Model Fitting and Estimating Uncertainties

Zarak Khan

Abstract

This report focused on the modelling of radioactive decay from a single and multiple isotope source. The goals were to highlight the importance of model fitting and explore different methods of estimating uncertainties in model parameters. The χ^2 method was used to apply exponential decay fits to the data, whilst uncertainties were estimated using either bootstrapping, the χ^2 grid method or Markov Chain Monte Carlo (MCMC) method. Model parameters corresponded to the initial activities and decay constants of the isotopes under investigation. Half lives were determined from decay constants, enabling identification said isotopes. For the multiple isotope sample, a single isotope and dual isotope model were fit to the data. Various statistics such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were calculated to determine which model provided the better fit, serving as a method of inferring how many isotopes were present in the sample. Results from the single isotope sample showed a half life of $12.04 \pm 0.13s$, enabling its identification as ^{138}Eu. From the multiple isotope sample, the dual isotope model provided a better fit, as indicated by AIC and BIC values of 122 and 133 versus 688 and 693 for the single isotope model. Half lives for the individual isotopes were determined to be $19.53 \pm [1.97, 2.25]s$ and $157.91 \pm [14.89, 13.81]s$, enabling their identification as either ^{166}W or ^{139}Eu for the first isotope, and ^{137}Ba, ^{121}Cs or ^{143}Eu for the second isotope. Exact identification was not possible given the breadth of the half life error ranges. Results pertaining to the identification of isotopes from parameters of the dual isotope model will also need to be interpreted with caution given the high degree of correlation between parameters. Finally, by determining the power law index of radiation as a function of distance to be roughly -2, radiation was shown to be emitted isotropically.

Introduction

A fundamental procedure in the analysis of any experimental dataset is model fitting, as the parameters of a well fitted model reveal relationships between quantities under investigation. Additionally, models can provide predictive power over the outcomes of unseen data.

Model selection is a complex problem, and typically requires formulating an initial hypothesis to predict expected relationships. If a selected model fits the data well, the initial hypothesis can be accepted with its presumed relationships seen as valid. It is therefore important to have a robust means of determining how well a model fits a dataset, and whether the initial hypothesis should be accepted or rejected.

As no dataset is free from error, the parameters determined from models will also have associated errors. Hence, in order to accurately the report results obtained from a fitted model, suitable errors ranges must be defined.

This report will focus on fitting models to datasets involving the measurement of radioactive decay counts as a function of time for unknown isotopes. As radioactive decay is known to exponentially decrease over time, a logical initial hypothesis is that an exponential decay function will fit the data well

$$A = A_0 e^{-\lambda t} \tag{1}$$

where A is the activity of the isotope, A_0 is the initial activity, λ is the decay constant and t is time. However, the number of isotopes present in the sample is unknown, necessitating the exploration of models which incorporate multiple isotopes.

$$A = \sum_{i=1}^{n} A_{0i} e^{-\lambda_i t} \tag{2}$$

The models will be fit using the χ^2 method, with the χ^2 statistic serving as an indicator of the 'goodness of fit'. Model selection will be further guided by the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Determining best fit parameters of the models will reveal the initial activities and decay constants of the isotopes, enabling identification. Uncertainties in the parameters will be obtained through either bootstrapping, the χ^2 grid method or the Markov Chain Monte Carlo (MCMC) method. These methods will be elaborated upon in the following sections.

Least Squares Fitting

The method of least squares is a standard procedure in regression analysis, used to find the best fitting curve through a set of data points. The goal is to minimize the sum of the squared differences between data points and respective values of a proposed model. For example, when fitting a linear model, this summation would be minimized

$$\sum_{i=1}^{n}[y_i - (mx_i + c)]^2 \tag{3}$$

where y_i and x_i correspond to values of the dependent and independent variables in the dataset, with m and c representing parameters of the linear model, the gradient and intercept respectively. To be more rigorous, the accuracy of each data point should be accounted for, with data points associated with a high degree of error being weighted less strongly in the summation, as shown below

$$\sum_{i=1}^{n}\frac{[y_i - (mx_i + c)]^2}{\sigma_i^2} = \chi^2 \tag{4}$$

where σ is the error on each data point. The summation of the squared differences between data points and predicted values of a model, weighted by the errors on each data point is the χ^2 statistic. This provides an indication of how well the model fits the data, with a lower value indicating a better fit. The fitting method which involves minimizing the χ^2 statistic is known as the χ^2 method, or weighted least squares method.

Parameters which minimize χ^2 for a linear model can be deduced analytically by setting their respective derivatives to zero and solving.

$$\frac{\partial}{\partial m}\sum_{i=1}^{n}\frac{[y_i - (mx_i + c)]^2}{\sigma_i^2} = 0 \tag{5}$$

$$\frac{\partial}{\partial c}\sum_{i=1}^{n}\frac{[y_i - (mx_i + c)]^2}{\sigma_i^2} = 0 \tag{6}$$

Solutions to these expressions will be shown in Appendix 1 for reference. Once the parameters that minimize χ^2 are found, their uncertainties need to be estimated, with a popular choice of method being bootstrapping.

Bootstrap Method

The bootstrap method is a resampling technique that enables the sampling distribution of a statistic to be estimated. It involves the repeated resampling with replacement of a dataset to provide new datasets from which the same statistic can be calculated, with the key idea being that the resampled datasets mimic variability in the original data. Following central limit theorum, for large numbers of resampled datasets, the sampling distribution of a statistic will approach a normal distribution and suitable confidence intervals can be determined.

In the case of model fitting, the sampling distributions of model parameters can be calculated, the mean and median values determined and an appropriate confidence interval can serve as an estimate of uncertainty.

Further methods of estimating the uncertainty in parameters include the χ^2 grid method and MCMC method which will be explained as follows.

χ^2 Grid Method

This method involves minimizing χ^2 values of a model over an n dimensional grid, where n is the number of parameters in the model. Suitable ranges of values for each of the parameters are defined beforehand, and χ^2 values for each combination of parameters are calculated. These are then displayed on the n dimensional grid.

