1 11

A simulation study to compare Pb dating data analyses

Marco A Aquino-López[[1]](#footnote-1)1 [[2]](#footnote-2)2

Nicole K. Sanderson[[3]](#footnote-3)3

Maarten Blaauw[[4]](#footnote-4)4

J Andrés Christen[[5]](#footnote-5)5

Joan-Albert Sanchez-Cabeza[[6]](#footnote-6)6

Ana Carolina Ruiz-Fernandez[[7]](#footnote-7)7

01

**A simulation study to compare Pb dating data analyses**

/

Abstract

The increasing interest to understand anthropogenic impacts on the environment have led to a considerable amount of studies that focus on sedimentary records of ∼ 100 - 200 years. Dating this period is often complicated by the poor resolution and large errors associated with radiocarbon (14C) ages, which is the most popular dating technique. Instead, sediment dating with lead-210 (Pb) is widely used it provides absolute and continuous dates for ∼ 100 – 150 years. The Pb dating method has traditionally relied on the Constant Rate of Supply (CRS, also known as Constant Flux - CF) model which uses the radioactive decay equation as a age-depth relationship resulting in a restrictive model to approximate dates. In this work, we compare the classical approach to Pb dating (CRS) and its Bayesian alternative (*Plum*). For this, we created simulated Pb profiles following three different sedimentation processes, complying the assumptions imposed by the CRS model, and analysed them with both approaches. Results indicate that the CRS model does not capture the true values even when the sediment is heavily dated, nor improves its accuracy as more information is available. On the other hand, the Bayesian alternative (Plum) provides consistently accurate results even with few samples, and its accuracy and precision constantly improves as more information is available.

*Keywords:* Plum, Age-depth models, Chronology, Constant Rate of Supply, Comparison.

1.45

## 1 Introduction

Lead-210 (Pb) is a radionuclide, part of the U decay chain, which forms naturally in the atmosphere as well as in sediments. This isotope, with a half-life of 22.23±0.12 years, is commonly used to date recent recently accumulated sediments (<150years). In recent decades, increasing amounts of palaeoecological and pollution studies have focused on these recent sediments (e.g., Mustaphi et al., 2019) in order to evaluate human impacts on the environment. Unlike to other dating techniques such as C (radiocarbon dating), dating sediment is impossible from single measurement of Pb; it is only when a suitable portion of the excess-Pb (atmospheric Pb) decay curve (total inventory) is measured and with certain assumptions about the sedimentation process are met that a chronology can be established. These studies strongly rely on the accuracy of their chronologies in order to correctly assign dates to chemical, biological and ecological changes. That is, unlike other dating techniques, an analysis of a series (data set) of Pb measurements must be carried out in order to obtain meaningful dates. In a lake sediment, or any other, sedimentation process, samples are taken along a core at different depths, from which Pb activity is measured. The whole series of Pb measurements need to be analyzed to attempt to produce a coherent chronology, see Aquino-López et al. (2018).

A range of traditional data analyses, or so called “models”, are available for dating recent sediment using Pb; e.g. the Constant Initial Concentration (CIC, Goldberg, 1963), also known as Constant Activity (CA, Robbins and Edgington, 1975), the Constant Flux : Constant sedimentation (CFCS, Crozaz et al., 1964) and the Constant Rate of Supply (CRS, Appleby and Oldfield, 1978; Robbins, 1978; Sanchez-Cabeza and Ruiz-Fernández, 2012) also known as the Constant Flux model (CF). The CRS model is by far the most popular (see Figure 1) and has the most flexible assumptions. It assumes a constant supply of atmospheric Pb (also known as excess Pb) to the sediment and allows for changes in the sedimentation rate. In order to build a chronology, the CRS model uses a ratio between the complete “inventory” (the excess activity accumulated in the sediment column, between the surface and the equilibrium depth, where excess Pb can no longer be found) and the remaining inventory from depth *x* to the previously defined equilibrium depth, (, where is the complete inventory and λ the decay constant of the Pb ≈0.03118±0.00017 yr).

[Sorry. Ignored \begin{centering} ... \end{centering}]

Other, more restrictive, models such as CFCS and CIC also require the assumption of a constant flux of Pb as well as other assumptions of the sedimentation process, as well as that of a constant supply of Pb. The flexibility of the CRS model, regarding its assumptions, comes at the cost of the need to measure a sufficient portion of the inventory or the use of interpolation in order to properly estimate the complete inventory of Pb in the sediment.

The CRS model has undergone several revisions in the last decade in order to improve its applicability and precision. There are two types of revisions to this model: (1) revisions to its uncertainty quantification (eg. Binford, 1990; Appleby, 2001; Sanchez-Cabeza et al., 2014) and (2) to its application where extra information is available, such as external independent dating markers (e.g. Cs profiles), laminated sediments, tephras, contaminated layers (known sedimentary events) (eg. Appleby, 1998, 2001, 2008).

A recent inter-laboratory model comparison experiment (Barsanti et al., 2020) presented concerning results. A series of Pb measurements were send to 14 laboratories around the world. Each laboratory was ask to provide a chronology, given the same data, it is important to note that each laboratory applied their preferred model, in most cases the CRS model was calculated. This experiment resulted in a wide range of chronologies, independently of the model used, providing different chronologies even when the same model and dataset was used. The authors reinforced the need to use of independent time markers (independent dating sources) to validate of the chronologies, as suggested previously by (Smith, 2001). This comparison experiment clearly and critically shows the impact that user decisions have on the resulting chronologies. In order to replicate and/or update any given chronology these user decisions becomes extremely important. Additionally to the user’s decisions the raw data is also required; unfortunately, both the raw data sets and/or user’s decisions are rarely reported.

