

RESEARCH ARTICLE

Reevaluation of radiocarbon counting statistics on the MICADAS AMS system: Evidence and implications of non-Poisson distributions on robust uncertainty calculation

Gary Salazar  and Leonard I. Wassenaar

André E. Lalonde AMS Laboratory, University of Ottawa, 25 Templeton Street, Ottawa, Ontario, K1N 6N5, Canada
Corresponding author: Gary Salazar; Email: gsalazar@uottawa.ca

Received: 16 September 2024; **Revised:** 21 May 2025; **Accepted:** 01 June 2025

Keywords: counting statistics; internal error modeling; MICADAS; non-Poisson distribution; quasi-Poisson models; radiocarbon AMS

Abstract

A reassessment of radiocarbon counting statistics in accelerator mass spectrometry (AMS) at the Andre E. Lalonde National Facility revealed that the traditionally assumed Poisson distribution may not always apply. An extensive analysis of 2.5 years of ^{14}C and ^{12}C data was conducted on a MICADASTM AMS. This study found that only 63% of results adhered to Poisson statistics, while 34.2% showed slight deviations, and 2.8% exhibited strong non-Poisson behavior. This finding challenges the classic assumption that radiocarbon AMS is inherently a Poisson process. This study recommends considering non-Poisson models, specifically quasi-Poisson and negative binomial models, to better account for internal error and improve the accuracy of the reported error. Integrating ^{12}C current noise into error calculations is also suggested as it plays a significant role in measurement variability. We would like to ignite curiosity on other AMS laboratories to test the non-Poisson error framework with the broader aim of assessing its applicability in improving conventional statistical models, error expansion methods, and in ensuring more accurate and reliable ^{14}C results.

Introduction

The Poisson distribution is a fundamental principle for estimating the variance of ^{14}C count rates in radiocarbon accelerator mass spectrometry (Burr et al. 2007; Stuiver and Becker 1986; Stuiver and Polach 1977). The Poisson probability density function (Eq. 1) computes the probability of counting a certain number of ^{14}C events during a specific time lapse (Pinsky and Karlin 2011) or cycle time. The number of counts per cycle is the same as count rate or counts per block (N_i) where i is the cycle or block. The term μ_i is the expected counts per cycle and can be approximated to the average count rate. AMS has long used Poisson uncertainty estimation from classical radioactivity-decay counting methods (Povinec et al. 2009).

$$\text{Pois}(N; \mu) = \frac{\mu^N e^{-\mu}}{N!} \quad (1)$$

Typically, an AMS radiocarbon measurement is conducted by counting during s number of cycles (e.g. $s = 15$ cycles). A group of cycles is called a *run* or *pass* (of the sample target wheel) and the entire sample measurement consists of n passes. In statistical terms, the counts per cycle are observations drawn from a population of unknown distribution, and the size s is the size of the statistical sample. If μ is constant across the passes or stationary, then μ can be estimated from the global mean of counts per cycle defined as

$$\mu_{global} = N_T / (n \cdot s) \quad (2)$$

where N_T is the total counts for the entire measurement. If the Poisson model is applicable, then the absolute scatter (uncertainty) of the distribution of the whole set of count rates is given by

$$\sigma_{pois} = \sqrt{\mu_{global}} \quad (3)$$

If the mean count rate μ is not constant and behaves as a moving average across the passes of the measurement, then the Poisson process is deemed non-stationary. The uncertainty of non-stationary Poisson processes cannot be calculated with Eq. 3, but with $\sigma_{pois} = \sqrt{\mu_i}$, because the uncertainty changes over time. Therefore, a model for μ_i must be estimated to capture the trend of the counts per cycle (N_i). The non-stationary data can then be transformed into stationary data around μ_{global} by detrending with an appropriate model μ_i (Poplová et al. 2017). Stationary and non-stationary processes fail to meet a Poisson distribution when the measured scatter is higher than $\sqrt{\mu_i}$. This situation is called overdispersion. In short, counting statistics is the proposed model that dynamically describes the uncertainty at a certain time point along the measurement is being executed.

Some papers discussing the counting of radioactive decay observed unusual overdispersion attributed to unexplained events occurring during the counting process. Consequently, several alternative models have been suggested (Currie et al. 1998). In the case of AMS, the ion source physically destroys the graphite target during counting (Litherland et al. 1987). Therefore, owing to sample destruction, AMS assays should be prone to events that lead to a non-Poisson counting distribution. Nevertheless, most AMS studies do not report deviations from Poisson statistics.

In the field of radiocarbon by AMS, the Poisson processes have been investigated by measuring the time between the arrival of each ^{14}C particle at the detector, known as the interarrival time (Vogel et al. 2004). Vogel et al. (2004) showed a limited set of experiments where the interarrival time distribution corresponded to Poisson processes. They measured only a few ($n = 5$) samples with different $^{14}\text{C}/^{12}\text{C}$ ratios ($^{14,12}R$) and three measurements at different cathode potentials. However, we believe there is still insufficient evidence to assume that radiocarbon AMS counting follows a Poisson distribution. Rather, radiocarbon Poisson distribution should be directly evaluated from the observed count rates, considering the stability and conditions of the AMS signal.

From the point of view of the actual spread of the measurand of interest, $^{14}\text{C}/^{12}\text{C}$ ratio ($^{14,12}R$), the relative standard error of the whole measurement (Eq. 4) is calculated from the standard deviation of the passes means (σ_{means}) factored by \sqrt{n} (Malonda and Carles 2020) and such sampling means should be normally scattered around the global mean $\langle^{14,12}R\rangle_{global}$ as stated by the Central Limit Theorem (CLT) (Evans and Rosenthal 2004; Malonda and Carles 2020). The factor \sqrt{n} comes from the error propagation of the uncertainty of n statistical samples. In Radiocarbon, Eq. 4 is the “external error” (McNichol et al. 2001), and in this paper, it is denoted as the passes-based external error.

$$SE_{\langle X \rangle, rel} = \frac{\sigma_{X means}}{\langle X \rangle_{global} \sqrt{n}} = \frac{\sigma_{R means}}{\langle^{14,12}R\rangle_{global} \sqrt{n}} \quad (4)$$

The proposed and expected relative standard error of the measurement should be calculated from the Poisson uncertainty (σ_{pois}) of the distribution of the entire set of ^{14}C count rates. For the count rates, the distribution of the means is usually not performed. However, another property of the Central Limit Theorem (Evans and Rosenthal 2004) tells that the σ_{means} for the count rates can be approximated by the dispersion of the population factored by the sampling size (σ_{pois}/\sqrt{s}). The approximation of σ_{pois}/\sqrt{s} and definitions of μ_{global} from Eqs. 2 and 3 were inserted into Eq. 5 to calculate the widely accepted Poisson relative standard error (“internal” error). We were unable to find any reference for this demonstration except in Vogel et al. (2004) who used an abridged approach. We assume that the inherited Poisson counting statistics have been so prevalent over time in radiocarbon AMS because of the simplicity of Eq. 5 to quickly predict the error of the measurement.

$$SE_{^{14}\text{C}, \text{rel}} = \frac{\sigma_{^{14}\text{C means}}}{\langle X \rangle \sqrt{n}} = \frac{\frac{\sigma_{\text{pois}}}{\sqrt{s}}}{\mu_{\text{global}} \sqrt{n}} = \frac{\sqrt{\mu_{\text{global}}}}{\mu_{\text{global}} \sqrt{ns}} = \sqrt{\frac{1}{\mu_{\text{global}} \cdot ns}} = \frac{1}{\sqrt{N_T}} \quad (5)$$

Aerts-Bijma et al. (2021) clearly defines both errors: “The internal errors are the ^{14}C uncertainties, as we calculate them along with the measurands, whereas the external errors are the uncertainties observable from the spread in measurands.” The internal error is the expected uncertainty predicted from counting statistics while the external error is the actual obtained uncertainty of $^{14,12}\text{R}$. It is important to note that internal and external errors are both in the format of standard errors. It has been argued in the literature that when the internal and external errors are about the same then the measurement is dominated by counting statistics (Bonani et al. 1987; Tumey et al. 2004; Vogel et al. 2004). However, these two errors should follow a chi squared distribution and sometimes they differ within this chi-squared framework. In short, the internal error (Eq. 5) is obtained by assuming a stationary Poisson counting model $\mu = \mu_{\text{global}}$ and by assuming the CLT is applicable.

The non-Poisson distribution models we employed are the quasi-Poisson (*qp*) and the negative binomial (*NB* type 1) models, which account for overdispersion. These models are different, as reflected by their absolute counting uncertainty shown in Eqs. 6 and 7 (Dupuy 2018; Greene 2008; Tjur 1998). For simplicity, the general form of the non-Poisson absolute uncertainty can be seen as the product of a dispersion parameter (*D*) with the root of the overall mean counts per cycle ($D\sqrt{\mu_{\text{global}}}$).

$$\sigma_{qp} = \sqrt{\varphi_{qp}} \sqrt{\mu_{\text{global}}} \quad (6)$$

$$\sigma_{NB} = \sqrt{(1 + \varphi_{NB1})} \sqrt{\mu_{\text{global}}} \quad (7)$$

For the *NB* model, the stationary data is fitted by iteratively maximizing the likelihood where the parameter ϕ_{NB1} is optimized to the stationary data and the mean is assumed constant. The *NB* model is a Poisson model where μ_i has a Gamma distribution with scale parameter ϕ_{NB1} . While the *qp* model has a dispersion calculated from the Pearson χ^2 and the degrees of freedom. The dispersion *D*, which is defined here as higher or equal to 1 and decided by $\sqrt{\varphi_{qp}}$ or $\sqrt{(1 + \varphi_{NB1})}$, quantifies how many times wider the actual distribution (histogram of detrended count rates affected only by a random process) would be relative to a Poisson distribution. Similarly to Eq. 5, the relative standard error for the models with overdispersion can be written as Eq. 8, which is a multiplied version of the Poisson error. Thus, the non-Poisson model is versatile and supports Poisson or non-Poisson conditions depending on the calculated *D*.

$$SE_{\text{non-pois}, \text{rel}} = \frac{D}{\sqrt{N_T}} \quad (8)$$

The dispersion *D* is exclusively related to the counting statistics which is the most basic error level, and the counting error must be propagated with other uncertainty components related to $\delta^{13}\text{C}$ correction, standard and blank corrections to calculate the next level of error, the quoted error (Aerts-Bijma et al. 2021). Several laboratories have adopted the concept of error multiplier (Scott et al. 2007) to capture sources of variation that appear at the replication level and are not accounted for in the quoted error. The non-Poisson error and the error multiplier are difficult to compare, however, their magnitudes are indirectly related. The quoted errors can be further expanded by empirical functions comparing the replicated measurements to infer error multipliers or other top-down parameters (Salazar et al. 2021). Examples of error expansion applications are: intercomparison studies that evaluate the quoted uncertainty and the error multiplier. Calibration curves and age modeling use Poisson-calculated expanded quoted errors (e.g. additive expansion) during the model construction (Blaauw et al. 2024; Heaton et al. 2020; Scott et al. 2018). Thus, conceptually, non-Poisson quoted errors may create differences in the expansion parameters, quoted errors, and inferences from intercomparison studies and radiocarbon models.