With this method, the combination of parameters which give the lowest χ^2 value are the best fit parameters. The uncertainty in one parameter can be calculated by fixing the other parameters at their best fit values, varying the parameter in question through a suitable range and returning the χ^2 values. This will reduce the n dimensional grid to a 1 dimensional plot, and the uncertainty in the parameter in question can be estimated by defining an appropriate confidence interval. Confidence intervals are typically defined by the change in a parameter that causes a specific change in χ^2 value. For example, in a 2 parameter model where parameter 1 is fixed and parameter 2 is varied, the change in parameter 2 from its best fit value which causes a change in χ^2 of 2.3 defines the 68.3% confidence interval in parameter 2. This is illustrated in Figure 1 and a table of the changes in χ^2 value corresponding to different confidence intervals is provided in Appendix 1 for reference.

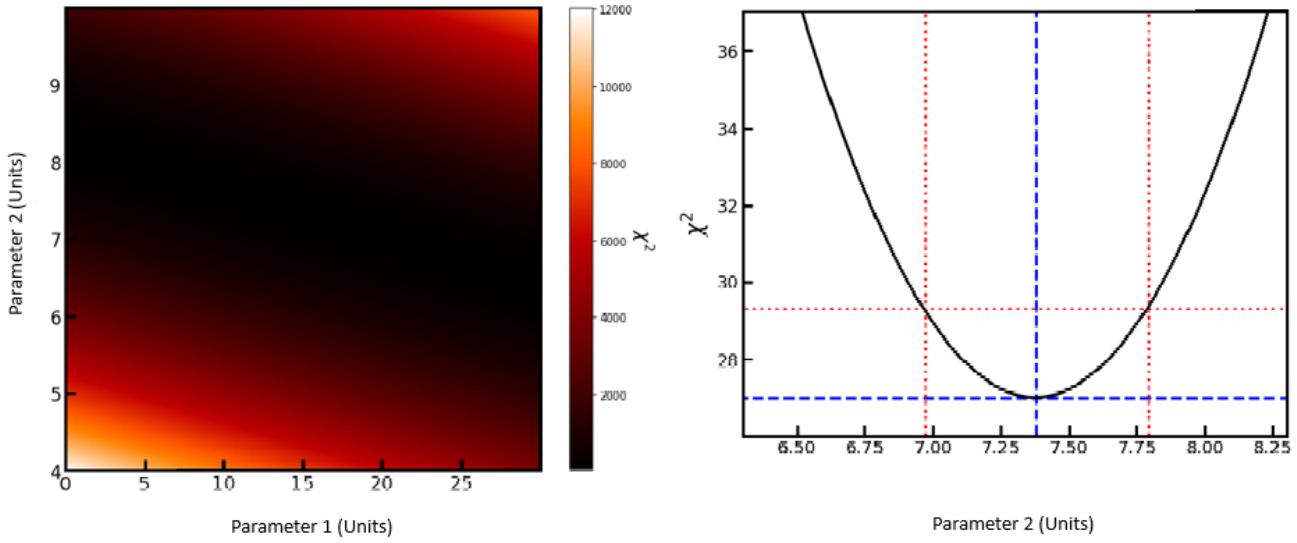


Figure 1: On the left is a 2D grid of χ^2 values for a 2 parameter model being fit to arbitrary data. The χ^2 values corresponding to each combination of parameters in specified ranges can be seen. On the right is a 1D plot with parameter 1 fixed at its best fit value and parameter 2 varied over a specified range with corresponding χ^2 values plotted. The horizontal blue line indicates the minimum χ^2 value, with the vertical line showing the corresponding value of parameter 2. The horizontal red line shows a change in χ^2 of 2.3 from the minimum, with vertical lines showing the corresponding values of parameter 2. The range between the two values of parameter 2 causing a change in χ^2 of 2.3 represents the 68.3% confidence interval.

This method can quickly become computationally demanding for more complex models, as the dimensionality of the grid will increase proportionate to the number of parameters. As the number of dimensions increases, the amount of combinations of parameter values for which the χ^2 statistic needs to be evaluated increases exponentially, and it is for this reason that the MCMC method is typically favoured over grid based approaches.

Markov Chain Monte Carlo Method

The MCMC method is a powerful computational technique that is used for sampling from complex probability distributions and enables efficient exploration of high-dimensional spaces. It is particularly useful in Bayesian statistics where it can approximate the posterior distribution of model parameters given the observed data. The idea is to construct a Markov chain that iteratively explores parameter space such that the stationary distribution of the chain converges upon the posterior distribution[1]. There are multiple algorithms that can be employed to achieve this, with the most common being the Metropolis-Hastings algorithm.

A brief overview of the Metropolis-Hastings algorithm:

1. An initial guess of fit parameters for the model is provided.

2. A likelihood function is defined and used to evaluate the probability of observing the experimental data based on a model incorporating the proposed parameters (calculating the χ^2 statistic can be used in this regard).
3. A new set of parameters is proposed using a predefined proposal function.
4. The likelihood of observing the experimental data with the new parameters is calculated.
5. If the likelihood is higher, then a new set of parameters is proposed and the chain jumps to a new position in parameter space.
6. Otherwise, the probability of jumping to a new position will be set by the proposal function. A random number is selected between 0 and 1 and if it is less than or equal to this probability, the chain will jump to a new position, otherwise it remains in place until the next iteration where the process repeats.
7. These steps are repeated a large number of times until the distribution converges.

This algorithm enables a random walk through parameter space, where the probability of moving to a new location is based only on information related to its current location (the χ^2 statistic in this case), hence the name Markov Chain Monte Carlo method. Depending on the initial guess and how well the proposal function is defined, the chain should spend the majority of its time at locations with high likelihoods, whilst still exploring areas with lower likelihoods. This enables an accurate approximation of the posterior distribution to be obtained. Statistics such as the mean and median of sampled parameters can be calculated to estimate best fit parameters, with suitable confidence intervals serving as uncertainties.

Once the best fit parameters of candidate models have been determined, it is necessary to assess which model provides the most accurate fit to the data. This is where the AIC and BIC become important, and will be discussed briefly as follows.

