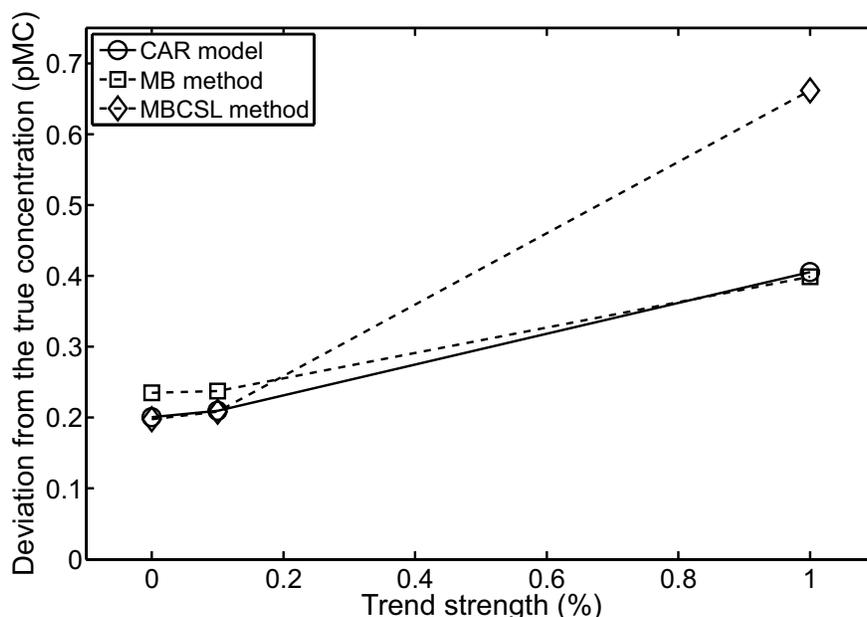


Figure 3 The mean absolute values of the differences between a method's results and the true concentration as a function of increasing machine drift. The graph is drawn for modern samples. The trend strength is the random walk step std of the machine drift relative to the mean of the standard level. The CAR model is more accurate than MB for a typical run of a modern high-precision AMS. MBCSL (where the standard level is assumed to be constant and the same value is used for all measurements) is as accurate as CAR in the case optimal for MBCSL.

### Old Samples

To ascertain the performance of the methods for old samples, no-drift measurements of old samples were simulated and analyzed with both methods. The number of counts were drawn from Poisson statistics with  $\lambda$ , the mean number of counts for 1 measurement, being 0.1, 0.5, 1.0, 1.5, 2.0, 3.5, 5.0, or 10 for "unknown" samples. The standards had  $\lambda = 18,000$ . Again, 6 measurements were made per sample. Note that in this case, the counting statistical uncertainty will be so large that many of the previously noted advantages of CAR will not show up.

As an example of CAR results for very old samples, Figure 4 shows 4 PDFs for the parameter  $O_m$ , which represents the  $^{14}\text{C}$  concentration. Due to Poisson statistics, the PDFs are tailed for measurements with only a few total counts. The samples of the solid PDFs represent the rather extreme case where the total count average  $\lambda$  was 0.6. Here, the solid PDFs represent samples with 1 count and zero counts. Despite the difference in the amount of counts, both PDFs give a reasonably high value for the true value  $5.6 \times 10^{-6}$  fraction modern. This is to be expected because, by taking into account the  $^{14}\text{C}$  measurement time, the  $^{13}\text{C}$  current, and the underlying Poisson statistics, the CAR model is able to give a reliable PDF for the true  $^{14}\text{C}$  concentration in spite of the randomness in the measurement result.