The aim of this paper is to demonstrate that radiocarbon AMS counting uncertainty is not necessarily Poisson, and to accordingly show the advantage of using non-Poisson counting to comparably report internal and external errors. We furthermore consider if the uncertainty or noise of the ^{12}C current coming from instrumental instability should be accounted for in the internal error. Our objective is to develop a non-Poisson framework for radiocarbon AMS uncertainty estimation. A time-dependent model J_i for the ^{12}C current was inferred, and then a time-dependent model μ_i was inferred from J_i . The count rates were detrended to obtain stationary counting data distributed around μ_{global} to finally measure the non-Poisson overdispersion with Eq. 8. The core calculations were not novel but well-established tools in the field of statistics, and an abridged code can be found in GitHub (github/oPAC). Simulations of ^{14}C counting were carried out to validate our equations and to explain the empirical results. We hypothesized that non-Poisson counting statistics applied to our ^{14}C counts would perform better than conventional Poisson statistics by comparing their respective internal errors relative to the external errors. This reassessment gives a more robust error calculation framework related to the AMS instrument itself during the days of measurement in a basic level of error estimation. The broader motivation is to show the non-Poisson error can affect the traditional quoted error and thus, as explained previously, the potential to affect parameters of existing calibration curves, age modeling, and interlaboratory studies.

Materials and methods

MICADAS AMS instrument

The MICADAS main ion optics were tuned before every measurement, but the range of running conditions did not change considerably over 2.5 years of data acquisition. The main instrumental running conditions ranges were: source cathode potential 7.146 kV (fixed); source ion potential 30.32 kV (fixed); source ionizer heating 122 ± 10 W depending on the age of the ionizer and conditions of the ion source, and Cs reservoir $120\text{--}130^\circ\text{C}$ to keep the HE ^{12}C beam current between $40\text{--}60\ \mu\text{A}$; extractor potential -0.87 ± 0.57 kV; box lens -15 ± 1.9 kV; and accelerator 184.86 kV (fixed). The measurements mainly were 14 passes, with a few measurements using 13 or 15 passes. The cycles were fixed at 20 seconds and 15 cycles per pass.

Data query and preprocessing

Two computer scripts written in the **R** language (R Development Core Team 2013) automatically queried or simulated the variables corresponding to each measurement obtained with the MICADAS AMS. The variables were: ^{12}C current at the high energy side (I_i), ^{14}C counts per cycle (N_i), $^{14,12}\text{R}$, and $^{13,12}\text{R}$ ratios. The **R** script analysed these variables at the cycle and pass levels. In our system, the $^{14,12}\text{R}$ accuracy degrades for C masses lower than 0.4 mg when using conventional high-mass standards. Thus, samples with C mass lower than 0.5 mg were rejected to ensure that the mass was not a variable that affected the value and uncertainty of the ratio and counts. The models μ_i and J_i were inferred from the analysis of N_i , I_i at the cycle level, which ultimately were used to calculate the internal error using the distribution models quasi-Poisson and negative binomial type II. A more detailed description of the calculations can be found in the online Supplementary materials.

Our shortest counting time unit was a 20-second cycle length. The count rate was the total ^{14}C particle counts corresponding to each cycle, while the ^{12}C current was the mean. During each pass, the cycles are repeated s times (e.g. $s = 15$) and then the graphite target is changed to the next sample in the sequence. The passes were repeated n times (e.g. $n = 14$). Thus, a typical ^{14}C AMS measurement had 210 cycles, and the datasets (N_i , I_i) each had 210 entries.

Determination of J from the ^{12}C current and μ from the current and count rates

Because the passes of each graphite target were measured periodically, the target heated up during sputtering and cooled down during idling. Consequently, the trends of the ^{12}C and ^{14}C signals were not continuous relative to their neighbor passes, as shown in Figure 1a, b. Therefore, individual models for J_i and μ_i were inferred for each pass. The J_i model was a cubic smooth spline calculated on the ^{12}C current data (I_i). J_i was proportionally scaled and converted into a count model to obtain a first estimate for μ_i using Eq. 9. In Eq. 9, J_i is scaled by the mean of the counts per cycle and a scalar A divided by the mean ^{12}C current of the pass. Each pass had its corresponding μ_i and J_i models.

$$\mu_i = A \cdot J_i \cdot \frac{\langle N \rangle_{\text{pass}}}{\langle I \rangle_{\text{pass}}} \quad (9)$$

$$\mu_i = B \cdot J_i \cdot {}^{14,12}R \quad (10)$$

If the ${}^{14,12}R$ value was known, then, theoretically, the ^{14}C count rate μ_i could be calculated with Eq. 10, which includes the constant B , which incorporates beam transmission, electron charge, detector efficiency, and live counting time. Eq. 10 justified our selection of Eq. 9 to model the count rate from the data. Eq. 10 was employed during the simulations to convert the ratio and current into the count rate model.

The goodness-of-fit for μ_i was evaluated with the Akaike Information Criterion (AIC) calculated with Eq. 11 where P is the number of parameters and $\mathcal{L}_{\text{pois}}$ is the Poisson loglikelihood calculated with Eq. 12. The AIC was minimized by optimizing the proportional constant A and by varying the smoothness parameter of the J_i cubic spline which was scaled to infer μ_i . Graphically, the models can be seen in the insets of Figure 1a, b corresponding to the first pass.

$$\text{AIC} = 2P - 2\mathcal{L}_{\text{pois}} \quad (11)$$

$$\mathcal{L}_{\text{pois}} = \sum \ln[\text{Pois}(N_i; \mu_i)] \quad (12)$$

Transforming a non-stationary model into a stationary model for counting statistics

After determining the optimum J_i and μ_i models for each pass, the counts per cycle (N_i) were transformed into a stationary model. As mentioned above Poplová et al. (2017) did this by fitting a trend model of their count rates, which is akin to our μ_i . Using a Z-score, they could detrend the original signal and found the stationary transformation which is the set of scattered counts (N_i') relative to a constant global mean (μ_{global}). The Z-score for a Poisson process is

$$[N_i - \mu_i] / \sqrt{\mu_i} = [N_i' - \mu_{\text{global}}] / \sqrt{\mu_{\text{global}}} \quad (13)$$

The numerator represents the detrending of the raw signal and the denominator normalizes it relative to the Poisson uncertainty. The left side of Eq. 13 uses the non-stationary model, and the right side is the stationary model. In our case, instead of the Z-score, our transformation (Eq. 14) was based on the Pearson's χ -squared: $\chi^2 = [N_i - \mu_i]^2 / \mu_i$ which is the square of the Z-score. The solution of Eq. 14 gave the transformed counts per cycle as shown in Eq. 15 and Figure 1c where N_i' is an integer and the root sign is the sign of $N_i - \mu_i$.

$$[N_i - \mu_i]^2 / \mu_i = [N_i' - \mu_{\text{global}}]^2 / \mu_{\text{global}} \quad (14)$$

$$N_i' = \left[\mu_{\text{global}} \mp \sqrt{\mu_{\text{global}} \cdot [N_i - \mu_i]^2 / \mu_i} \right] \quad (15)$$

Once the counts per cycle of a pass were transformed then the process was applied to the next pass. A one-time Grubb's outlier detection at a confidence level $\alpha=0.05$ was applied to the transformed counts (Alrawashdeh 2021) to ensure outliers were eliminated to avoid instabilities in the fitting calculations.

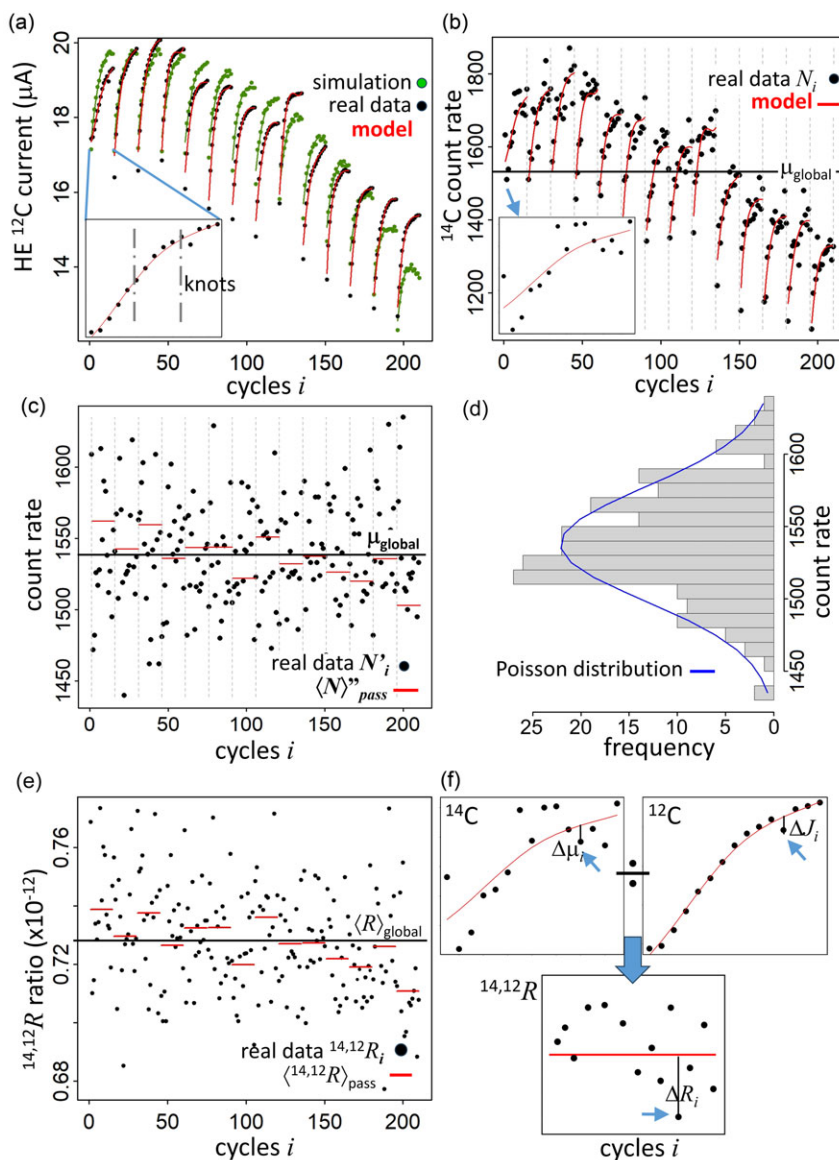


Figure 1. Graphical representation of the AMS data and model fits. Because the simulations fit the empirical data, simulations are only shown in Figure (a). (a) Empirical approximation of the simulation with spline models to the empirical and simulated behavior of the HE ^{12}C current. The inset shows the details of the smoothness of the model and data points for the first pass which are further separated in knots by the algorithm. (b) Non-stationary behavior of the raw ^{14}C count rate with the model inferred by proportionally bringing the ^{12}C model into count rate scale. (c) Stationary transformation of the raw ^{14}C counts (N_i) at the cycle level and second type of transformation at the passes level ($\langle N \rangle$). (d) Histogram of the stationary-transformed counts (N_i) with a superimposed theoretical Poisson distribution. (e) $^{14,12}\text{R}_i$ values for each cycle i and the value of the mean ratio corresponding to each pass $\langle ^{14,12}\text{R} \rangle_{\text{pass}}$. (f) Schematic explanation of the need of including the ^{12}C current uncertainty into the accounting of the internal error.