Recently Aquino-López et al. (2018) presented an alternative to these classical models, by introducing *Plum*, a Bayesian approach to Pb dating. This model treats every data point as originating from a forward model that includes both the sedimentation process and the radioactive decay process. *Plum* also uses an assumption of constant rate of supply to the sediment (this assumption can be relax at the cost of computational power), similar to the CRS model. Another important difference between the CRS and *Plum* is that this last one incorporates the supported Pb, which naturally forms in the sediment and is normally threaded as a hindrance variable. *Plum* assumes that there exists an (unknown) age-depth function *t*(*x*) that relates depth *x* with calendar age *t*(*x*). Conditional on *t*(*x*), the following model is assumed for the measured Pb between depths to

(1)

Here is the supported Pb in the sample and the supply of excess Pb to the sediment, the age-depth model *t*(*x*) is based on a piece-wise linear model constrained by prior information on the sediment’s accumulation rates (Blaauw and Christen, 2011), see Aquino-López et al. (2018) for details. This treatment of the data allows for a formal statistical inference on a well defined model with specific parameters. In order to infer the parameters of the model Bayesian approach used. This differs from the CRS model as the latter does not provide a formal statistical inference. This model uses the decay equation to obtain the age-depth function resulting in a more restrictive age-depth model and only dears with the excess Pb, the estimated supported Pb is previously removed before modelling. *Plum* has shown to provide accurate results with a realistic precision on different case scenarios (Aquino-López et al., 2018, 2020) - both in simulations as well as for real cores. Under optimal conditions *Plum* and the CRS model have shown to provide similar results (Aquino-López et al., 2020), with *Plum* providing more realistic uncertainties, with minimal user interaction.

Blaauw et al. (2018) presented a comparison between classical and Bayesian age-depth models construction, both for real and simulated C-dated cores. They concluded that Bayesian age-depth models provide a more accurate result and more realistic uncertainties under a wide range of scenarios. The objective of this study is to test whether the results obtained by Blaauw et al. (2018), concerning the accuracy and precision of the Bayesian approach, are maintained in a more complex modelling situation, such as the construction of Pb-based age-depth models. In this study, we compare Pb dates and uncertainties from the widely applied CRS model (by far the most popular age-depth model for Pb) against *Plum* using simulated cores, i.e. sedimentation “scenarios”. We also wish to observe the learning process of each of the models and estimate the amount of information is needed to obtained a reasonable chronology for each model.

Provided that the CRS model has several revisions, as shown by Barsanti et al. (2020), we decided to apply the original version of the equations provided by Appleby (2001), with its suggested error propagation calculation, we will call this version of the CRS the “classical implementation of the CRS" (CI-CRS). We acknowledge that this implementation may be the less suitable for some particular cases and then expert knowledge can greatly improve the precision and accuracy of the model, but this will reduce the impact of any particular implementation has on our results.

The paper is organized as follows: first section sets the tools for the comparison, it describes the simulations of the three different scenarios and we described a parameter which will facilitate the comparison called information percentage. Section 3 describes the comparison for both the overall chronologies and by single depths. Lastly section 4 presents the conclusions and discussion of the results obtained in section 3.

## 2 Experiment Setup: Simulations, Model Considerations

In order to observe the accuracy and precision of any chronology, a known true age-depth function is required. Blaauw et al. (2018) presented a methodology for simulating radiocarbon dates and their uncertainties, while Aquino-López et al. (2018) presented an approach for simulating Pb data given an age-depth function *t*(*x*). It is important to note that these simulations follow the equations presented by Appleby and Oldfield (1978); Robbins (1978) guaranteeing that the CRS assumptions are met. By using the approach presented by Aquino-López et al. (2018) for simulating Pb data and the structure of uncertainty quantification presented by Blaauw et al. (2018), reliable Pb simulated data can be obtained.

### 2.1 Simulation Construction

Three different scenarios (see Table 2.1) were chosen to simulate sedimentation processes, with their own age-depth functions and parameters. These scenarios were selected as they provide three key challenges for the models: Scenario 1 presents an age-depth function which is the result of increasing sedimentation and less compaction towards present (surface), this is quite common for recent sediments; Scenario 2 presents a challenging core structure as the function has a drastic and rapid shift in sediment accumulation around depth 15 cm depth; and lastly Scenario 3 presents a cyclic and periodic change in accumulation rates. Using the age-depth functions and parameters defined in Table 2.1, we obtain the Pb activity, or concentration, at any given depth or interval, by integrating the age-depth curve for that interval. Although these concentrations may be interpreted as error-free measurements (see Figure 2), we replicated the Pb activity uncertainty, following a similar methodology to Blaauw et al. (2018). This methodology was chosen as it introduces different sources of uncertainty related to different steps of the measurement process. Other uncertainty quantification methodologies could be used, but as longest the same methodology and uncertainty is provided to both models the comparison remains valid.

|  |  |  |  |
| --- | --- | --- | --- |
| Label | Age-depth | Φ | Supported Pb |
|  | function *t*(*x*) | () | () |
| Scenario 1 |  | 100 | 10 |
| Scenario 2 |  | 50 | 25 |
| Scenario 3 |  | 500 | 15 |