Akaike Information Criterion

Two issues that arise when fitting models are those of underfitting and overfitting. Underfitting occurs when a model is too simple to capture the underlying structure of the data. This can result in the model failing to accurately reveal relationships between quantities under investigation and generalizing to unseen data. Overfitting arises when a model is too complex, containing too many parameters. Although the model may fit the observed data well, it will also capture patterns within the noise or random fluctuations in the dataset. As a result, the model will fail to generalize to new datasets.

The AIC is used to estimate the amount of information lost when a given model is fitted to a dataset, providing a single numerical value for each model. However, it is important to note that the absolute value of the AIC does not have a meaningful interpretation, and only the relative differences in AIC values between models are significant, with lower values indicating a better model[2]. It is calculated as follows

$$AIC = 2k - 2ln(L) \tag{7}$$

where k is the number of parameters and L is the maximized likelihood function. A model will be penalized for either having too many parameters or being a poor fit to the data. By quantifying the trade-off between 'goodness of fit' and model complexity, a good estimate for the relative amount of information lost by a model can be achieved.

Bayesian Information Criterion

The BIC serves the same purpose as the AIC, and is calculated as follows

$$BIC = kln(n) - 2ln(L) \tag{8}$$

where k and L are as before, and n is the number of data points. Similar to the AIC, the absolute BIC values have no meaningful interpretation, and only the relative differences in BIC values between models matter, with lower values indicating a better model. The main difference between the AIC and BIC is that models with more parameters tend to be penalized more heavily by the BIC, making it more stringent in avoiding overfitting[3]. Nonetheless, both the AIC and BIC serve as effective tools in guiding model selection.

Method

Data in relation to the activity of a single radioactive isotope as a function of time was collected by 4 detectors, each at a distance of 0.05, 0.1, 0.18 and $0.3 \pm 0.0005m$ from the source. The detectors each had a cross section of $3.142 \pm 0.001\ cm^2$. Similar data was then collected using a source containing multiple radioactive isotopes. The methodology behind this data collection is not important, as the focus of this report is the statistical analysis and modelling of these datasets.

The objectives of analysis on the datasets relating to the single isotope source were to:

1. Determine the initial activity of the isotope at each of the 4 detectors.
2. Calculate the decay constant of the isotope.
3. From the decay constant, calculate the half life of the isotope.
4. Identify the isotope using its half life.
5. Determine whether the radiation is emitted isotropically
6. Estimate the number of particles in the isotope at $t = t_0$

In order to achieve these aims, a suitable model needs to be fitted to the data. As mentioned in the introduction, radioactivity is known to exponentially decay as a function of time, shown in equation 1. Taking the natural logarithm of this equation gives the following

$$lnA = -\lambda t + lnA_0 \tag{9}$$

where, as mentioned prior, A is the activity of the isotope, λ is the decay constant and A_0 is the initial activity. From equation 9, it can be seen that lnA and time have a linear relationship, with gradient equal to the decay constant and intercept equal to the natural logarithm of the isotopes initial activity. Hence, fitting a linear model to lnA vs t should enable the initial activity at each of the 4 detectors to be found, as well as the isotopes decay constant. A linear model was fitted using the χ^2 method, with best fit parameters determined via bootstrapping. Best fit parameters will be taken as the mean bootstrapped values with standard deviations serving as uncertainties.

The isotopes half life can be determined from the decay constant as follows

$$t_{\frac{1}{2}} = \frac{ln2}{\lambda} \tag{10}$$

with uncertainties in the decay constant being propagated into $t_{\frac{1}{2}}$ using the general principle

$$\sigma f(x) = |\frac{df(x)}{dx}|\sigma x \tag{11}$$

thus enabling the isotope to be identified.

To determine whether radiation is being emitted isotropically, the power law index of activity as a function of distance can be calculated using the following expression

$$LnA = xLnD \tag{12}$$

where x is the index and D is distance. From 12 it can be seen that the gradient of this plot will be the power law index. As before, a linear model was fitted to this data with best fit parameters and uncertainties determined via bootstrapping. For emission from a point source to be considered isotropic, this index is expected to be -2.

The activity of a radioactive isotope is directly proportional to the number of radioactive particles within the sample, as shown in the following expression

$$A = \lambda N \tag{13}$$

where N is the number of radioactive particles in the sample. Once the initial activity and decay constant are determined as detailed prior, the number of particles can be calculated at $t = t_0$ using equation 13. It is important to note that the isotopes activity is only being measured at the detector, whereas radiation is being emitted in all directions. Assuming radiation is emitted in a sphere with radius equal to the distance between source and detector, the number of particles calculated from equation 13 needs to be multiplied by the ratio of this spheres surface area with the detector cross section. This accounts for activity not being picked up by the detector and will provide a more reasonable estimate for the number of particles within the sample. Uncertainties from the initial activity, decay constant and detector cross section were propagated into the uncertainty for N using percentage error propagation.

The objectives of analysis on the datasets relating to the multiple isotope source were similar, only that a single isotope and dual isotope model were fitted to the data, again using the χ^2 method. Uncertainties in best fit parameters were determined using a χ^2 grid for the single isotope model and MCMC method for the dual isotope model. AIC and BIC values were calculated for both models to determine which provides the most accurate description of the observed data.

Results

Single Isotope Sample

This section will focus on the results obtained from analysing data from the 4 detectors using the single isotope sample. As mentioned in the Method section, the decay constant and initial activities at the 4 detectors were obtained by applying a linear fit to the natural logarithm of activity vs time. These results were input into a standard exponential decay function (equation 1) and overplotted on the original data, as shown in Figure 2. Residuals associated with each of these fits are shown in Figure 3, with decay constants and initial activities used in each of the fits shown in Table 1. Additionally, half lives, number of particles at $t = t_0$ and χ^2 values for each of the fits are shown in Table 1.