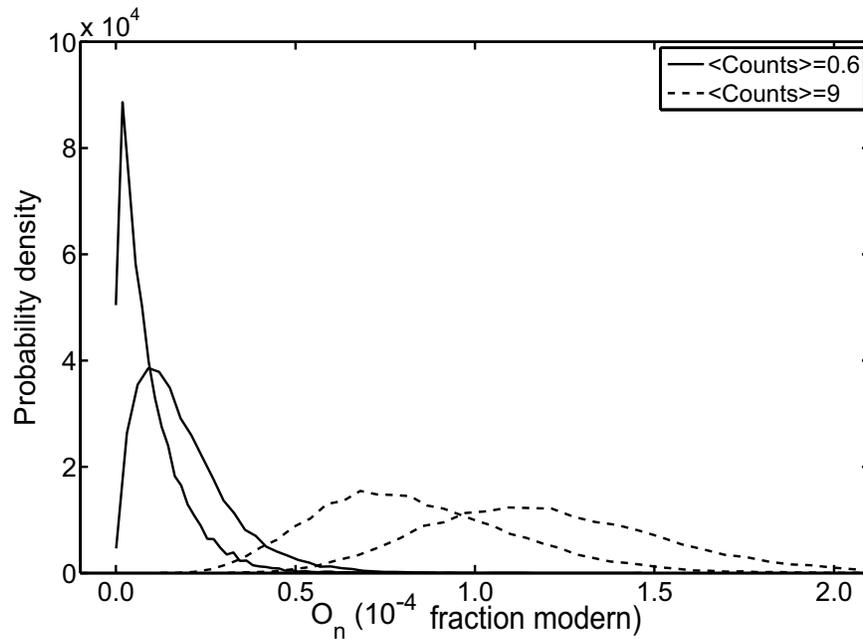


Figure 4 PDFs for the fraction modern values  $O_n$  of 4 very old samples from a CAR model analysis of a simulated measurement. Two samples were selected from 2 (the count average  $\lambda = 0.6$  and  $\lambda = 9.0$ ) groups each. The 2 groups correspond to the fraction modern values of  $6 \times 10^{-6}$  and  $8 \times 10^{-5}$ . Due to the underlying Poisson statistics, the distributions are tailed.

Figure 5 shows the means of relative deviations between each method’s results and the known true values as a function of the mean of the total counts from the sample. The location of highest posterior density was used as a point estimate for CAR. MBCSL is not included because the uncertainty from of the counting statistics is much larger than the error from the handling of the standard level. Based on point estimates, the CAR model may be just slightly more accurate for very old samples. Contrary to our preliminary claim (Palonen and Tikkanen 2007b) (which was the unfortunate result of a coding error and was noted upon further verification of the results), the MB method does not suffer from a significant bias for very old samples. (Note that MB has some bias when the result is converted to an age for very old samples. This is because the change of variable is not done for the whole probability distribution.)

Figure 6 gives the uncertainties from the methods relative to the (normally approximated) counting statistical uncertainty. The uncertainties from CAR are given as 68.3% highest posterior density regions and the uncertainties from MB as  $1-\sigma$  intervals. For very old samples, CAR gives uncertainties slightly above the counting statistical uncertainties, approaching the counting statistical uncertainty. MB gives erroneously low uncertainties, with the mean of the uncertainties below the counting statistical uncertainty. (The reason for the underestimation of the uncertainty is that while the sampling variance is “unbiased,” the corresponding sampling standard deviation is not.) Note also that the uncertainty estimates from MB are much more unstable (the light gray area) than the uncertainties given by CAR (the dark gray area). The difference in stability comes from the fact that CAR uses the information about the underlying counting statistics in addition to the sampling variance, resulting in stable uncertainties, whereas the MB method’s standard error of the mean from 6 data points has random scatter. It is noted that the difference in the stability of the uncertainty estimate is larger when more modern samples are measured or when fewer measurements are made per sample.