Table 1: Simulated age-depth function and parameters used in each scenario

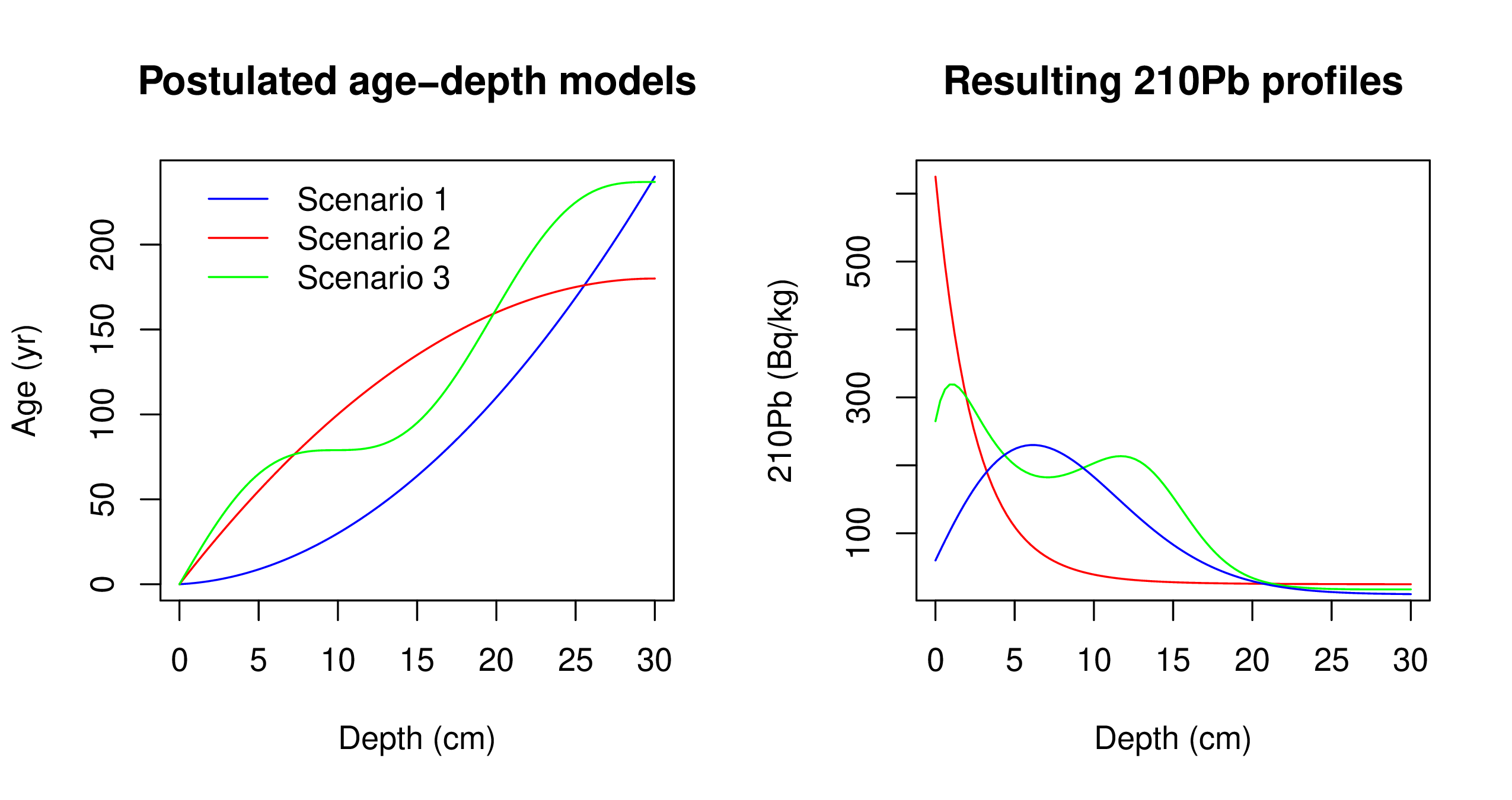


Figure 1: Simulated sedimentation scenarios with their corresponding Pb profiles. Left: Age-depth functions for the three different scenarios (Table 2.1). Right: Corresponding Pb activity profiles in relation to depth.

Let be the true Pb concentration in the interval ̂*x*=[*x*−δ,*x*), given the age-depth function *t*(*x*) and parameters Φ and in each scenario. To simulate disturbances in the material, we can introduce scatter centred around the true value, , where is the amount of scatter for this variable (in this case ). Now, to replicate outliers, a shift from the true value () is defined, which occurs with a probability . This results in a new variable Θ' which is defined as

[Sorry. Ignored \begin{cases} ... \end{cases}]

. (2)

Finally, to simulate the uncertainty provided by the laboratory, we can define the simulated measurements as , where is the standard deviation reported by the laboratory. is defined as , where is the minimum standard deviation assigned to a measurement. This variable differs between laboratories,we use a default value of 1 *Bq*/*kg*. Finally, ε is the analytical uncertainty (default .01) and an error multiplier (default 1.5). The default parameters were set in accordance with Blaauw et al. (2018).

For this this study we created a data set for each of the three simulation by integrating in intervals of δ=1 cm, for depths from 0 to 30 cm where radioactive equilibrium was guaranteed (Aquino-López et al., 2018). The complete simulated Pb data sets can be found in the Supplementary Material ?.