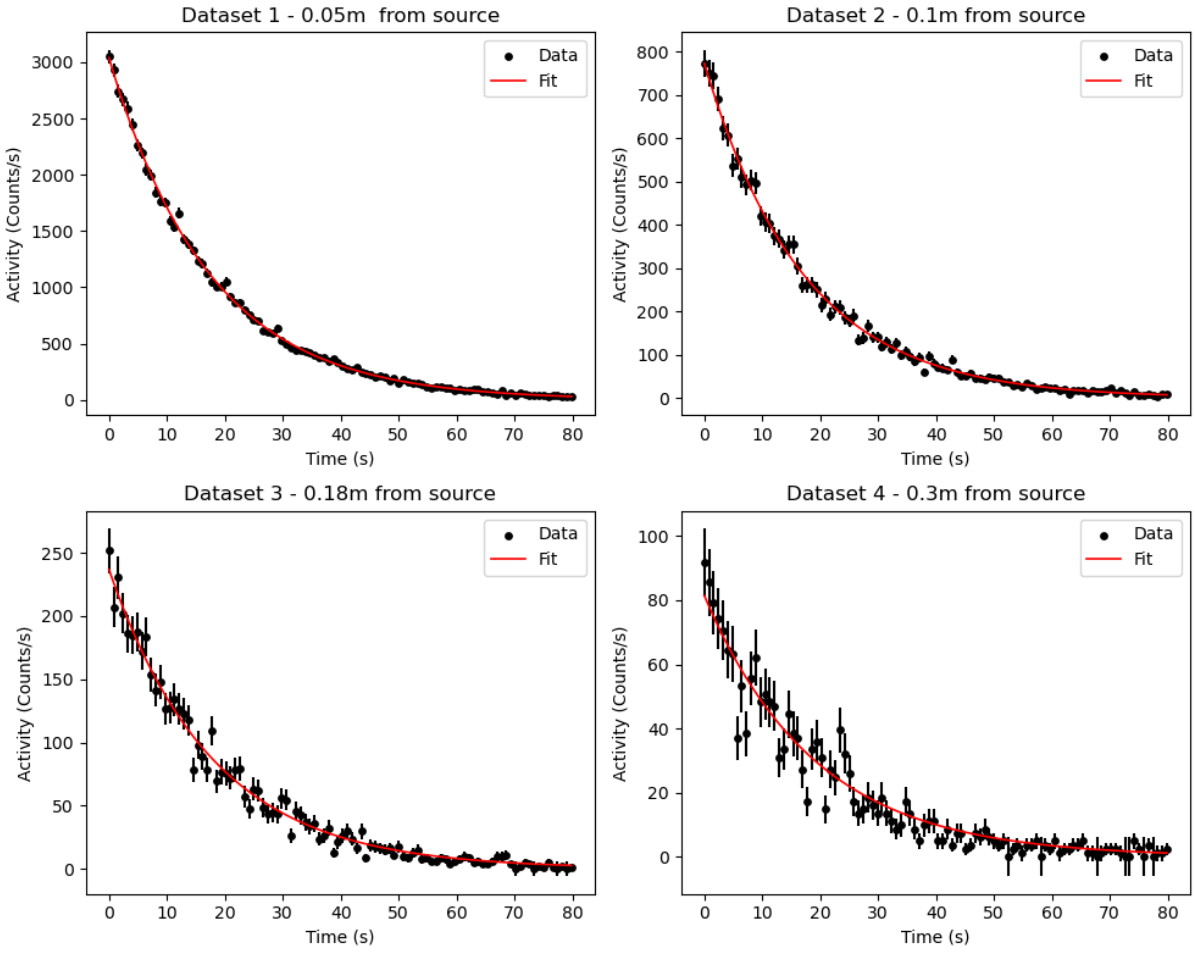


Figure 2: Exponential fits obtained using decay constants and initial activities from linear fit to lnA vs t, shown in red.

Assessing the plots in Figure 2 by eye show that an exponential decay model assuming a single isotope source fits reasonably well to the data, with even scattering of residuals around 0 counts/s (Figure 3) supporting this claim more quantitatively. Reduced χ^2 values of roughly 1 (Table 1) for each of the fits further supports this.

From dataset 1, the decay constant was calculated to be $0.05757 \pm 0.00006\ s^{-1}$, yielding a half life of $12.04 \pm 0.13\ s$. This enabled the isotope to be identified as ^{138}Eu from a list of candidates, provided in Appendix 1. Dataset 1 was primarily used for this identification given the higher quality of data, with results from the other 3 datasets serving as confirmation. The number of radioactive particles in the sample was estimated to be 5.38 ± 0.21 x10^6, with results from the other 3 datasets again agreeing within errors.