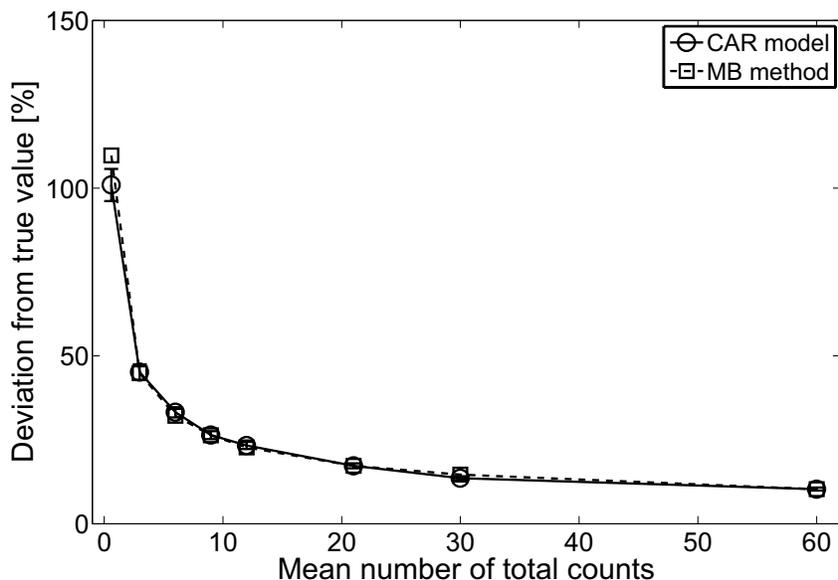


Figure 5 The mean absolute values of the differences between a method's results and the true concentration as a function of total counts from the sample. The differences are given in percent of the true concentration. The mean number of counts in the simulated measurements was 0.1, 0.5, 1.0, 1.5, 2.0, 3.5, 5.0, or 10, with 6 measurements for each sample.

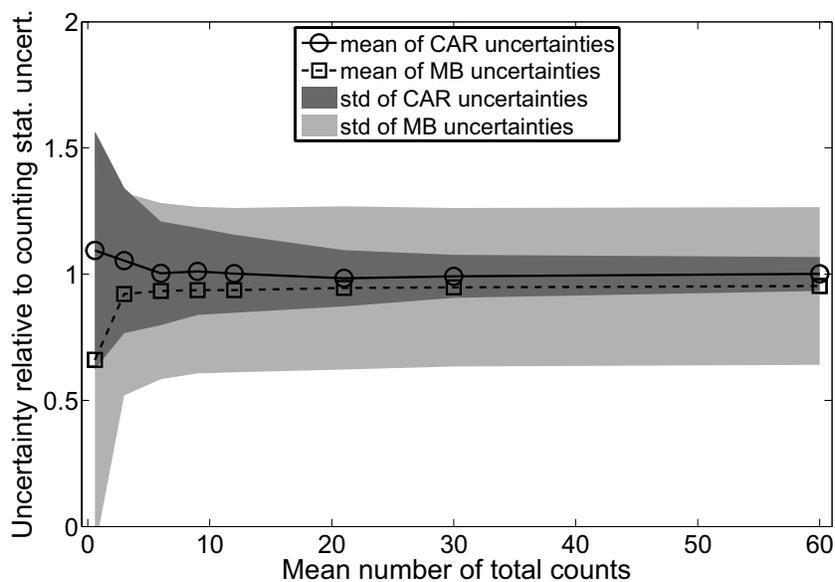


Figure 6 Ratio of the uncertainty estimate (for the  $^{14}\text{C}$  concentration of the sample) given by the method to the (normally approximated) counting statistical uncertainty as a function of the mean number of total counts. Overall, CAR gives more reliable uncertainties. Especially for the very old samples, MB gives erroneously low uncertainties. Importantly, the random scatter in the uncertainties is much lower for CAR (dark gray area) compared to MB (light gray area). The scatter was calculated as the sample standard deviation of the uncertainties given by each method.

Figure 7 shows the relative amount of results that do not contain the true value for both methods. The expected fraction of true values outside the given uncertainties is 31.7% for both methods. CAR delivers reliable uncertainties with roughly the ideal 31.7% fraction falling outside the given uncertainties, except for the samples with very low amount of counts, for which CAR uncertainties are “too safe.”