The non-stationary to stationary transformation basically aligned the mean count rate of each pass to the global mean, see black line in Figure 1c. This alignment removed any scatter or drift of the mean count rate occurring along the passes, leaving only the dispersion of the counts. Next, we calculated the quasi-Poisson overdispersion using the Pearson χ^2 (Tjor 1998) and the degrees of freedom (df) with Eq. 16.

$$\varphi_{qp} = \frac{\chi_{pearson}^2}{df} = \frac{1}{df} \sum_{i=1}^{ns} [N_i - \mu_i]^2 / \mu_i = \frac{1}{df} \sum [N'_i - \mu_{global}]^2 / \mu_{global} \quad (16)$$

The quasi-Poisson is not a real distribution but a model to which an overdispersion and quasi-likelihood calculations have been developed to deal with overdispersion and under-dispersion (non-Poisson situations). The application of Eq. 14 ensured that the dispersion was preserved from the non-stationary model into the stationary model, at least, from a quasi-Poisson point of view. In both situations, the degrees of freedom (df) were taken as $ns - n$.

Finally, we fitted the negative binomial distribution to the transformed stationary count rates using the generalized linear model option “gamlss” from the package “gamlss” in **R** (Rigby and Stasinopoulos 2005; Stasinopoulos and Rigby 2007). The generalized linear method (GLM) is a fitting of the counts which outputs ϕ_{NB1} , but ϕ_{NB1} must be corrected for the right df (see Online Supplementary Materials). Finally, we calculated the dispersion parameter D as $\sqrt{\varphi_{qp}}$ or $\sqrt{1 + \varphi_{NB1}}$; the internal error from the counting statistics and the external error from the $^{14,12}R$ ratios. The code in **R** program language to infer the models J and μ can be downloaded from github (github/oPAC) for using csv data files.

Cycles-based external error

The conventional method for expressing external error for AMS consists of calculating the standard error of the means $\langle^{14,12}R\rangle$ corresponding to each pass, as depicted in Eq. 4. The conventional passes-based external error is the scatter of $\langle^{14,12}R\rangle_{pass}$ (red lines in Figure 1e) which contains $n-1$ degrees of freedom. Furthermore, another external error is the standard error of all the $^{14,12}R_i$ values corresponding to all the individual cycles which account for $ns-1$ degrees of freedom as defined in Eq. 17.

$$SE_{\langle X \rangle, rel} = \frac{\sigma_{R_{cycles}}}{\langle^{14,12}R\rangle_{global} \sqrt{ns}} = \frac{\sqrt{\sum_{i=1}^{ns} [^{14,12}R_i - \langle^{14,12}R\rangle_{global}]^2 / (ns - 1)}}{\langle^{14,12}R\rangle_{global} \sqrt{ns}} \quad (17)$$

The equivalency of Eq. 17 to Eq. 4 originates from the Central Limit Theorem which equates the standard deviation of the means to the standard deviation of the population factored by the root of the sample size $\sigma_{R_{means}} = \sigma_{R_{cycles}} / \sqrt{s}$ where the number of cycles per pass (s) is the sample size (Brussolo 2018; Ruggieri 2016). The same approach was applied to the counting uncertainty to deduce internal error in Eq. 5. Here, we compared the cycles-based external error to the conventional passes-based external error to test if the results were obeying the Central Limit Theorem.

Calculation of the ^{12}C noise (uncertainty)

The high-energy ^{12}C beam current (I_i) relative uncertainty, shown in Eq. 18, was the noise of the current around a perfectly smooth model J_i . We can show that the form of Eq. 18 is similar to Eq. 17 if the value of J_i was constant as $\langle^{14,12}R\rangle_{global}$ then it would be possible to take J_i out and by approximating ns to $ns-1$ then both equations would be almost the same.

$$SE_{C12, rel} = \frac{\sqrt{\sum_{i=1}^{ns} \left[\frac{I_i - J_i}{J_i} \right]^2}}{ns - 1} \quad (18)$$

The $^{14,12}R$ inset (Figure 1f) shows that a perfect ratio without scatter (horizontal red line) could be obtained across a pass by dividing the models of the count rates and current (μ_i/J_i) which are also smooth and proportional to each other in their respective scales. Therefore, the count rate scatter and the noise of the current are both translated into the scatter on the $^{14,12}R$ values. Mathematically, the ratio is a function of the models and their uncertainties: $^{14,12}R_i = (\mu_i \pm \Delta\mu_i)/(J_i \pm \Delta J_i)$ where $\Delta\mu_i = \mu_i N_i$ and $\Delta J_i = J_i I_i$. For this reason, the effect of the ^{12}C uncertainty on the overall internal error calculation was taken into consideration in this paper.

Novel passes-based internal error

$$D_{passes}^2 = \varphi_{passes} = \frac{s}{n} \sum_{pass=1}^n \frac{\left[\left(\frac{\langle J_{T,pass} \rangle}{I_{T,pass}} \langle N_i \rangle_{pass} \right) - \langle \mu \rangle_{global} \right]^2}{\langle \mu \rangle_{global}} \quad (19)$$

$$SE_{passes, rel} = D_{passes} / \sqrt{N_T} \quad (20)$$

The derivation of the passes-based internal error (Eq. 19) is shown in the Supplementary Materials. In Eq. 19 and 20; N_T is the total ^{14}C counts of the whole measurement of n passes, each pass containing s cycles. N_T is the same concept defined for Eq. 2 and 5. $N_T = \sum_{i=1}^{ns} N_i$. $I_{T,pass}$ is the accumulated ^{12}C current magnitude corresponding to each pass. $I_{T,pass} = \sum_{i=i0}^{i0+s-1} I_i$. Each pass is defined between $i0$ and $i0+s-1$ cycles. For example, the first and second passes for $s=15$ are defined by cycles 1 to 15 and 16 to 30 respectively. $\langle J_{T,pass} \rangle$ is the average of accumulations of the J model current corresponding to each pass defined as $\langle J_{T,pass} \rangle = \frac{1}{n} \sum_{pass=1}^n J_{T,pass}$ where the cumulative $J_{T,pass}$ is defined exactly as $I_{T,pass}$. Both $\langle J_{T,pass} \rangle$ and $I_{T,pass}$ are accumulations at the passes level with similar magnitude, and there is no need for conversion into electrical charge. $\langle N_i \rangle_{pass}$ is the average ^{14}C counts per cycle for each pass calculated as the total counts corresponding to each pass divided by the number of cycles. $\langle N_i \rangle_{pass} = \frac{1}{s} N_{T,pass}$. $\langle \mu \rangle_{global}$ is also an average ^{14}C counts per cycle of the model calculated as the global accumulation relative to all the cycles (ns). $\langle \mu \rangle_{global} = \frac{1}{ns} \mu_T$. One $I_{T,pass}$, $J_{T,pass}$ and $\langle N_i \rangle_{pass}$ value is calculated for each pass. $I_{T,pass} = \{I_{T,1} \dots I_{T,n}\}$, $J_{T,pass} = \{J_{T,1} \dots J_{T,n}\}$, $\langle N_i \rangle_{pass} = \{\langle N_i \rangle_1 \dots \langle N_i \rangle_n\}$.

A closer inspection of Eq. 19 showed that the average counts per cycle of each pass, $\langle N_i \rangle_{pass}$, is weighted by a corresponding ratio $\langle J_{T,pass} \rangle / I_{T,pass}$ which depends on the ^{12}C current model J and the current data I . These ratios distribute around 1. In other words, $\langle N_i \rangle_{pass}$ is being transformed (see Eq. 21) and consequently, Eq. 19 can be rewritten as Eq. 22. The transformation $\langle N \rangle_{pass}''$ corresponding to each pass is illustrated in Figure 1c as horizontal red lines.

$$\langle N \rangle_{pass}'' = \frac{\langle J_{T,pass} \rangle}{I_{T,pass}} \langle N_i \rangle_{pass} \quad (21)$$

$$D_{passes}^2 = \varphi_{passes} = \frac{s}{n} \sum_{pass=1}^n \frac{[\langle N \rangle_{pass}'' - \langle \mu \rangle_{global}]^2}{\langle \mu \rangle_{global}} \quad (22)$$

Simulation of the ^{14}C count rates

The raw counts of an AMS radiocarbon measurement, as noted above, were simulated by arbitrarily drifting the high-energy ^{12}C current (I_i) in a similar fashion to real measurements. The drift function shape and magnitude were chosen from our long experience running graphite samples. The simulation assumed the observed ^{14}C count rates were random integers that follow a non-Poisson distribution. The expected ^{14}C count rate was taken as the local count rate model (μ_i). The expected count rate was calculated for each cycle i multiplying the HE ^{12}C current with or without Gaussian noise by the $^{14,12}R$ absolute value and the cycle time; dividing by the elementary charge in Coulombs and supposing 100%

transmission from the high energy side until the counting detector (see Eq. 10). At each cycle, a random count rate N_i was generated around the calculated expectation μ_i using the negative binomial random integer generator of the **R** program $N_i \sim NB(\mu_i, \varphi_{in})$. A Poisson distribution could be simulated when $\varphi_{in} = 1$. In this way, the underlying dispersion (φ_{in}) was known. Like the empirical data, the models J and μ were inferred from the simulated current and counts rates, which were transformed into N'_i . The GLM fitting was carried out and the dispersion D was calculated. The parameters of interest, like ^{12}C noise, HE ^{12}C current, non-Poisson dispersion and other parameters, were systematically and uniformly sampled, and the simulation was repeated in a Monte Carlo fashion. This method allowed the simulated uncertainty of the $^{14,12}\text{R}$ ratio to depend on the combination of the dispersion of the ^{14}C counts and the ^{12}C uncertainty or noise.