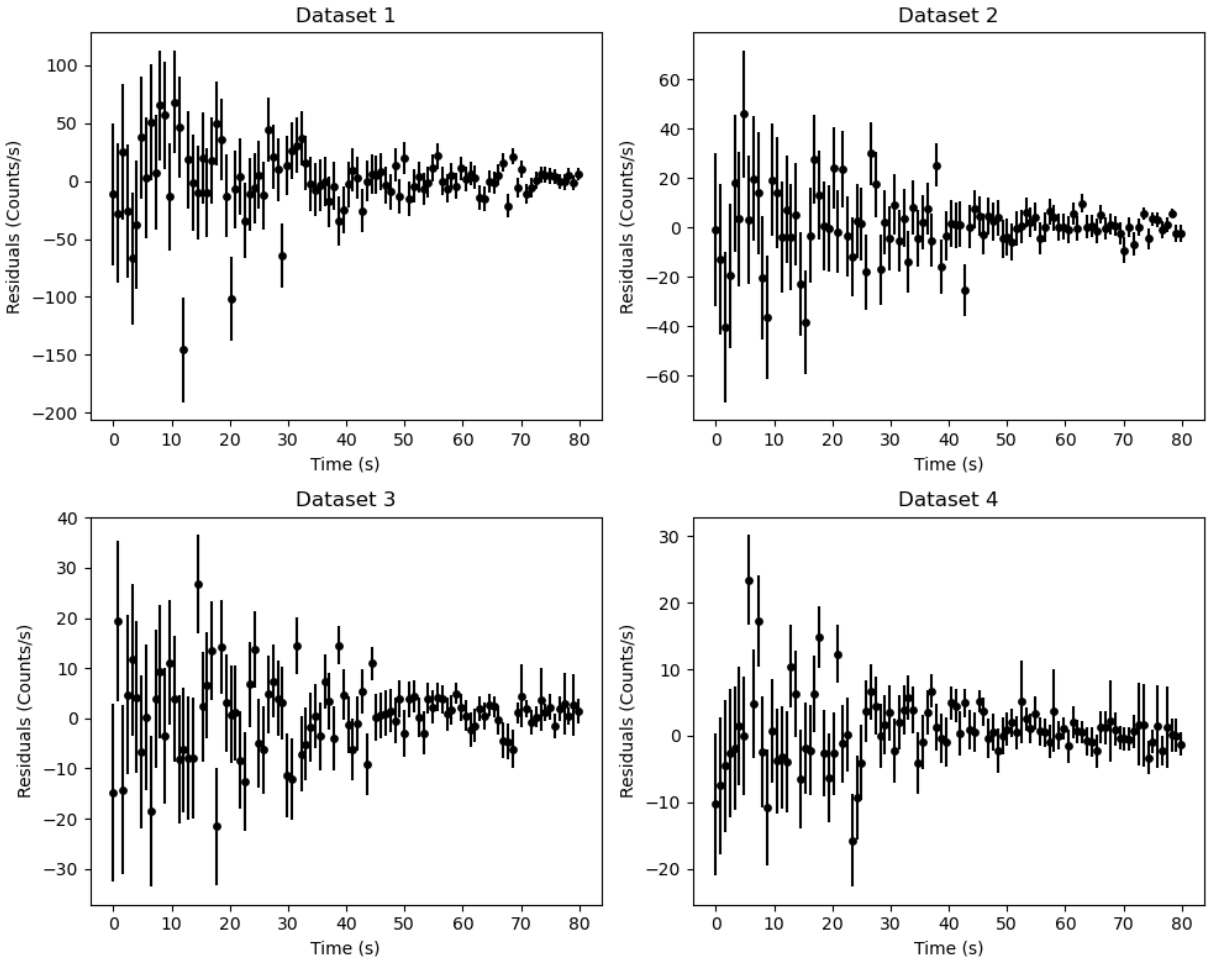


Figure 3: Residuals from exponential fits shown in Figure 2.

Table 1 – Decay constants, initial activities, half-lives, number of particles and χ^2 values calculated from exponential fits to single isotope data

Dataset	λ (Mean & St Dev) (s^{-1})	A_0 (Mean & St Dev) (Counts/s)	Half Life (s)	Number of Particles (x10^6)	χ^2 Value	Reduced χ^2 Value
1	0.05757 ± 0.00006	3077 ± 82	12.04 ± 0.13	5.38 ± 0.20	100	1.02
2	0.05757 ± 0.00143	795 ± 49	11.90 ± 0.29	5.49 ± 0.47	100	1.02
3	0.05593 ± 0.00485	293 ± 61	12.39 ± 1.08	6.82 ± 2.02	112	1.14
4	0.05230 ± 0.00554	109 ± 26	13.31 ± 1.42	7.59 ± 2.65	116	1.18

Multiple Isotope Sample

Figure 4 shows single and dual isotope models fitted to the 4 datasets from a multiple isotope source. Residuals for each of the models are shown in Figure 5. A table containing decay constants, initial activities and half lives of the isotopes

calculated with both models is shown in Appendix 2. Table 2 shows the χ^2, AIC and BIC values for the single and dual isotope models.