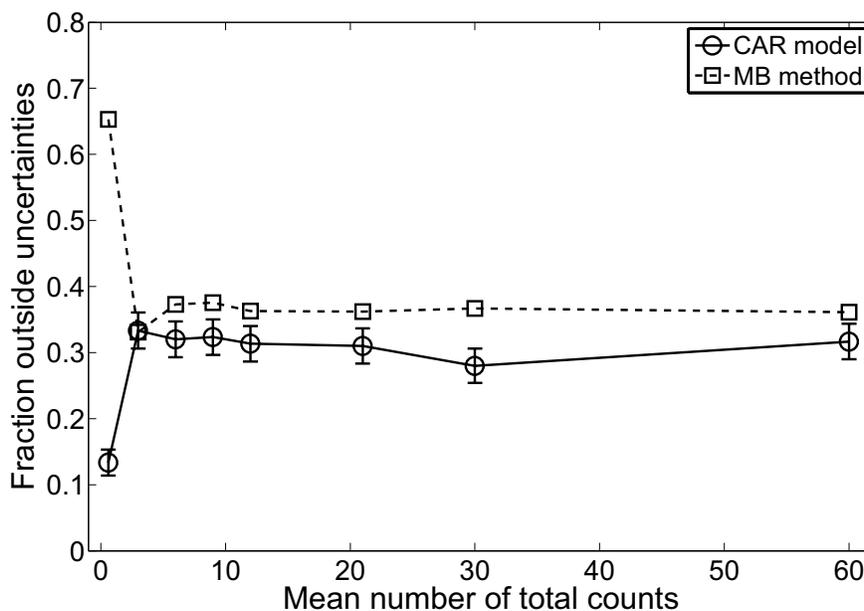


Figure 7 Fraction of true values that are outside the uncertainties given by the methods. A 68.3% highest posterior density region is used for CAR and a 1- $\sigma$  interval is used for MB. (Uncertainties for CAR were calculated from the corresponding binomial distribution.)

### Analysis of Actual AMS Measurements

About a dozen background samples were prepared for AMS measurements as preliminary tests for both the sample treatment background and the AMS machine background in the Helsinki AMS facility (Palonen et al. 2004; Tikkanen et al. 2004). Several materials were measured without sample preparation by simply pressing the material (e.g. diamond, graphite, or anthracite) to the sample holder. To see the effect of sample preparation, some graphite and anthracite samples were put through different sample treatment steps. Figure 8 gives the inferred probability distributions for the  $^{14}\text{C}$  concentrations of the measured samples. Sample materials and pretreatments done are given in the figure. Figure 9 gives the corresponding probability distributions for the  $^{14}\text{C}$  ages. Table 2 gives the points of maximum posterior density and the 68.3% central posterior interval for the  $^{14}\text{C}$  concentrations and ages for each sample. During the measurements, our standard MC-SNICS ion source had a maximum  $^{12}\text{C}$  current of 40  $\mu\text{A}$ . For the diamonds, the current was roughly 20  $\mu\text{A}$ .

Because the samples could have been measured longer and since at present we are losing some of the carbon during the slow switching of the injector magnet's field, even lower concentrations are expected to be measurable. If one takes the lowest measured concentration as background level, it should be possible to attain  $^{14}\text{C}$  ages up to 80 kyr BP for some samples. The background sample  $^{14}\text{C}$  concentration is of course also given as a PDF and the subtraction of the background can be done probabilistically, taking both uncertainties into account correctly.