Results and discussion

Validation of the simulations

Our strategy for validating the simulations was to compare our simulated uncertainty to empirical results from the literature and our own measurements. The studied parameters that affect the simulated uncertainty were the sample $^{14,12}\text{R}$ value, the number of passes (data points), and the beam currents. Furthermore, we assessed how much the simulated internal error explained the external error under Poisson conditions.

The Poisson distribution demonstrated a better fit to the count rate data and produced a more realistic model than the Normal distribution, as shown in Figure S1 (Supplementary Materials). The symmetrical position of the confidence intervals of the Normal distribution created a small but nonsensical probability of negative counts for a fossil sample. The Poisson distribution shape not only changed with the mean, but it was asymmetric, bounded at zero counts (Figure S1b), and this property gave zero probability of negative counts. Modern samples were exempt from the Gaussian failure because their means were far from zero, as shown in Figure 1d. Nevertheless, the Poisson distribution was chosen because it was applicable to every type of sample.

We characterized the beam current as its maximum value at the high energy (HE) side, e.g. $^{12}\text{C}=20\ \mu\text{A}$ in Figure 1a. The error magnitudes from the simulations (Figure 2a) at 14 passes were comparable to the empirical results extracted from our AMS database (see Figure S2). The range of errors in the simulation, $(0.5\text{--}3) \times 10^{-3}$ in units of absolute $^{14,12}\text{R}$ (1×10^{-12}), was a bit higher than the empirical range $(0.4\text{--}2) \times 10^{-3}$, which was also in units of absolute error. In addition, the magnitude of the simulated errors (Figure 2a) was comparable to empirical errors from the literature, Figure 5 in Salazar and Szidat (2021). The error scale in the literature was $(0.5\text{--}4) \times 10^{-3}\ \text{F}^{14}\text{C}$. The conversion of our simulations absolute errors into F^{14}C should result in a similar range to the literature empirical range because our *oxa2* always has consistently given an absolute $^{14,12}\text{R}$ of $(1.45\text{--}1.5) \times 10^{-12}$ and the $\delta^{13}\text{C}$ correction is too small to change the errors scale. It must be noted that the literature results were a compendium of four publications. The parabolic shape of Figure 2a agrees with empirical measurements and theory from Nadeau and Grootes (2013).

Another interesting point about our simulations was that the scatter of the measured errors depended on the number of passes (Figure 2a, b). Every data point in Figure 2a, b came from the simulation of a certain number of passes (everything in Figure 1). This meant that repeating the simulation for the same conditions did not return the same measured error but a distribution of errors. This variability of the error measurement was due to the stochastic nature of the process. The parameter output of stochastic simulations is a distribution of values. A very simple example is generating a number of random values around a straight line with an underlying slope. Every time the simulation is performed, a slightly different slope is measured. In our case, when the number of simulated data groups was increased from 14 to 90 passes, then the variability of the measured errors decreased or in other words, the width of the output distribution of the errors became narrower.

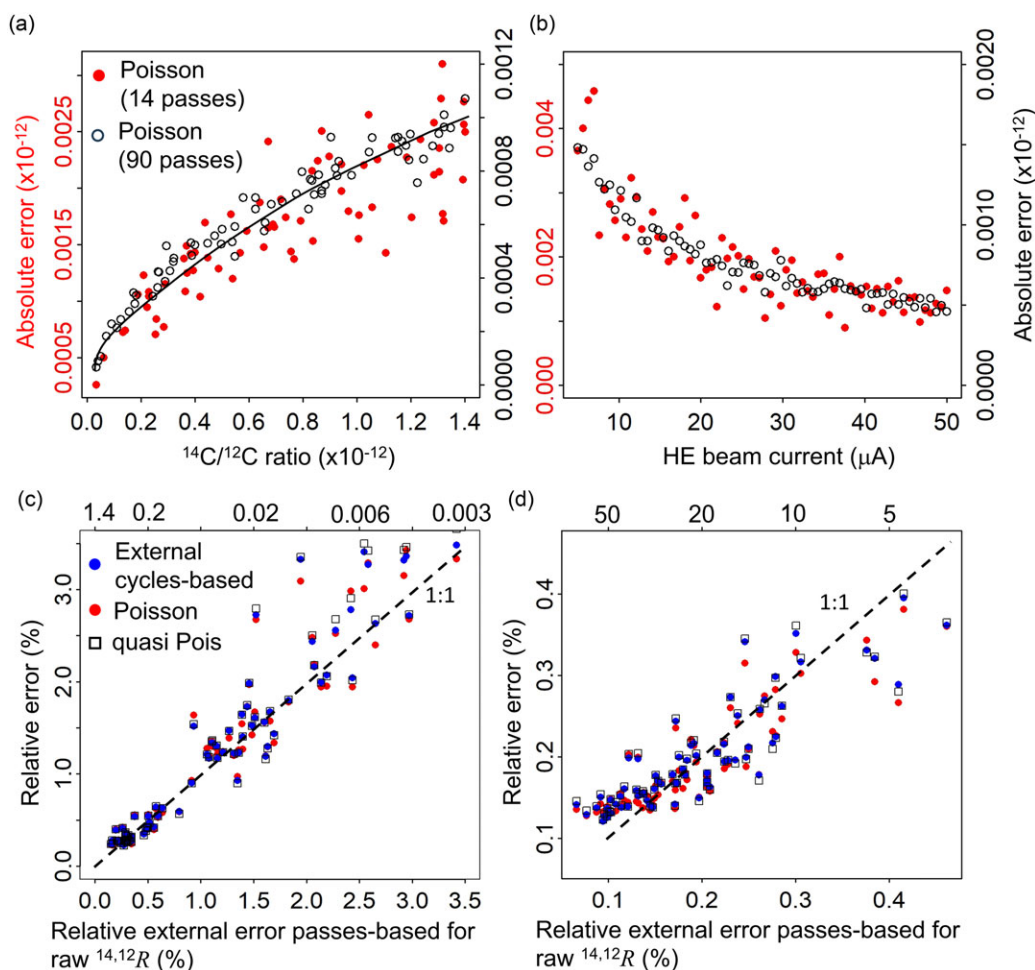


Figure 2. Validation of the model simulations with Poisson distribution. (a) Absolute Poisson error behavior across a range of $^{14,12}\text{R}$ values ($0.003\text{--}1.4$) $\times 10^{-12}$ for two sets with different number of passes. ^{12}C current = $20 \mu\text{A}$ max. (b) Absolute Poisson error behavior across a range of maximum ^{12}C currents for $^{14,12}\text{R}=1 \times 10^{-12}$ for two sets with different number of passes described in (a). The legends for (a) and (b) are the same. (c) 1:1 plot of relative errors vs. relative external error for 14 passes. The $^{14,12}\text{R}$ sample range was $(0.003\text{--}1.0) \times 10^{-12}$ and the same conditions as (a). (d) 1:1 plot of the relative errors vs. relative external error for 14 passes and the same conditions as in (b). The legends for (c) and (d) are the same.

The behavior of the absolute Poisson error showed a decaying pattern with the increase of the ^{12}C beam currents (Figure 2b). The decaying shape of the plot was expected because at higher beams, the number of counts increased, and the measurement uncertainty decreased. One method to proportionally increase the beam ^{12}C current, empirically, is to increase the sample carbon content within a certain range (Pearson et al. 1997). These authors showed how the error decayed by increasing the sample C mass; thus, we can interpret that their error decayed with ^{12}C current, giving a similar shape as our simulations (Figure 2b).

We compared several measurement errors with the conventional passes-based external error calculated using Eq. 4 (see Figure 2c, d). In both figures, the errors are scattered around the 1:1 line. The fact that there was a trend between the cycles-based external error and the passes-based external error

demonstrated that the Central Limit Theorem was applicable to the simulated data for these Poisson conditions as explained in the introduction. The errors that lay exactly on the 1:1 line presented a reduced χ^2 of 1, while the distribution around the 1:1 line should reflect a χ^2 distribution. This distribution is presented in the last part of the Results section. In some papers, it had been argued that only when the internal and external errors are about the same then the uncertainty is mostly due to counting statistics (Bonani et al. 1987; Tumey et al. 2004; Vogel et al. 2004). However, here we argue that because both errors obey a χ^2 distribution, then sometimes these errors can differ but still the counting statistics apply. One important condition of the simulations of Figure 2 is that they were fixed to Poisson distributions, and the measured quasi-Poisson error was similar to the Poisson error. This demonstrated the accuracy and versatility of the equation for the quasi-Poisson error.

As expected, the Poisson error, on average, followed the passes-based external error in Figure 2c, d. These results reproduced the traditional equivalency of the internal and external errors in AMS which relies on the CLT as shown in Eq. 4 and 5.

Simulated non-Poisson counting statistics

In this subsection, the simulation was carried out for count rate distributions spread away from the Poisson distribution. The accuracy and precision of measuring the dispersion D was checked and how this precision depended on the number of passes under non-Poisson conditions. Finally, a new passes-based internal error was demonstrated.

Under non-Poisson conditions, the histogram of the particles of Figure 1c became wider than a Poisson distribution (see Poisson example in Figure 1d) when the dispersion D was increased in the simulation. This generated data (input) was fed into the statistical model calculations to output the dispersion D and the errors. The measured dispersion D (D output) presented a good agreement with the input dispersion D for 14 and 90 passes, as shown in Figure 3a, b. This agreement showed that our method of calculating the non-Poisson dispersion was accurate. The uncertainty to measure the dispersion D was quantified with the confidence interval width ($\sigma_D = 0.050$ and 0.021 for 14 and 90 passes, respectively) corresponding to the D input- D output plots (Figure 3a, b). The magnitude of the scatter σ_D translates into higher uncertainty to measure D . The difference in scatter between Figure 3a, and 3b was explained above in “validation of the simulations.” If more data points (passes) were available, then the measurement of the error of the whole measurement would become more robust (better reproducibility and accuracy of measuring the error). Figure 3a, b also showed that the Negative Binomial method produced a similar dispersion D to the quasi-Poisson method. This is significant because the Negative Binomial is computationally exhaustive and involves optimization of the NB model recursively, while the quasi-Poisson just outputs the Pearson χ^2 equation once. Thus, the quasi-Poisson method is an accurate and simple equation to calculate the non-Poisson error.