### 2.2 Model Considerations

In order to create a comparison with minimal user interaction each model will be run automatically. In the case of *Plum*, default settings will be used in order to minimize user interaction. As the CI-CRS model assumes that background has been reached, in order to reduce user manipulation, we decided to fix the last sample (30 cm depth) for every case. This step not only guarantees the consistent application of the CI-CRS model, it also provides the model with a single bottom-most depth to be removed as it is common practice when using the CI-CRS model. Because of this, *Plum*’s resulting chronology will always reach 30 cm and as default 1 cm sections will be used for every simulation. On the other hand, the CI-CRS model only models the excess Pb (the total Pb minus the supported Pb), because of this if certain excess activities fall below zero, the chronology will only be calculated up to such depth. Regarding the supported Pb, *Plum* deals with this variable automatically, as part of the inference, and it only requires a user decision if it should be calculated at every depth or as constant throughout the core. In order to provide the best possible estimate for this variable, for both models, a constant level of supported Pb was assumed. For the CI-CRS model, the mean of the supported Pb measurements was calculated and then subtracted from the total Pb to obtain the excess Pb, as it is common practice when using the CI-CRS model.

In order to provide an objective comparison, the offset of the true age-depth model (in yr), length of the 95% intervals (in yr) and normalized accuracy was calculated (the normalized offset indicating the distance of modelled ages from the true value given the model’s own uncertainty). The main discussion will revolve around the normalized offset as it provide a intuitive measure of the accuracy a model by taking into account the levels of uncertainty provided by each model.

## 3 Model Comparison

To allow for a reasonable comparison between models, and to evaluate the effect that different amount of information may have on the accuracy and precision of Pb models, we used our three simulated data sets (see Supplementary Material ?). For these simulated cores, samples were randomly created given a percentage of information (e.g. for a 20% information a dataset with 6 random 1-cm samples -of a possible total 30 1-cm samples- is created) in order to create a sub-dataset, which is then use to create a chronology. 100 of these sub-datasets were created for information percentages from 10% to 95% at 5% intervals (i.e., 10%, 15%, 20%,...,95%), the complete dataset was also used (i.e 100% percentage of information sample). After a dataset is created, both the CRS model and *Plum* were applied. Both sets of outputs were then compared against the true known age value, see Figure 3.

[Sorry. Ignored \begin{centering} ... \end{centering}]

Figure 3 shows an example of the comparison between the Pb models against the true value. As we are dealing with a total of *n*= 5333 simulations, in order to evaluate the overall precision and accuracy of both models, we decided to calculate the mean offset to the true age-depth model (in yr), the mean of length of the 95% intervals (in yr), as well as the mean normalized accuracy indicating the distance of modelled ages from the true value given the model’s own uncertainty at each depth.