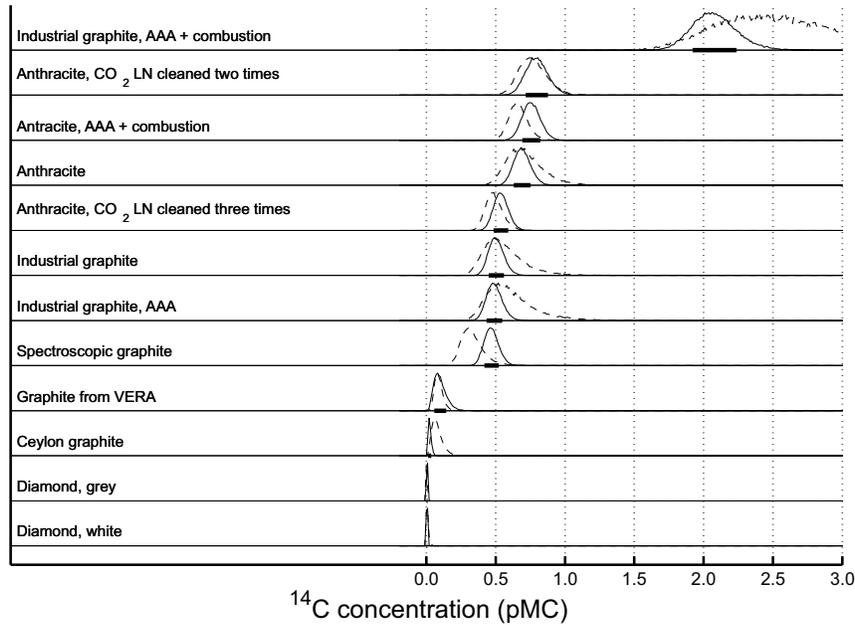


Figure 8 Probability density distributions (PDFs) of the  $^{14}\text{C}$  concentrations for the different background samples measured with the Helsinki AMS. The dashed line represents short measurements done in April 2006 and solid lines represent longer measurements done in November 2006 for the same sample wheel. The thick horizontal solid lines give the 68.3% central posterior intervals for each PDF. Sample pretreatments were done in the Dating Laboratory, University of Helsinki.

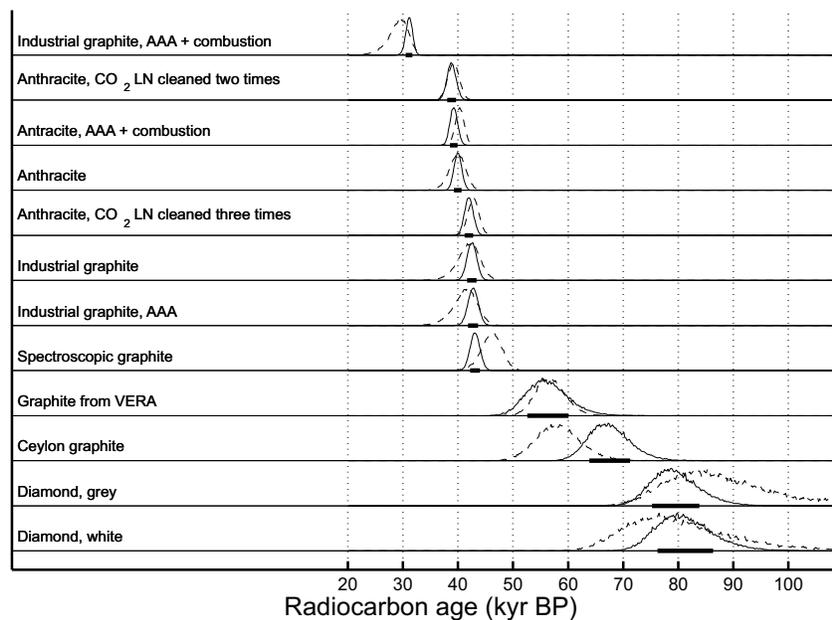


Figure 9 As for Figure 8, but the PDFs here are for conventional  $^{14}\text{C}$  ages for the different background materials.

Table 2 Measured  $^{14}\text{C}$  concentrations and corresponding  $^{14}\text{C}$  ages for different background materials. The value given is the point of maximum posterior density, and the limits represent the 68.3% central posterior interval.

Sample	$^{14}\text{C}$ concentration (pMC)	$^{14}\text{C}$ age (kyr BP)
Industrial graphite, AAA + combustion <sup>a</sup>	2.04 +0.19/−0.13	31,180 +560/−670
Anthracite, CO <sub>2</sub> LN cleaned twice	0.801 +0.076/−0.085	38,740 +910/−710
Anthracite, AAA + combustion	0.743 +0.079/−0.051	39,220 +710/−670
Anthracite	0.684 +0.067/−0.054	40,010 +660/−750
Anthracite, CO <sub>2</sub> LN cleaned 3 times	0.528 +0.063/−0.043	42,030 +750/−840
Industrial graphite	0.492 +0.066/−0.043	42,630 +760/−990
Industrial graphite, AAA	0.481 +0.066/−0.047	42,820 +840/−1000
Spectroscopic graphite	0.461 +0.061/−0.044	42,990 +990/−790
Graphite from VERA	0.078 +0.064/−0.023	55,200 +4700/−2700
Ceylon graphite	0.020 +0.014/−0.007	67,000 +4200/−3200
Diamond, gray	0.0041 +0.0043/−0.0013	78,600 +5100/−3400
Diamond, white	0.0028 +0.0046/−0.0008	79,500 +6700/−3400

<sup>a</sup>Combusted with an experimental high-temperature setup.