The errors of Figure 3c, d corresponded to the simulations of Figure 3a, b; therefore, the increase of the error was due to the increase of the ^{14}C dispersion D . As expected, the measured Poisson error of the simulations was not able to capture the increase in count rate dispersion, while the quasi-Poisson and the cycles-based external errors (Eq. 17), to some degree, correlated with the conventional passes-based external as shown by their slope and scatter around the dashed line on Figure 3c. The Pearson correlation coefficient (r) of the quasi-Poisson vs. passes-based external errors was 0.47. When the number of passes was increased from 14 to 90, then the quasi-Poisson errors showed a better agreement with the passes-based external error with a $r = 0.78$. The same increase in agreement occurred for the cycles-based external (Figure 3d). The increase in correlation was mainly due to the increase in slope which became closer to the 1:1 line. The agreement of the cycles-based external error at higher number of passes indicated that the CLT required more data points to be applicable under non-Poisson conditions. In practice, to measure 14 passes, it takes 2.4 days but 90 passes would take 15 days and consume entirely the sample target. The 90-passes results are depicted primarily for demonstration purposes. Nevertheless, the quasi-Poisson error, to some degree, can follow the external error much

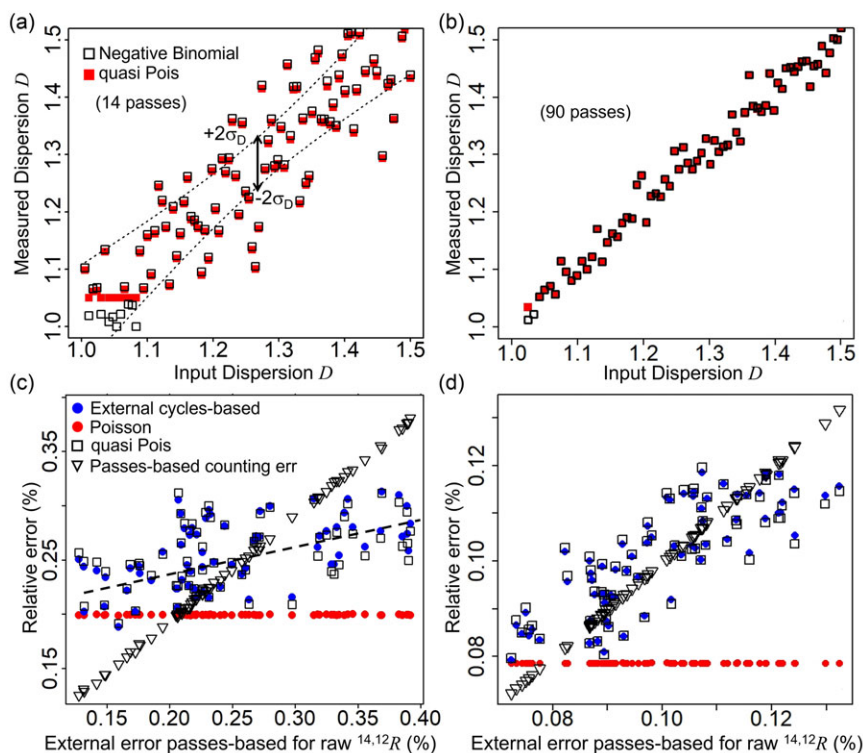


Figure 3. Simulated non-Poisson ^{14}C counting statistics for $^{14,12}\text{R} = 1.0 \times 10^{-12}$ and $20 \mu\text{A}$ which resulted in very similar N_T counts for each dispersion D . (a) D input- D output plot is the plot of the measured dispersion D vs. input dispersion D for 14 passes. 2σ confidence interval is shown with dashed lines. (b) Measured dispersion D vs. input dispersion D for 90 passes, (a) and (b) shares the same legend. (c) Diverse relative errors vs. conventional passes-based external error for 14 passes. (d) Diverse relative errors vs. conventional passes-based external error for 90 passes. The error increase on the x,y axis on (c) and (d) was created by increasing the dispersion D of the ^{14}C distribution shown in (a) and (b).

better than the Poisson calculations under non-Poisson conditions. The uncertainty to measure the dispersion D (σ_D) decreased very fast, starting at 7 passes and stabilizing at 90 passes (see Figure S3 in the Supplementary Materials). The quasi-Poisson error correlation r with the external error increased very fast for the same range of passes.

We suggest that the difficulty in capturing the conventional external error using the quasi-Poisson internal error and the cycles-based external error arises because they are calculated at the level of the cycles and conventional external errors are based on the passes. Therefore, we developed a novel internal error which was based on the count rates analysed at the passes level (see Eq. 19, 20 and 22). We emphasize the similarity between our novel passes-based internal error equation (Eq. 22) with the literature quasi-Poisson equations (Eq. 16). This similarity of outcome validated Eqs. 19 and 22. The passes-based internal error was derived from the conventional passes-based external error converted into ^{14}C count rates and ^{12}C current. However, this conversion required that the models μ and J were accurately measured. Furthermore, the passes-based internal error defined the $^{14,12}\text{R}$ ratio averages at the global and passes levels as the ratio of the total accumulation of each isotope at the respective level, whereas the conventional external error uses true averages of the ratios.

The novel passes-based counting internal error (Eq. 22) showed excellent agreement with the conventional passes-based external error (Eq. 4) as shown in Figure 3c, d. The passes-based counting

internal error was robust and independent of the number of passes in contrast with the other errors. Eq. 22 was back-engineered from Eq. 4 but Eqs. 19, 22 have a different definition for the global and passes ratios. In addition, Eqs. 19, 22 only work if the counting and current models are well inferred.

Simulated effect of the ^{12}C uncertainty on the $^{14,12}\text{R}$ error

The ^{12}C uncertainty in Figure 4a was created by adding Gaussian noise into the input ^{12}C current function of the simulation and then quantifying the ^{12}C uncertainty with Eq. 18. As noted in the “Calculation of the ^{12}C noise (uncertainty)” of the Methods, the ^{12}C noise or uncertainty caused the passes-based external error to increase in the range of 0.15%–0.45% for a modern ^{14}C sample. However, the increase of ^{12}C noise did not cause the same effect in a low- ^{14}C sample. The reason is the ^{12}C uncertainty magnitude was too low (0.05%–0.35%) for the old sample compared to the overall error of 2.5%–5.0%. This result contrasted with the effect of increasing the ^{14}C dispersion, which is the same for both fossil or modern samples.

As expected, neither the quasi-Poisson and Poisson internal errors accounted for the ^{12}C uncertainty, as shown in Figure 4b, but a simple error propagation between the quasi-Poisson and the ^{12}C uncertainty

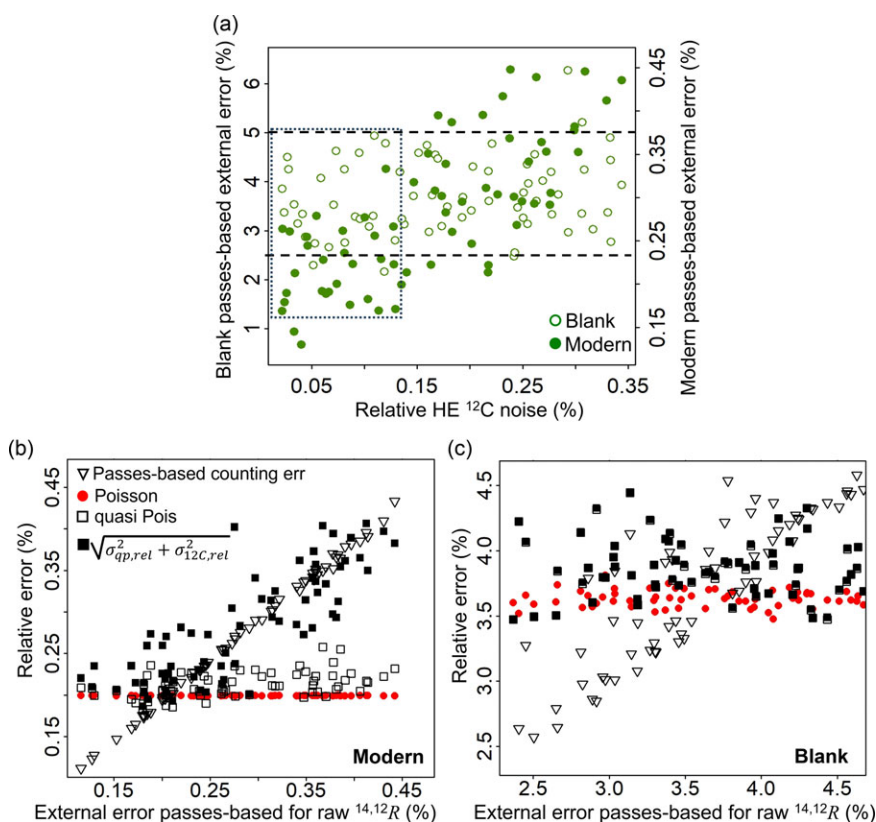


Figure 4. Simulated ^{12}C noise impact on error at 20 μA . (a) passes-based external error vs. the ^{12}C current noise for blank sample ($^{14,12}\text{R}=0.003\times 10^{-12}$) and modern sample ($^{14,12}\text{R}=1.0\times 10^{-12}$). The external error of the blank shows no correlation with the ^{12}C noise and stays within the dashed lines, while for the Modern sample there is no increase within the dash square until high ^{12}C noise. (b) Diverse relative errors vs. the passes-based external error, including the error propagation of the quasi-Poisson error and ^{12}C uncertainty. The external error increase is due to the large increase of ^{12}C noise from 0.05% to 0.35%. (c) Diverse relative errors vs. passes-based external error for a blank sample. The variation of the errors in (b) and (c) were due to the variation of the ^{12}C uncertainty as shown in (a).

showed good agreement for a modern sample. The error propagation did not work for the fossil sample (Figure 4c) due to the magnitude of ^{12}C uncertainty being too small to cause any perceptible increase in the overall error. On the other hand, the passes-based internal error from Eq. 22 worked well for both cases. Figure 4 shows the importance of accounting for the ^{12}C uncertainty to the internal error for non-fossil samples for the simulated conditions.