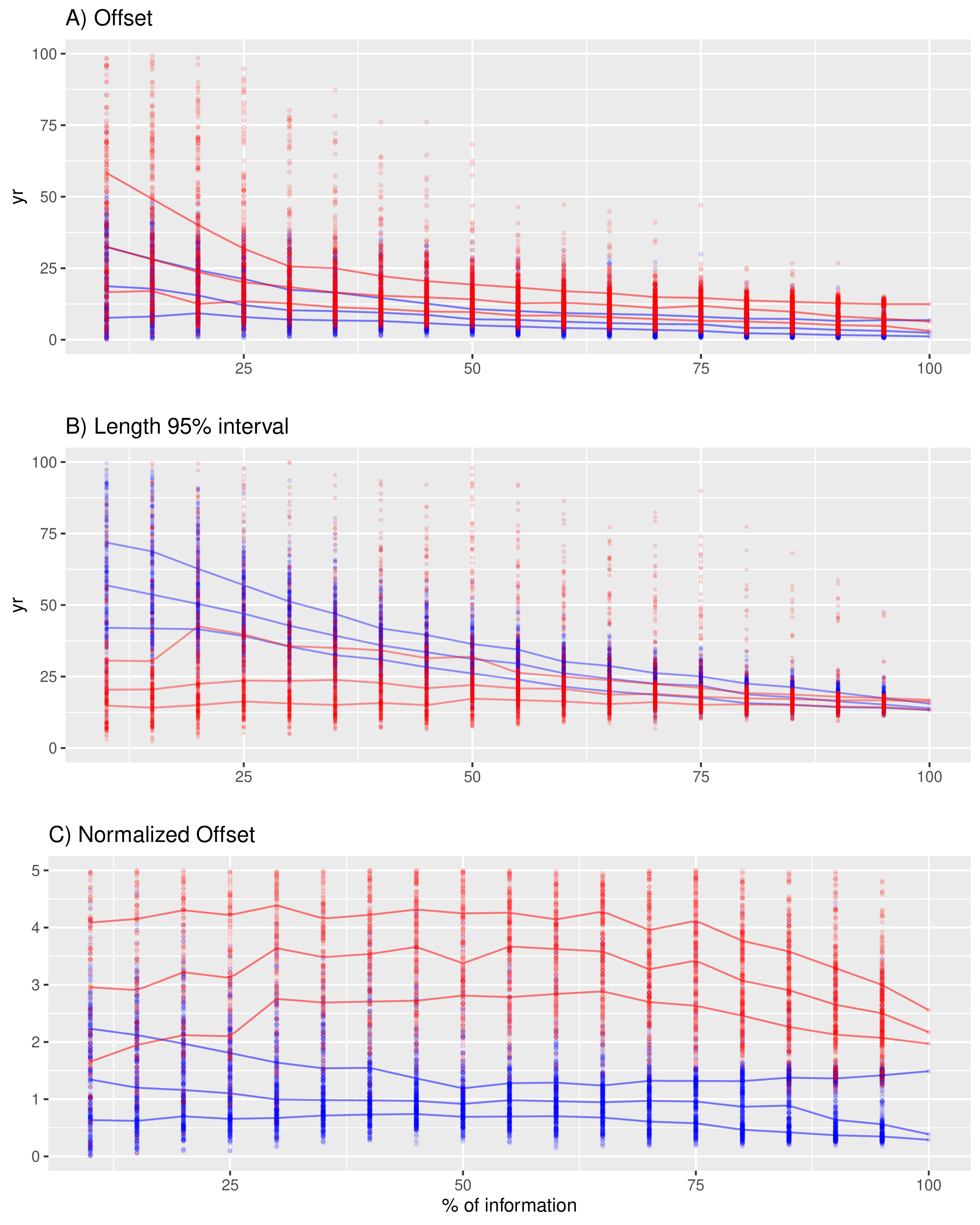


Figure 2: Top panel A) shows the offset between the modelled and true age of the CI-CRS (red) and *Plum* (blue). This panel shows how *Plum* provides a small offset in almost every scenario with both models improving their offset as more information is available. Middle panel B) shows the 95% confidence intervals. It is clear, from this panel, than the uncertainty provided by *Plum* is a lot bigger for low percentage of information and it constantly improves as more data is available, whereas the length of the intervals provided by the CI-CRS appear to stay constant regardless of the available information. Bottom panel C) shows the normalized offsets, presenting the distance between the modelled age and the true age normalized divided by the standard deviation (in the case of *Plum*, the length of the 95% interval divided by 4). This panel presents a worrying situation where the CI-CRS model’s calculated standard deviation (on average) is incapable of of capturing the true age. On the other hand, *Plum*’s credible intervals almost always capture the true age even when little information is available.

Figure 4 show results similar to those presented by Blaauw et al. (2018). The classical model (CI-CRS) at first appears to provide a similar results (similar offsets) to the Bayesian alternative (*Plum*), but at higher estimated precision (if we only look at the length of the 95% interval). It is important to note that the CI-CRS model’s offset improves as more information is available. Unfortunately, if we do not consider both the effects of both the offset and length of the interval together the results are not favorable to the CI-CRS. To have a more realistic representation of how the models capture the true age-depth models relationship, we should observe the normalized offset. This variable shows the degree the average models contain the truth within their uncertainty intervals (normalized to one standard deviation). Any model with a normalized offset larger than two (two standard deviations) is incapable of capturing the true ages within its uncertainty intervals. This means that, while the CI-CRS estimates smaller uncertainties and its ages improve as more data is available, it does so at the cost of its accuracy and the improvements are not sufficient to capture the true age. It also appears that the length of the 95% interval and offset are not affected by how much information is provided to the CRS model.

On the other hand, *Plum* seems to provide increasingly accurate results as more information is added to the model. This again coincides with the results outlined by Blaauw et al. (2018). When we observe the regular offset (not normalized), we find that *Plum* provides a smaller offset in comparison to the CI-CRS model; this in combination with slightly larger modelled uncertainties (more realistic) results in more consistently accurate age-depth model which are capable of capturing the true values within their uncertainty intervals. This result supports the claim that *Plum* provides more realistic uncertainties compared those obtained by the CI-CRS. Another important statistic to take into account is that 87.86% (4686/5333) of *Plum*’s runs remain within the 2 standard deviations, opposed to 7.48% (399/5333) for the CI-CRS model. Furthermore, only 0.54% (29/5333) of the CI-CRS model runs remain under the 1 standard deviation, which is the most commonly reported interval when reporting CI-CRS results. We can also observe a clear structure in the way that *Plum* increases its accuracy and precision to obtain a better chronology as more information is available, whereas the CI-CRS model does not appears to improve its capability of capturing the true value from additional data.

These results are valid for the overall chronology (the mean offset, interval and normalized offset of the overall chronology). In order to evaluate whether certain models are better predicting ages at certain section of the sediment cores, we have to look at the normalized offset of every depth.

[Sorry. Ignored \begin{centering} ... \end{centering}]

Figure 5 shows the normalized accuracy of every simulation according to depth for both models. *Plum* shows a clear learning structure which depends on the information available to the model. The information percentage appears to be irrelevant to the normalized accuracy of the CI-CRS model, contrary to the results obtained by *Plum*. It is important to note that the inaccuracies of the CI-CRS model are not exclusive to any particular sections of the chronology; this is most likely driven by the small uncertainties estimated by the CI-CRS model.See below for a discussion of how *Plum* behaved in sedimentation simulation 2.

## 4 Improvements to the CI-CRS

Since the 70’s, when the CRS method was first introduced (Appleby and Oldfield, 1978; Robbins, 1978), the CRS has received several improvements. Some of these improvements rely on independent dates, other isotopes or techniques, and/or require user manipulation to “force" the method to agree with these independent dates. One recent improvement, which does not requires user manipulation and/or independent dates, is the comprehensive explanation, with expert notes, on the practical used of the CRS model by Sanchez-Cabeza and Ruiz-Fernández (2012). Soon after the same authors presented an improvement to the uncertainty quantification of the age estimates by using the Monte Carlo method (Sanchez-Cabeza et al., 2014). As part of this publication, the authors made publicly available an Excel spreadsheet, which facilitate the calculation of their age estimates and improved uncertainties. Barsanti et al. (2020) showed that there exist several alterations and improvements to the CRS. Considering that this research focuses on the methods with minimal user manipulation, and given that these modifications and alterations are usually not made publicly available, an R implementation of the improved CRS by Sanchez-Cabeza et al. (2014), here labelled as revised CRS (R-CRS), was used to calculate a chronology of a particular, randomly selected, subsample with 95% of information and then compare against *Plum* and CI-CRS. The goal of this experiment is to quantify the improvement than these alterations have on the age estimates provided by the CRS methodology.