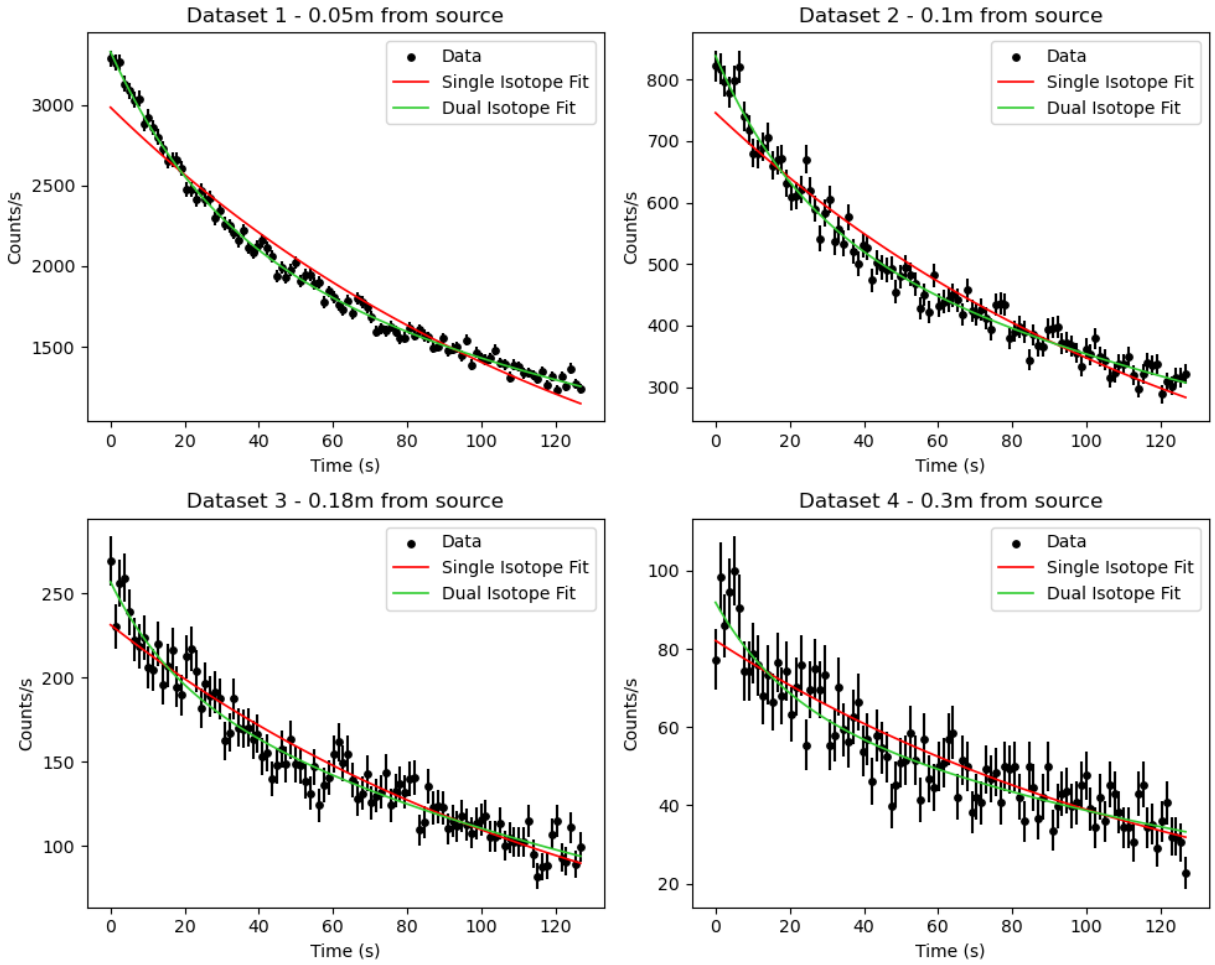


Figure 4: Exponential fits applied to datasets. The single isotope model is shown in red, and the dual model in green.

A cursory assessment of the plots in Figure 4 indicates the dual isotope model provides a better fit to the data. This is supported by the even scattering of residuals around 0 counts/s for the dual isotope model, whilst residuals for the single isotope model are not as evenly scattered (Figure 5). This is most apparent in dataset 1 which contains the highest quality data, reflected in the smaller error bars. AIC, BIC and χ^2 values are also lower for the dual model, further supporting its validity. It is important to note that the most significant changes in AIC and BIC values between models can be seen in dataset 1, where they drop roughly by a factor of 7 for the dual model, whilst remaining roughly similar in datasets 3 and 4. This is likely due to the higher scattering and errors in datasets 3 and 4, allowing the single isotope model to erroneously provide a good fit. Hence, the differences in AIC and BIC values between models for dataset 1 provide more accurate information to guide model selection.

From dataset 1, using the single isotope model, the decay constant was calculated to be $0.000753 \pm 0.00011 s^{-1}$ (full results shown in Appendix 2), with a corresponding half life of $92.07 \pm 1.30 s$. This was the best fit decay constant determined using the χ^2 grid method, with a 99% confidence interval quoted as the uncertainty (the χ^2 grid is shown in Appendix 3 alongside 1D plots for gradients and intercepts with 99% confidence intervals). These results are in agreement with the other 3 datasets within errors, however, this half life could not be matched to an isotope on the candidate list. This was to be expected given that the AIC and BIC values indicate this is not the most appropriate model for the data.

From dataset 1, using the dual model, decay constants for each of the isotopes were calculated to be $0.03548 \pm [0.00357, 0.00409] s^{-1}$ and $0.00439 \pm [0.00041, 0.00038] s^{-1}$ (full results shown in Appendix 2). These are the best fit decay constants obtained using the MCMC method, with 16 - 84% confidence intervals quoted as uncertainties. Corresponding half lives were determined to be $19.53 \pm [1.97, 2.25] s$ and $157.91 \pm [14.89, 13.81] s$, assisting in their identification, although associating each half life with a single isotope was not possible given the error ranges. The first isotope could be either ^{166}W or ^{139}Eu, as their respective half lives of $17.9 s$ and $19.2 s$ lie within the error range. Similarly, the second isotope could be identified as ^{137}Ba, ^{121}Cs or ^{143}Eu, with respective half lives of $153.1 s$, $154.8 s$ and $155.4 s$ lying within the error

range. The experiment would need to be repeated with a more accurate detection method in order to narrow this error range and enable the true isotopes to be discerned from these candidates. It should also be noted that there is conflict between results from different datasets, as some of the calculated half lives lie outside each others error ranges. This would also need to be addressed with more accurate measurement techniques.