In summary, with the developed equations and demonstration plots, we propose a framework for a modified internal error calculation that include two novel aspects to explain the external error: the non-Poisson distribution and ^{12}C uncertainty. It was important to infer models for the count rate and the ^{12}C current. Afterwards, it was possible to calculate the quasi-Poisson and ^{12}C errors. The longer version of the simulation involved the conversion to stationary count rates and the Negative Binomial modeling.

Our simulations showed that non-Poisson dispersion did not fully mirror the external error. However, it was still, on average, closer to the external error than the Poisson error. We also showed that if the ^{12}C uncertainty is low, then it has no impact on the overall error. It is necessary to quantify these errors for each ^{14}C measurement to check if they are considerable or not because the errors depend on many conditions, such as the sample's $^{14,12}\text{R}$, the ^{12}C current magnitude, the number of passes, and instrument stability.

Empirical results

We analyzed the number of Poisson, slightly non-Poisson and strongly non-Poisson results; we measured the degree of ^{12}C relative noise in the empirical data; examined the agreement of the internal errors with the passes-based external error for Poisson and non-Poisson results from the point of view of 1:1 plots and the chi-squared distribution; and how the $\delta^{13}\text{C}$ corrections improved the agreement in average.

The number of Poisson and non-Poisson results were recorded in the form of a histogram, shown in Figure 5a. We selected the threshold for Poisson results at a dispersion D range of 1–1.05. This selection was based on the uncertainty of calculating the dispersion D ($\sigma_D = 0.05$) shown in the 2σ confidence interval of the D input- D output plot (Figure 3a). We were not able to find a precedent for defining the uncertainty of measuring the dispersion D using D input- D output simulation plots, therefore using σ_D is a novel decision-making method. The proportion of Poisson results was 63% out of 7985 samples and reference materials within the range of dispersion D of 1.0 to 1.05 (1 to $1 + \sigma_D$). Around 34% corresponded to samples with slightly non-Poisson dispersion D in the range of 1.05–1.2 ($1 + \sigma_D$ to $1 + 4\sigma_D$). The number of samples with D higher than 1.2 ($> 1 + 4\sigma_D$) was 227, or 2.8%. Overall, most of the time, our instrument produced Poisson results, but the number of highly non-Poisson results cannot be ignored since each reported error must be independently and carefully determined for each sample. We think that the degree of non-Poisson error must be determined for each sample case by case. A dispersion of 1.2 or higher means that the real error is considerably high, 1.2 times the Poisson error or even higher. We can explain the presence of non-Poisson errors to the random instability of the AMS instrument, but this is a preconceived explanation that needs to be studied in the future.

The number of sample measurements in previous work (Vogel et al. 2004) supporting Poisson distribution was ca. 8. In our study, we found 227 radiocarbon measurements with strong non-Poisson behavior ($D > 1.2$). However, we can only generalize that Poisson behavior is observed more frequently with non-Poisson exceptions, and we recommend that counting statistics should be measured for each sample.

Figure 5a also shows that histograms for fossil samples (black trace) were similar to the whole population (red trace). This similar behavior meant that the degree of dispersion D was independent of the sample $^{14,12}\text{R}$ value. Otherwise, the shape of these histograms would have been different. As a routine procedure, each magazine measurement included at least four fossil samples. If the AMS instrument was behaving “non-Poisson” during a certain measurement, then samples and fossil samples

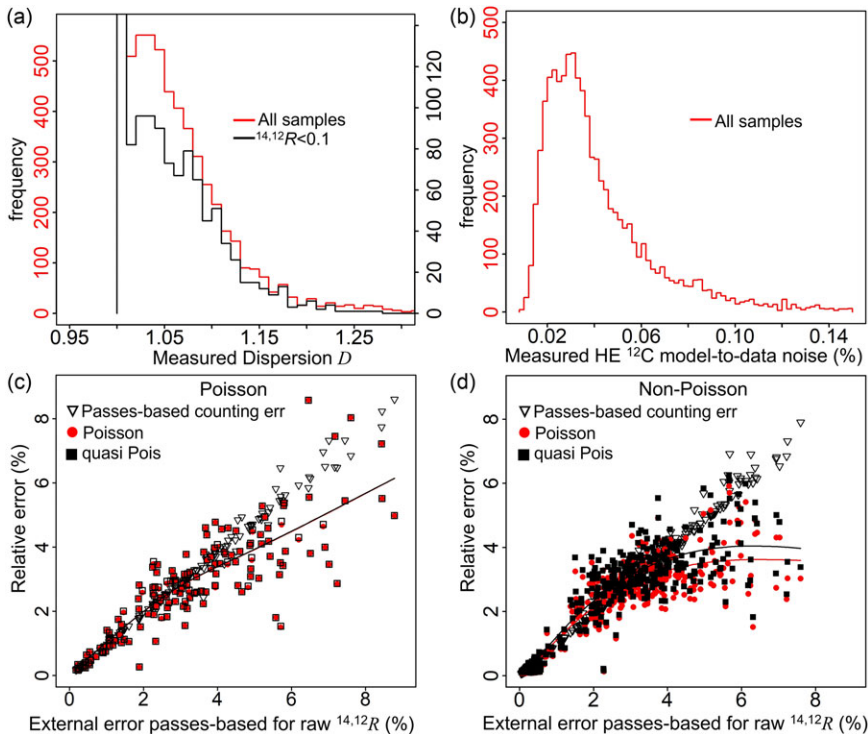


Figure 5. Empirical results for ^{14}C non-Poisson dispersion and ^{12}C uncertainty. (a) Histogram of dispersion D for all the samples and standards in our database and for fossil samples. (b) Histogram of ^{12}C uncertainty (noise) for all the samples. (c) 1:1 plot of relative errors vs. passes-based external error at Poisson conditions: $1 < D \leq 1.05$, ^{12}C uncertainty $> 0\%$, $20 \mu A < ^{12}C$ current $< 25 \mu A$. (d) 1:1 plot of relative errors vs. passes-based external error at non-Poisson conditions. Same conditions as (c) but $D > 1.05$.

would have presented non-Poisson results or Poisson results depending on the instrument condition. The same is true for all the other sample magazines; therein, the similarity of the histogram shapes.

On the other hand, the ^{12}C uncertainty did not exceed a low value of 0.14%, as shown in the histogram of Figure 5b for all the ^{14}C samples measured in our laboratory. The simulations of Figure 4a already showed that ^{12}C uncertainty in the range of 0–0.14% is too low to affect the external error, thus the ^{12}C uncertainty can be discarded from the internal error of our empirical results. A low ^{12}C uncertainty cannot overcome the scatter or variability of the measurement of the external error, even for a modern sample. Nevertheless, we cannot assume that ^{12}C uncertainty can be neglected for radiocarbon AMS. Every laboratory should measure their ^{12}C noise to quantify it and decide.

The Poisson and quasi-Poisson internal errors, to some degree, agreed with the passes-based external error (see Figure 5c) when the experimental conditions favoured Poisson statistics ($D < 1.05$). There is agreement between the internal and external errors as predicted by the simulations in Figure 2c. While, for non-Poisson conditions, internal errors strongly deviated from a 1:1 agreement after 4% (Figure 5d) with the quasi-Poisson error slightly better than the Poisson. The simulations in Figure 3c predicted a stronger deviation under non-Poisson conditions vs. less deviation under Poisson conditions. The reason that the internal and external errors, on average, agree is that their underlying relationship is the chi-squared distribution which allows for many data points to strongly differ within a chi-squared range.

Empirical internal vs. external errors chi-squared relationship

The Central Limit Theorem tells us that the average ratios, corresponding to each pass ($^{14,12}R$ values) are normally distributed because each pass is a statistical sample. Therefore, from the theory of the chi-squared distribution (Anderson et al., 1994), the quotient Q should follow a chi-squared distribution (Eq. 23) because Q is the sum of squared normally distributed values standardized by the proposed error or expectation error (internal error). The proposed error is the hypothetical uncertainty that the data should follow along the measurement, and we take the proposed error as a constant value over n passes. For radiocarbon, the proposed error comes from the internal error shown in Eq. 5 ($\text{proposed err.} = \sigma_{\text{pois}} / \sqrt{s}$). If the numerator of Q is multiplied and divided by $n-1$, the numerator becomes the squared standard deviation ($\sigma_{R \text{ means}}^2$) explained in Eq. 4. If we divide the numerator and denominator by their mean values to set the errors in relative form and divide both sides by \sqrt{n} to set both errors as standard errors, then Q becomes $(n-1)$ times the squared quotient of the external and internal relative errors. The external error is the observed realization of the hypothesized proposed uncertainty then the external and internal errors are comparable.

$$Q = \sum_{i=1}^n \frac{\left({}^{14,12}R_i - \langle {}^{14,12}R_i \rangle \right)^2}{\text{proposed err.}^2} \equiv \frac{n-1 \left(\sigma_{R \text{ means}} / \langle {}^{14,12}R_i \rangle \sqrt{n} \right)^2}{\left(\sigma_{\text{pois}} / \sqrt{s} / \mu_{\text{global}} \sqrt{n} \right)^2} = \frac{(n-1) \text{ ext. err.}^2}{\text{internal err.}^2} \sim \chi_{n-1}^2 \quad (23)$$

Over a 2.5-year period, many target samples were consistently measured using 14 passes ($n = 14$), and a reduced Q value was calculated for each sample (Figure 6). The reduced Q value is Q divided by $n - 1$. The theoretical reduced chi-squared distribution was also calculated for $n = 14$ and its distribution shape was similar to that reported in the literature (Wendt and Carl 1991). Most of the reduced chi-squared values are near to 1.0, but other values are higher or lower. This variability is due to normal random events. If the histogram of the reduced Q values was different than the theoretical reduced chi-squared distribution, then other sources of errors were affecting the results, outliers were present, or the expectation error was incorrectly selected.