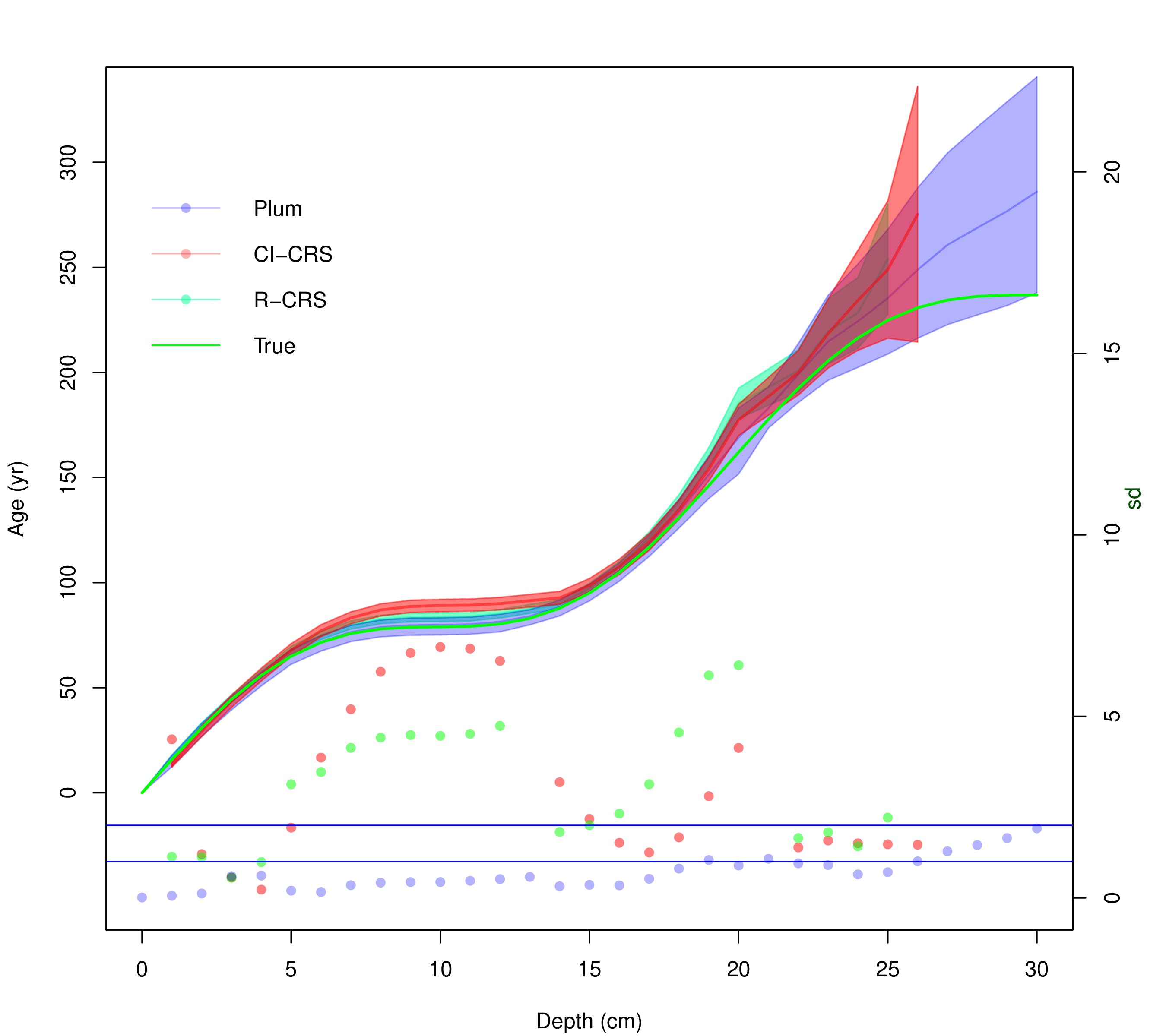


Figure 3: Comparison between the CI-CRS, R-CRS and *Plum*. .

Figure 6 shows the resulting chronologies for the three methodologies. From this figure it is clear that both versions of the CRS model (CI-CRS and R-CRS), with a good percentage of information, are capable of replicating the cyclic changes in accumulation that this scenario presents. It is also important to note that the offset of the R-CRS is improved throughout the whole chronology. The smaller uncertainties provided by the Monte Carlo method (Sanchez-Cabeza et al., 2014) appear to be the reason that the normalized offset, of the R-CRS, is bigger in some parts of the chronology, when compared to the CI-CRS.

The discussion on how large the standard deviation has to be in order to properly represent the uncertainty of the CRS model goes beyond the scope of this research. Nevertheless, it is important to point out that the realistic uncertainties, provided by *Plum*, are the result of a proper statistical inference where a likelihood is properly defined. In the case of the two version of the CRS model used in this research (CI-CRS and R-CRS) a proper likelihood is never defined and the uncertainties are calculated as a separate part of the model, either by error propagation or the Monte Carlo method. This could be the reason that the uncertainty quantification is not properly defined.