## CONCLUSIONS

The continuous autoregressive (CAR) model was improved by representing the AMS machine drift with hidden variables instead of assuming a normal distribution for the measurement error and marginalizing over the trend. The use of hidden variables enables the use of more realistic Poisson distribution for the measurement error.

With simulated measurement data, it has been shown that the CAR model gives slightly more accurate results compared to the mean-based (MB) method even for typical measurements from modern AMS machines and measurement sequences favorable to MB. MB with the assumption of a constant standard level is as accurate as CAR when the measurements have no machine drift, enough  $^{14}\text{C}$  counts, and no correlations due to the use of the same standard measurements. While CAR uses the data to learn how to handle the standards, MB would seem to need a tool rather similar to CAR to ascertain for certain when the assumption of a constant standard level can safely be made.

The inclusion of the Poisson-distributed measurement error improves the performance of AMS data analysis for very old or very small samples. It seems that when the number of total counts from a sample is very small, the conventional MB method has unreliable uncertainties. MB uncertainties also seem to be somewhat unstable even for more modern samples. In contrast, due to the correct handling of the counting statistics, CAR is able to give reliable and stable uncertainties even for the samples with very few  $^{14}\text{C}$  counts.

Analyses of real measurements of background samples indicate that  $^{14}\text{C}$  ages as old as 80 kyr BP can be measured for some samples. Of course, in many cases significant improvement of sample treatment is also needed in order to attain the low intrinsic  $^{14}\text{C}$  levels.

## REFERENCES

- Broemeling LD, Cook P. 1997. A Bayesian analysis of regression models with continuous errors with application to longitudinal studies. *Statistics in Medicine* 16:(4):321–2.
- Brooks SP, Gelman A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7(4):434–55.
- Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7(4):457–511.
- Gelman A, Carlin JB, Stern HS, Rubin DB. 2003. *Bayesian Data Analysis*. 2nd edition. Boca Raton: Chapman & Hall/CRC. 668 p.
- Gilks WR, Richardson S, Spiegelhalter DJ, editors. 1996. *Markov Chain Monte Carlo in Practice*. Boca Raton: Chapman & Hall/CRC.
- Jaynes ET. 2003. *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press. 758 p.
- Jensen FV. 1996. *An Introduction to Bayesian Networks*. London: UCL Press. 178 p.
- Jones RH. 1993. *Longitudinal Data with Serial Correlations: A State-Space Approach*. London: Chapman & Hall/CRC. 240 p.
- Kass RE, Wasserman L. 1994. Formal rules for selecting prior distributions: a review and annotated bibliography. Technical report #583, Carnegie Mellon University, Pennsylvania, USA.
- Palonen V, Tikkanen P. 2007a. An information-efficient Bayesian model for AMS data analysis. *Radiocarbon* 49(2):369–77.
- Palonen V, Tikkanen P. 2007b. A shot at a Bayesian model for data analysis in AMS measurements. *Nuclear Instruments and Methods in Physics Research B* 259(1):154–7.
- Palonen V, Tikkanen P, Keinonen J. 2004. Ion-optical modelling of the Helsinki AMS tandem. *Nuclear Instruments and Methods in Physics Research B* 223–224:227–32.
- Särkkä S, Vehtari A. 1999. MCMC diagnostics toolbox for MATLAB [software]. URL: <http://www.lce.hut.fi/research/mm/mcmcdiag/>.
- Sivia DS. 1996. *Data Analysis: A Bayesian Tutorial*. Oxford: Clarendon Press. 240 p.
- Smith BJ. 2005. Bayesian Output Analysis (BOA) program, version 1.1.5 [software]. URL: <http://www.public-health.uiowa.edu/boa>.
- Tikkanen P, Palonen V, Jungner H, Keinonen J. 2004. AMS facility at the University of Helsinki. *Nuclear Instruments and Methods in Physics Research B* 223–224:35–9.