Both reduced Q histograms calculated with the raw $^{14,12}R$ and the Poisson and quasi-Poisson errors did not produce good chi-squared distributions (Figure 6a). Both reduced Q histograms had a tail to the right side because the reduced Q values were higher than they should have been. In contrast, the histograms of reduced Q values with the numerator calculated with the $^{14,12}R$ values corrected by the

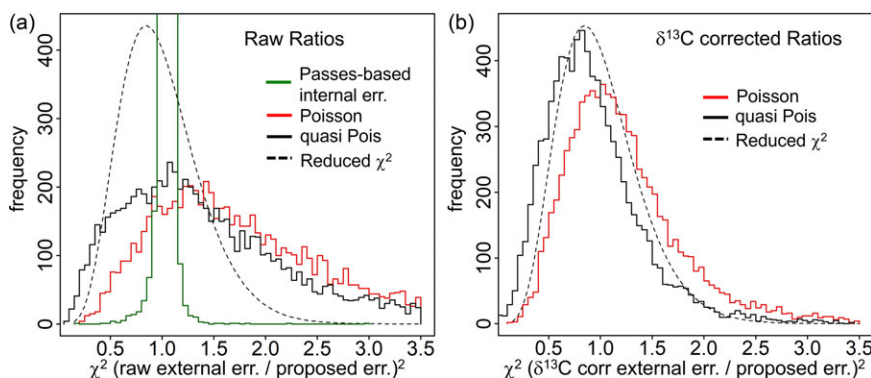


Figure 6. Reduced Q values histograms and reduced χ^2 distribution. The area of the χ^2 distribution was approximately equal to the area of the histograms. The reduced χ^2 distribution was generated by generating the χ^2 distribution for $n=14$ and then dividing by $n-1$. The reduced Q values were calculated using two types of external errors relative to several proposed errors shown in the legends. a) reduced Q values calculated with external error with raw ratios. b) reduced Q values calculated with external error corrected with the $\delta^{13}C$ information.

$\delta^{13}\text{C}$ presented a shape very similar to the chi-squared distribution (Figure 6b). Moreover, the quasi-Poisson error produced a better distribution than the Poisson error. The quasi-Poisson error almost mirrored the chi-squared distribution. As the internal errors were the same in Figure 6a and 6b, there is an indication that the $\delta^{13}\text{C}$ -uncorrected external errors were too high and not representative of the variability of the measurement in Figure 6a. The uncorrected external errors were too high because during the analysis, the measured $^{14,12}\text{R}$ values were not constant but usually drifted. This $^{14,12}\text{R}$ drift or fractionation caused by machine instability or ion source sputtering was another error source or uncounted error that was corrected by $\delta^{13}\text{C}$. The $\delta^{13}\text{C}$ correction for carbon isotope fractionation, including machine-derived fractionation was demonstrated by (Steier et al. 2004). The empirical evidence of the $\delta^{13}\text{C}$ -corrected external error agreeing with the internal error in a reduced χ^2 fashion, statistically, showed why the external error should be a realization of the expectation error or proposed internal error.

The reduced Q values represent the ratio of external to internal error squared. Therefore, examining the distribution of the reduced Q values offers another perspective on the data shown in the 1:1 plots of internal error versus external error in Figure 5c and 5d. The reduced Q histograms show more detail and clearer patterns than the 1:1 plots. All data pairs with reduced Q values close to 1.0 fell on the 1:1 line in Figure 5. Figure 6a shows the same data of Figure 5c, d; however, the difference between the quasi-Poisson and Poisson errors can only be realized in the reduced Q histograms. The ideal behavior of the $\delta^{13}\text{C}$ -corrected external error versus internal error in Figure 6b cannot be easily seen in a 1:1 plot. The mechanism of $\delta^{13}\text{C}$ correction consists on supposing that the raw $^{14,12}\text{R}$ suffers the same drift shape but twice in magnitude as the $^{13,12}\text{R}$ along the passes, so the $\delta^{13}\text{C}$ drifts from a universal constant value, -25‰ , decided by international convention (Steier et al. 2004). The magnitude of the $\delta^{13}\text{C}$ drift is used to push the raw $^{14,12}\text{R}$ in the opposite direction, correcting the drift and leaving the scatter of the ratios. On the other hand, as explained in the methods section, the non-stationary to stationary transformation eliminates the drift of the count rates from an absolute constant value, its global mean. The drift is eliminated by aligning the mean count rate of each pass to the global count rate; thus, doing a similar job to the $\delta^{13}\text{C}$ correction. For this reason, the $\delta^{13}\text{C}$ -corrected external error is more compatible with the internal error, being the non-Poisson error a slightly better representative of the count rate dispersion from a chi-squared perspective.

Another important takeaway from Figure 6b is that sometimes, the internal and external errors can differ as long as their squared ratios stay within the confidence region of a reduced chi-squared distribution and the system is still governed by counting statistics. This chi-squared confidence region is between 0.38 and 1.9 for degree of freedom of 13 and a significance level of 0.05. This takeaway contradicts claims that both errors should be similar if artifacts and drift effects are well corrected (Bonani et al. 1987; Tumey et al. 2004; Vogel et al. 2004). Furthermore, we propose that the definition of internal and external errors should be extended to include the ^{12}C uncertainty for completeness as follows: “The internal errors are the propagated ^{14}C and ^{12}C uncertainties, as we calculate them along with the measurand $^{14,12}\text{R}$, whereas the external errors are the uncertainties observable from the spread in measurands.” In practice, the ^{14}C counting uncertainty will continue dominating the internal error with the expectation that the ^{12}C uncertainty is, in general, too small as demonstrated in this work in Figure 5c.

On the other hand, the reduced Q values corresponding to our passes-based internal error produced a narrow histogram. The passes-based internal error ultimately failed to match the chi-squared distribution. Therefore, the passes-based internal error is not independent of the conventional passes-based external error. However, the novel passes-based internal error is an extremely good predictor of the passes-based external error, as shown by the empirical data in Figure 5 and simulations (Figures 3, 4). The extremely good agreement of the novel passes-based internal error with the conventional external error is indirect evidence that the models J_i , μ_i were accurately estimated.

Real-world applications

The quoted error absolute values were calculated with the equation proposed by Aerts-Bijma et al. (2021) using Poisson or non-Poisson internal errors and converted to ^{14}C years rounded to the nearest integer (see Figure 7a). The subtraction between the non-Poisson quoted error and the Poisson quoted error are shown in Figure 7b, c and this difference was considerable. Some specific examples with high dispersion D ($D > 1.2$) are written below in the format ^{14}C of age \pm Poisson quoted error, non-Poisson quoted error: 49869 \pm (437, 456) years BP; 40066 \pm (134, 147) years BP; 29986 \pm (73, 79) years BP and 20112 \pm (39, 43) years BP. The respective differences in the quoted errors are: 19, 13, 6 and 4 in years BP. The respective ratios of the non-Poisson to Poisson quoted errors are: 1.04, 1.10, 1.08 and 1.10 which is a non-Poisson factor of 1.1 on average. These differences are representative values from Figure 7b, c, but higher and lower values can also be expected.

As described in the introduction, it is usually observed that, at a higher processing level, the actual uncertainty of replicates turns out to be higher than the quoted error due to unknown errors. The quoted error has been an essential parameter for calculating expanded errors (by addition or by multiplication) for many applications. Therefore, the usage of non-Poisson quoted errors should decrease the existing known expansion parameters (e.g. error multipliers), helping to close the gap of unknown uncertainties. For example, previously for Figure 7b, c; we showed that on average the non-Poisson quoted error with high dispersion D was 1.1 times larger than the Poisson quoted error. While Aerts-Bijma et al. (2020)

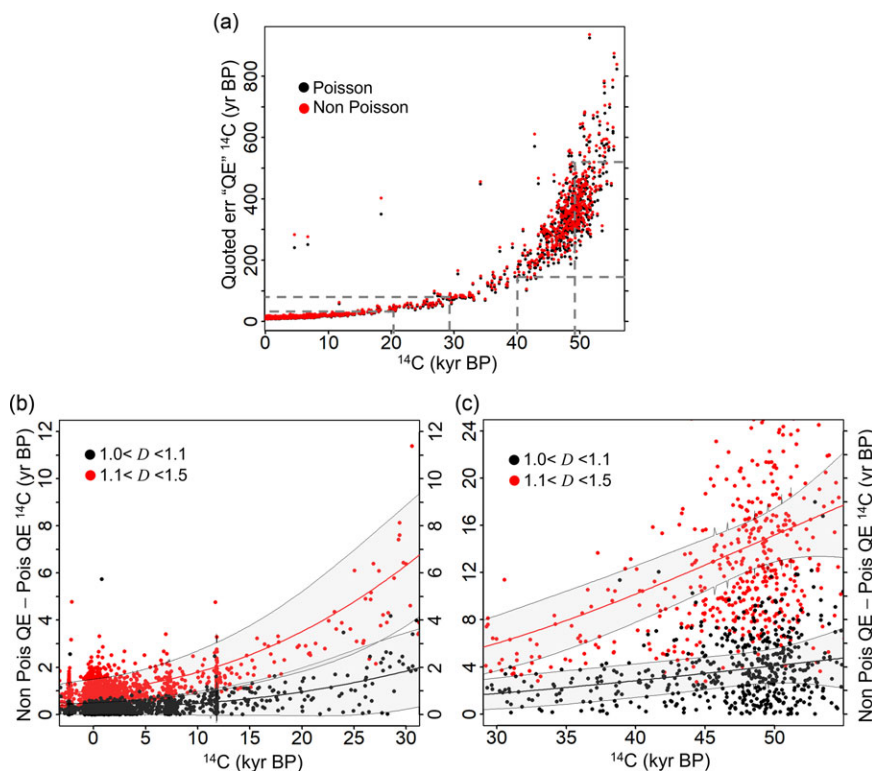


Figure 7. Non-Poisson error translation into radiocarbon age quoted error. (a) Quoted error was calculated from Poisson and non-Poisson counting statistics for different ages with dashed lines indicating representative examples. (b) Difference between the non-Poisson and Poisson quoted errors in radiocarbon years BP for the radiocarbon age range of 0–30,000 years BP. The differences are divided in two groups: differences calculated including quoted errors that included non-Poisson errors with dispersion D between 1 and 1.1 (black) and for dispersion D between 1.1 and 1.5 (red) (c) Difference between the non-Poisson and Poisson quoted errors in radiocarbon years BP for the radiocarbon age range of 30,000–50,000 years BP.

found errors multipliers for graphitization, graphite+combustion and graphite+combustion+pretreatment combinations to be 1.1, 1.4 and 1.6 respectively using Poisson quoted errors. Therefore, for this specific example, the non-Poisson quoted error explains 100% of the graphitization error multiplier, 25% for the graphite+combustion combination and 17% for the graphite+combustion+pretreatment combination. The apportionment explanation was calculated as $\frac{\text{non Poisson factor}-1}{\text{error multiplier}-1} \times 100$. The non-Poisson error seems to be more realistic than the Poisson, but it is more complicated to calculate. In contrast, the Poisson error is simpler to calculate for a quick and rough prediction of the error of the measurement. Finally, the non-Poisson cannot explain all the systematic effects of replications at high sample treatment level. These effects must be studied using a different method, top-down approach, in future.

Conclusions and outlook

A new mathematical framework has been developed to quantify two novel parameters that affect the internal error of radiocarbon AMS. These are the non-Poisson uncertainty of the ^{14}C count rate and the ^{12}C uncertainty or the noise of the ^{12}C beam. Simulations were developed based on inferring proportional models for the ^{14}C count rate and ^{12}C current and calculating the quasi-Poisson and ^{12}C errors.