## 5 Discussion and Conclusions

This research focuses on exploring the uncertainty and precision of the most popular Pb dating methods (CI-CRS and *Plum*). By using different scenarios, three different simulations were created. These simulations were then sub-sample at different percentage of information in order to observe the effects that different sample sizes have on the resulting chronology. This experiment provided an objective comparison of the accuracy and precision of both methods.

The experiment was measured in two different levels: (1) the overall accuracy and precision of the method. For this level, the mean of the; offset, length of the 95% confidence and credible interval as well as normalized offset was measure. (2) The second level focused on the capability of the method to capture the true value in their credible/confidence interval. To measure this level the normalized offset of every scenario per depth was calculated. These two levels provided a good picture of the difference in precision and accuracy between these methods.

From the overall accuracy (see Figure 4) it is clear that both methods reduce their offset as more data is available, with the Bayesian method providing, on average, smaller offset regardless of the sample size. On the subject of precision, the Bayesian method is providing much bigger uncertainties when small sample size are used. It is only with 60%, or more, of the information that the length of the intervals becomes comparable. This is a consequence of the linear interpolation, between data points, used by the CRS method and the Bayesian approach (*Plum*) using a proper statistical inference. As it is previously discussed by Aquino-López et al. (2020), the larger uncertainties provided by *Plum* are more realistic, and this experiment confirms it. Proof that these uncertainties are more sensible is that the length of the credible intervals becomes smaller as more data is available. On the other hand, the length of the confidence intervals provided by the classical model (CI-CRS) remain almost constant at any sample size. Lastly, the normalized offset, which shows the capability of the model to capture the true values within their intervals, shows that the classical model (CI-CRS) on average is incapable of capturing the true values within its 95% confidence interval. This results is alarming as the Pb dating community rarely report 95% confidence intervals and instead 65% confidence intervals (one standard deviation intervals) are reported. On the other hand, *Plum*’s normalized offset always remains under two, therefore guaranteeing that on average the true value is capture within its 95% credible intervals, even with small sample sizes. *Plum*’s normalized offset, for small sample size, is larger and it becomes much more stable between 50% and 60% of information percentage. This experiments shows that the Bayesian method, on average, provides more reliable results. These results coincide with those obtained by Blaauw et al. (2018), where Bayesian methods provide more accurate results on the overall chronology when compared to their classical counterpart.

Because the normalized offset shows the capability of capturing the true value within its intervals, this variable can be used to conclude if any given method is better at estimating certain time period. Figure 5, presents the performance of both the CI-CRS mode and *Plum* at every simulated scenario. It appears that, the normalized offset of many of the CI-CRS chronologies are bigger than two throughout the whole chronology, this means that the model does not have a period of time at which is more precise. Moreover, the model presents a chaotic structure, where the normalized offset appears to be indifferent to the information available. It appears that even high levels of information percentage provide normalized offsets bigger than two, in some cases closer to four for scenario 2 and 3. *Plum* on the other hand, shows a structure where more data is synonymous of a better model in scenarios 1 and 3. It is only in low levels of information where *Plum*’s normalized offset is bigger than two. Scenario 2 on the other hand, presents and example where *Plum* is both incapable of capturing the true value, for depths deeper than 15 cm, and it appears that as more data is available the model provides worse results. This may be of concern if we do not recognized that this scenario is unrealistic and presents a extreme change in the accumulation around 15 cm, which coincides with the depth at which the normalized offset becomes bigger than two. It is also important to acknowledge that this experiment was performed using default settings, in a real world scenario the user has a some prior knowledge of the sedimentation process, in the site of interest, which is prior information and can be incorporated to the model to improve the resulting chronology.

The results obtained by this experiment appear to persist even in the case of the improved version of the CRS model (R-CRS). The R-CRS model appears to improve the offset but this improvement appears to be nullify by the smaller uncertainties presented by Sanchez-Cabeza et al. (2014). The question of which version of the CRS provides the best result is beyond the scope of this research. Nevertheless, it is important to note that that the offset, related to the CI-CRS and R-CRS, are reasonably small at certain sections of the sediment, but the uncertainty quantification of both methods is insufficient. This is the results of a non-statistical quantification of the model’s uncertainty.

In conclusion, the use of the Bayesian age-depth models is preferred on the construction of sediment chronologies, not only on radiocarbon based chronologies as presented by Blaauw et al. (2018) but also in the more complex case of Pb as demonstrated by this research. The classical approach provides reasonable results, in function of the offset, unfortunately the uncertainty quantification in these methods needs improvements as they do not relied on a proper statistic structure. Unfortunately in a real case scenario, it is impossible to measure the true offset of a method and this is why a proper uncertainty quantification becomes extremely important. These results support the recommendations presented by Smith (2001); Barsanti et al. (2020) where the CRS method, or any dating methodology, should be validated using independent dating methods.

Lastly, it is important to highlight the benefits of the Bayesian methods. From both Blaauw et al. (2018) and the present work, it is shown that Bayesian methods constantly improve as more data is available, the uncertainty associated to the method is realistic and coherent with the amount of information available, this leads to a method which is able to capture the true age in their credible intervals. The capability of capturing the true value in the credible intervals becomes important when the problem is associated with decision making processes, as it provides a more realistic picture of the available knowledge of the process. Provided that Pb is used in pollution, environment and climate change studies, which has a high impact on both policy making and public perception, realistic uncertainties become extremely important.