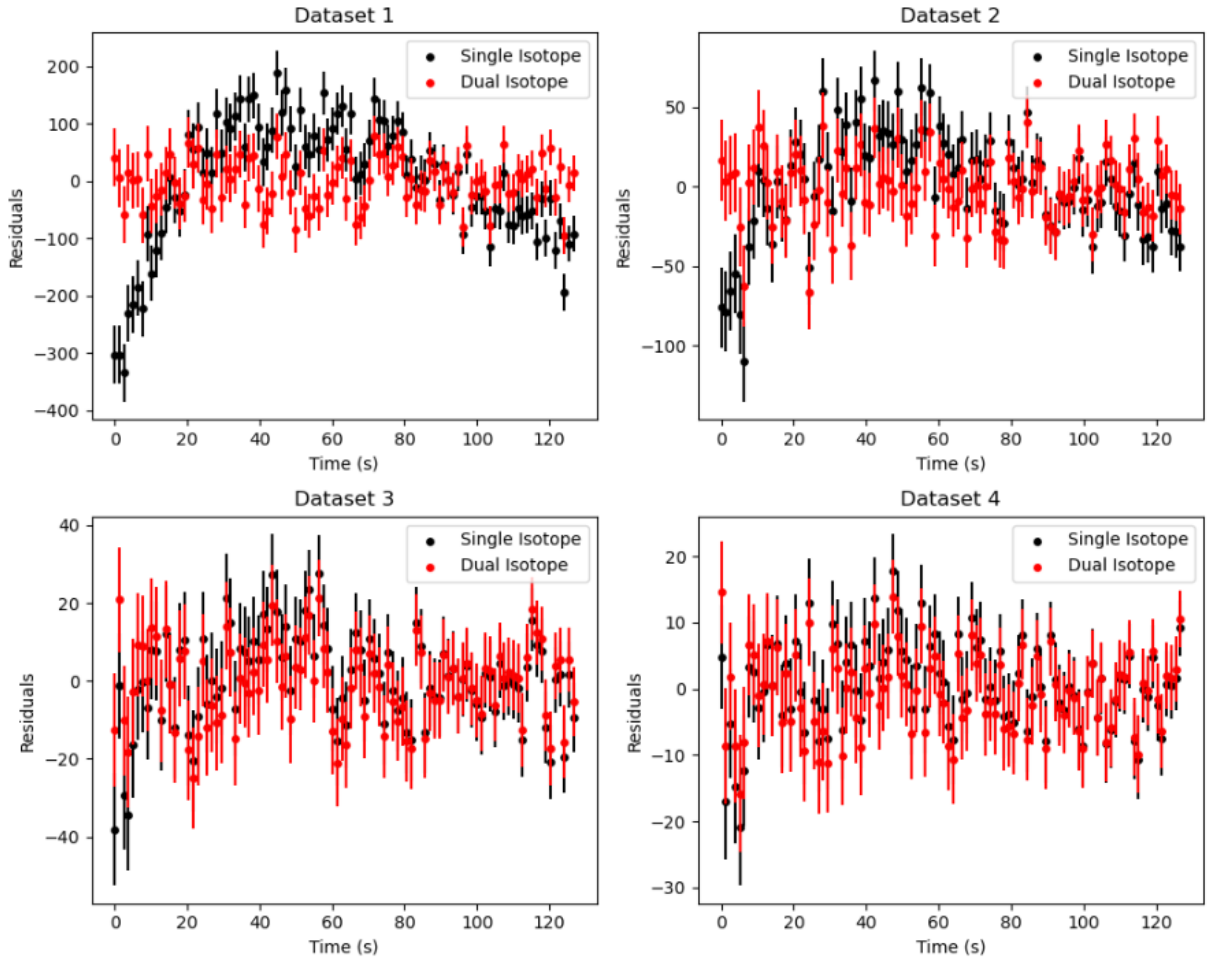


Figure 5: Residuals from exponential fits in Figure 4. In black are the residuals from the single isotope model, with residuals from the dual model in red.

Table 2 – χ^2 values, AIC and BIC for single and dual isotope models

Model	Single Isotope				Dual Isotope			
Dataset	1	2	3	4	1	2	3	4
χ^2 Value	683	255	116	103	114	110	87	91
AIC	688	259	120	107	122	118	95	99
BIC	693	265	125	112	133	128	106	109

Figure 6 shows a correlation plot for parameters resulting from MCMC analysis of dataset 1. Trace plots for each of the parameters are also provided in Appendix 4, where the even scattering of sampled parameters within a tight range suggests convergence. Additional statistics such as the Gelman-Rubin diagnostic can be calculated for a more quantitative assessment of chain convergence, although this is beyond the scope of this analysis. Histograms of parameters sampled by the chain are also shown in Appendix 4 for reference.

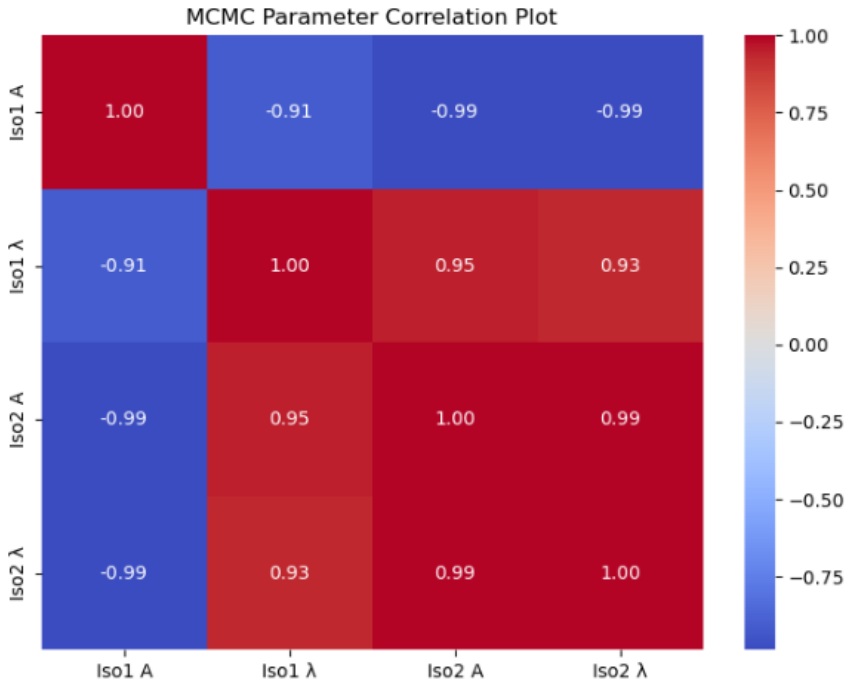


Figure 6: Correlation plot for parameters in MCMC analysis of dataset 1. λ and A labels refer to decay constant and initial activity parameters for each of the isotopes respectively. Each cell contains the Pearson correlation coefficient for the corresponding combination of parameters. The leading diagonal is trivially 1, as parameters correlate with themselves.

From Figure 6, it can be seen that the parameters used in this dual isotope model are heavily correlated or anti-correlated with one another. Although the overall fit to the data will not be affected, determining the individual impact of each parameter on the dependent variable can be difficult given this degree of correlation, making the interpretation of individual best fit parameters somewhat unreliable. Hence, the presence of two isotopes can be inferred using the dual isotope model, but results pertaining to the identification of each isotope will need to be interpreted with caution.

Isotropic Emission

Results for Ln Activity vs Ln Distance for the single and multiple isotope sources are shown in Figure 7.

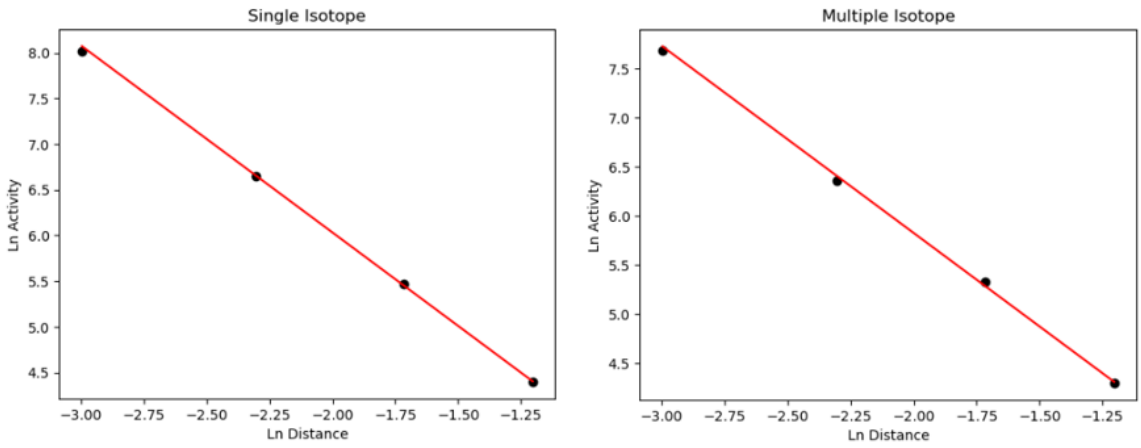


Figure 7: Ln Activity vs Ln Distance. Activities are the initial activities at each of the 4 detectors. Distances are the known distances between source and detector mentioned prior.