These simulations were verified to examine how these two new parameters were affected by the $^{14,12}\text{R}$ value of the sample, the ^{12}C current magnitude, and the number of passes. Besides the nature of the sample, the stability of the AMS instrument is primarily what controls these two parameters. The simulations produced results similar to those of the literature and our empirical results. The simulations helped to explain better the following relationships: internal and external error vs. sample $^{14,12}\text{R}$ value and vs. HE ^{12}C current magnitude; and the 1:1 correspondence, on average, between the internal error and external error for Poisson and non-Poisson conditions.

The analysis of more than two years of empirical radiocarbon results (7985 samples and reference materials) indicated that most of the results (63%) showed Poisson characteristics, 34.2% were slightly non-Poisson, and 2.8% had strongly non-Poisson characteristics. The ^{12}C uncertainty ranged between 0% and 0.15%, which was too small to affect the external error of any ^{14}C sample.

The comparison between the reduced Q value histograms and the reduced chi-squared distribution showed the importance of the $\delta^{13}\text{C}$ correction and within this frame of chi-squared distribution, we showed that the quasi-Poisson error is slightly better than the Poisson, in general, to explain the external error.

We emphasize that these results reflected the behavior of our MICADAS instrument. However, we have clearly demonstrated that radiocarbon AMS cannot be assumed to be a Poisson process. While we cannot generalize about how much non-Poisson behavior occurs on other Radiocarbon AMS instruments or that the ^{12}C uncertainty can be neglected, it would be fascinating to see these calculations applied to other laboratories results.

Specifically, we hope our work encourages interest in radiocarbon laboratories evaluating:

- Non-Poisson statistical models, such as quasi-Poisson and negative binomial models, to better account for internal error in radiocarbon AMS quoted errors, as traditional Poisson assumptions cannot be assumed always to be valid.
- The ^{12}C current noise, as it could significantly affect the measurement error, especially in modern samples.

Acknowledgments. The data used in this work came from graphite measurements funded by the Canadian Foundation for Innovation. We thank Sarah Murseli, Jennifer Walker and the radiocarbon team at the AEL facility for their hard work converting an enormous number of physical samples into valuable numerical data.

Definition of concepts. Statistical sample: sometimes referred as sample is a set of data observations that should represent the characteristics of the whole population. The population is an ideal concept, and it is the set of all possible data observations. In our context, the set of all possible observations (e.g. ratios) from a graphite target. n is the number of passes in a measurement and s is the size of the sample or the number of cycles per pass.

Internal error (cycles-based): uncertainty or error due to variability of the ^{14}C counting.

Passes-based internal error: novel error of the counting means corresponding to each pass.

External error: Normal distribution uncertainty of the $^{14}\text{C}/^{12}\text{C}$ ratios.

Passes-based external error: conventional error of the ratio means corresponding to each pass.

Cycle-based external error: error of the ratios corresponding to all the cycles.

$^{14,12}\text{R}$, $^{13,12}\text{R}$: variables that represents the $^{14}\text{C}/^{12}\text{C}$ or $^{13}\text{C}/^{12}\text{C}$ ratios.

I_i , N_i data: set of HE ^{12}C and ^{14}C count rate data respectively evaluated at each cycle i .

J_i , μ_i models: set of values of the mathematical models for the HE ^{12}C and ^{14}C count rate data respectively evaluated at each cycle i .

There is a single model for each parameter for each pass.

Dispersion D : model-dependant parameter that represents how many times the actual distribution of the count rates is wider than a Poisson distribution. D is root of the quasi poisson φ (φ_{qp}) or could be the root of $1 + \varphi_{\text{NB1}}$.

φ_{qp} : the parameter of the quasi-Poisson model and comes from the Pearson χ^2 of the count rates.

φ_{NB1} : the parameter of the negative binomial model which is obtained from fitting the count rate data to the model. φ_{NB2} is higher than zero.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/RDC.2025.10115>

References

- Aerts-Bijma AT, Paul D, Dee MW, Palstra SWL and Meijer HAJ (2021) An independent assessment of uncertainty for radiocarbon analysis with the new generation high-yield accelerator mass spectrometers. *Radiocarbon* **63**(1), 1–22. <https://doi.org/10.1017/RDC.2020.101>
- Alrawashdeh, M. J. (2021) An adjusted Grubbs' and generalized extreme studentized deviation. *Demonstratio Mathematica*, **54**(1): 548–557. <https://doi.org/10.1515/dema-2021-0041>
- Anderson DS, Sweeney D and Williams T (1994) *Introduction to Statistics: Concepts and Applications*, 3rd edn. St Paul, MN: West Publishing Company.
- Blaauw M, Aquino-López M and Christen JA (2024) Modelling chronologically ordered radiocarbon dates in R. *Radiocarbon*, 1–11. <https://doi.org/10.1017/RDC.2024.56>
- Bonani G, Beer J, Hofmann H, Synal H-A, Suter M, Wölfli W, Pfeleiderer C, Kromer B, Junghans C and Münnich KO (1987) Fractionation, precision and accuracy in ^{14}C and ^{13}C measurements. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **29**(1–2), 87–90. [https://doi.org/10.1016/0168-583X\(87\)90210-2](https://doi.org/10.1016/0168-583X(87)90210-2)
- Brussolo ME (2018) Understanding the central limit theorem the easy way: A simulation experiment. In *The 2nd Innovative and Creative Education and Teaching International Conference*, Basel, Switzerland: MDPI, 1322.
- Burr GS, Donahue DJ, Tang Y and Jull AJT (2007) Error analysis at the NSF-Arizona AMS facility. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **259**(1), 149–153.
- Currie LA, Eijgenhuijsen EM and Klouda GA (1998) On the validity of the Poisson hypothesis for low-level counting; investigation of the distributional characteristics of background radiation with the NIST individual pulse counting system. *Radiocarbon* **40**(1), 113–127.
- Dupuy J-F (2018) *Statistical Methods for Overdispersed Count Data*. Amsterdam: Elsevier.
- Evans M and Rosenthal JS (2004) *Probability and Statistics: The Science of Uncertainty*, 2nd edn. New York and Basingstoke: W.H. Freeman and Co.
- GitHub (2025) Online project for AMS coding (oPAC). <https://github.com/oPAC-2025/non-Poisson-calculations>
- Greene W (2008) Functional forms for the negative binomial model for count data. *Economics Letters* **99**(3), 585–590.
- Heaton TJ, Blaauw M, Blackwell PG, Bronk Ramsey C, Reimer PJ and Scott EM (2020) The IntCal20 approach to radiocarbon calibration curve construction: A new methodology using Bayesian splines and errors-in-variables. *Radiocarbon* **62**(4), 821–863. <https://doi.org/10.1017/RDC.2020.46>
- Litherland AE, Paul M, Allen KW and Gove HE (1987) Fundamentals of accelerator mass spectrometry. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **323**(1569), 5–21.
- Malonda AG and Carles AG (2020) Radioactivity counting statistics. In L'Annunziata M (ed), *Handbook of Radioactivity Analysis: Volume 2*, 4th edn. Academic Press, 627–667.
- McNichol AP, Jull AJT and Burr GS (2001) Converting AMS data to radiocarbon values: Considerations and conventions. *Radiocarbon* **43**(2A), 313–320.
- Nadeau M-J and Grootes PM (2013) Calculation of the compounded uncertainty of ^{14}C AMS measurements. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **294**, 420–425.
- Pearson A, McNichol AP, Schneider RJ, Von Reden KF and Zheng Y (1997) Microscale AMS ^{14}C measurement at NOSAMS. *Radiocarbon* **40**(1), 61–75.
- Pinsky MA and Karlin S (2011) *An Introduction to Stochastic Modeling*, 4th edn. Amsterdam: Academic Press Elsevier.

- Poplová M, Sovka P and Cifra M (2017) Poisson pre-processing of nonstationary photonic signals: Signals with equality between mean and variance. *PLoS One* **12**(12), e0188622.
- Povinec PP, Litherland AE and von Reden KF (2009) Developments in radiocarbon technologies: From the Libby counter to compound-specific AMS analyses. *Radiocarbon* **51**(1), 45–78.
- R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Rigby RA and Stasinopoulos DM (2005) Generalized additive models for location, scale and shape (with discussion). *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **54**(3), 507–554.
- Ruggieri E (2016) Visualizing the central limit theorem through simulation. *PRIMUS* **26**(3), 229–240.
- Salazar G and Szidat S (2021) Reassessment of uncertainty expansion by linear addition of long-term components from top-down information. *Radiocarbon* **63**(6), 1657–1671.
- Scott EM, Cook GT and Naysmith P (2007) Error and uncertainty in radiocarbon measurements. *Radiocarbon* **49**(2), 427–440. <https://doi.org/10.1017/S0033822200042351>
- Scott EM, Naysmith P and Cook GT (2018) Why do we need ^{14}C inter-comparisons?: The Glasgow - ^{14}C inter-comparison series, a reflection over 30 years. *Quaternary Geochronology* **43**, 72–82. <https://doi.org/10.1016/j.quageo.2017.08.001>
- Stasinopoulos DM and Rigby RA (2007) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* **23**(7). <https://doi.org/10.18637/jss.v023.i07>
- Steier P, Dellinger F, Kutschera W, Priller A, Rom W and Wild EM (2004) Pushing the precision limit of ^{14}C AMS. *Radiocarbon* **46**(1), 5–16.
- Stuiver M and Becker B (1986) High-precision decadal calibration of the radiocarbon time scale, AD 1950–2500 BC. *Radiocarbon* **28**(2B), 863–910.
- Stuiver M and Polach HA (1977) Discussion reporting of ^{14}C data. *Radiocarbon* **19**(3), 355–363.
- Tjur T (1998) Nonlinear regression, quasi likelihood, and overdispersion in generalized linear models. *The American Statistician* **52**(3), 222–227.
- Turney SJ, Grabowski KS, Knies DL and Mignerey AC (2004) Radiocarbon data collection, filtering and analysis at the NRL TEAMS facility. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **223–224**, 216–220. <https://doi.org/10.1016/j.nimb.2004.04.044>
- Vogel JS, Ognibene T, Palmblad M and Reimer P (2004) Counting statistics and ion interval density in AMS. *Radiocarbon* **46**(3), 1103–1109.
- Wendt I and Carl C (1991) The statistical distribution of the mean squared weighted deviation. *Chemical Geology: Isotope Geoscience Section* **86**(4), 275–285.