## 6 Acknowledgments

The authors are partially founded by CONACYT CB-2016-01-284451 and COVID19 312772 grants and a RDCOMM grant.

The corresponding author is founded by CONACYT through the postdoctoral residence program with CVU 489201.

References

Appleby, P. (2008). Three decades of dating recent sediments by fallout radionuclides: a review. *The Holocene*, 18(1):83–93.

Appleby, P. and Oldfield, F. (1978). The calculation of lead-210 dates assuming a constant rate of supply of unsupported 210pb to the sediment. *Catena*, 5(1):1–8.

Appleby, P. G. (1998). Dating recent sediments by Pb-210: Problems and solutions. *Proc. 2nd NKS/EKO-1 Seminar, Helsinki, 2-4 April 1997, STUK, Helsinki*, pages 7–24.

Appleby, P. G. (2001). Chronostratigraphic techniques in recent sediments. *Tracking Environmental Change Using Lake Sediments: Basin Analysis, Coring, and Chronological Techniques*, pages "171–203".

Aquino-López, M. A., Blaauw, M., Christen, J. A., and Sanderson, N. K. (2018). Bayesian analysis of 210pb dating. *Journal of Agricultural, Biological and Environmental Statistics*, 23(3):317–333.

Aquino-López, M. A., Ruiz-Fernández, A. C., Blaauw, M., and Sanchez-Cabeza, J.-A. (2020). Comparing classical and bayesian 210pb dating models in human-impacted aquatic environments. *Quaternary Geochronology*, 60:101106.

Barsanti, M., Garcia-Tenorio, R., Schirone, A., Rozmaric, M., Ruiz-Fernández, A., Sanchez-Cabeza, J., Delbono, I., Conte, F., Godoy, J. D. O., Heijnis, H., Eriksson, M., Hatje, V., Laissaoui, A., Nguyen, H., Okuku, E., Al-Rousan, S. A., Uddin, S., Yii, M., and Osvath, I. (2020). Challenges and limitations of the 210pb sediment dating method: Results from an IAEA modelling interlaboratory comparison exercise. *Quaternary Geochronology*, 59:101093.

Binford, M. W. (1990). Calculation and uncertainty analysis of dates for PIRLA project lake sediment cores. *Journal of Paleolimnology*, 3:253–267.

Blaauw, M. and Christen, J. A. (2011). Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis*, 6(3):457–474.

Blaauw, M., Christen, J. A., Bennett, K., and Reimer, P. J. (2018). Double the dates and go for Bayes — impacts of model choice, dating density and quality on chronologies. *Quaternary Science Reviews*, 188:58–66.

Crozaz, G., Picciotto, E., and de Breuck, W. (1964). Antarctic snow chronology with pb210. *Journal of Geophysical Research*, 69(12):2597–2604.

Goldberg, E. D. (1963). Geochronology with pb-210. *Radioactive Dating*, pages 121–131.

Mustaphi, C. J. C., Brahney, J., Aquino-López, M. A., Goring, S., Orton, K., Noronha, A., Czaplewski, J., Asena, Q., Paton, S., and Brushworth, J. P. (2019). Guidelines for reporting and archiving 210pb sediment chronologies to improve fidelity and extend data lifecycle. *Quaternary Geochronology*, 52:77–87.

Robbins, J. (1978). Geochemical and geophysical applications of radioactive lead. *The biogeochemistry of lead in the environment*, pages 285–393.

Robbins, J. A. and Edgington, D. (1975). Determination of recent sedimentation rates in lake michigan using pb-210 and cs-137. *Geochimica et Cosmochimica Acta*, 39(3):285 – 304.

Sanchez-Cabeza, J. A. and Ruiz-Fernández, A. C. (2012). 210Pb sediment radiochronology: An integrated formulation and classification of dating models. *Geochimica et Cosmochimica Acta*, 82:183–200.

Sanchez-Cabeza, J. A., Ruiz-Fernández, A. C., Ontiveros-Cuadras, J. F., Pérez Bernal, L. H., and Olid, C. (2014). Monte Carlo uncertainty calculation of 210Pb chronologies and accumulation rates of sediments and peat bogs. *Quaternary Geochronology*, 23:80–93.

Smith, J. N. (2001). Why should we believe 210pb sediment geochronologies? *Journal of Environmental Radioactivity*, 55(2):121 – 123.

## 7 Supplementary Material

1. Centro de Investigación en Matemáticas (CIMAT), Jalisco s/n, Valenciana, 36023 Guanajuato, Gto, Mexico. email: aquino@cimat.mx [↑](#footnote-ref-1)
2. Corresponding author. [↑](#footnote-ref-2)
3. GEOTOP Research Centre, Université du Québec à Montréal, Montréal, Québec, H2X 3Y7, Canada. email: sanderson.nicole@uqam.ca [↑](#footnote-ref-3)
4. School of Natural and Built Environment, Queen’s University Belfast, Belfast, BT7-1NN, UK. email:maarten.blaauw@qub.ac.uk [↑](#footnote-ref-4)
5. Centro de Investigación en Matemáticas (CIMAT), Jalisco s/n, Valenciana, 36023 Guanajuato, Gto, Mexico. email: jac@cimat.mx [↑](#footnote-ref-5)
6. Unidad Académica de Mazatlán, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de Mexico, 82040 Mazatlán, México [↑](#footnote-ref-6)
7. Unidad Académica de Mazatlán, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de Mexico, 82040 Mazatlán, México [↑](#footnote-ref-7)