Linear fits were applied using the χ^2 method and gradients were calculated to be -2.05 $\pm$ 0.03 and -1.94 $\pm$ 0.07 for the single and multiple isotope sources respectively. Best fit gradients and uncertainties were determined via bootstrapping. These power law indices suggest that radiation is being emitted isotropically from the single and multiple isotope sources.

Conclusion

This report has demonstrated the utility of fitting well described models to experimental data. The importance of quantifying the relative validity of different models using statistics such as the AIC and BIC has been explained. Various techniques including the χ^2 grid and MCMC methods to estimate uncertainties in model parameters have also been explored, with their limitations discussed. The MCMC method was identified as the most effective technique given its computational efficiency whilst exploring higher dimensional parameter spaces.

A limitation in extracting accurate information from the dual isotope model was the high degree of correlation between parameters. Further research may need to be performed to identify methods of handling models with correlated parameters.

References

[1] Speagle, J. S. (2020). A Conceptual Introduction to Markov Chain Monte Carlo Methods. Retrieved from http://arxiv.org/abs/1909.12313

[2] Hossein Bonakdari and Zeynoddin, M. (2022). Goodness-of-fit & precision criteria. Elsevier eBooks, pp.187–264. doi.org/10.1016/b978-0-323-91748-3.00003-3.

[3] Shen, N. and Gonzalez, B. (2021). Bayesian Information Criterion for Linear Mixed-effects Models. Retrieved from arXiv:2104.14725v1 [stat.AP] 30 Apr 2021.

Appendix 1

Solutions to gradient and intercept giving minimized χ^2 value for linear fit:

$$m = \frac{SS_{xy} - S_x S_y}{SS_{xx} - S_x^2}$$

$$c = \frac{S_{xx} S_y - S_x S_{xy}}{SS_{xx} - S_x^2}$$

where $S = \sum_i \frac{1}{\sigma_i^2}$, $S_x = \sum_i \frac{x_i}{\sigma_i^2}$, $S_y = \sum_i \frac{y_i}{\sigma_i^2}$, $S_{xy} = \sum_i \frac{x_i y_i}{\sigma_i^2}$, $S_{xx} = \sum_i \frac{x_i^2}{\sigma_i^2}$

List of $\Delta\chi^2$ values as a function of confidence interval.

$\Delta\chi^2$ as a Function of Confidence Level p and Number of Parameters of Interest ν						
	ν					
p	1	2	3	4	5	6
68.27%	1.00	2.30	3.53	4.72	5.89	7.04
90%	2.71	4.61	6.25	7.78	9.24	10.6
95.45%	4.00	6.18	8.02	9.72	11.3	12.8
99%	6.63	9.21	11.3	13.3	15.1	16.8
99.73%	9.00	11.8	14.2	16.3	18.2	20.1
99.99%	15.1	18.4	21.1	23.5	25.7	27.9

List of candidate Isotopes

Isotope	Half-life time
^{120}Cs	61.2 s
^{121}Cs	2.58 min
^{124}Cs	30.9 sec
^{123}Ba	2.7 min
^{124}Ba	11.0 min
^{125}Ba	3.5 min
^{137m}Ba	2.552 min
^{137}Eu	8.4 sec
^{138}Eu	12.1 sec
^{139}Eu	17.9 sec
^{143}Eu	2.59 min
^{166}W	19.2 sec
^{168}W	51 sec
^{169}W	1.27 min
^{229}Np	4.0 min
^{230}Np	4.6 min
^{232}Np	14.7 min

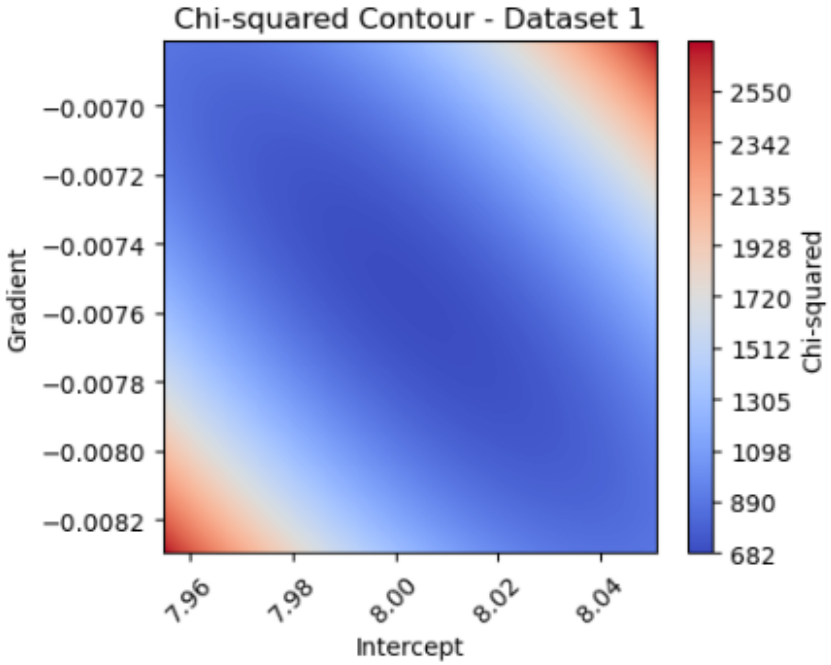
Appendix 2

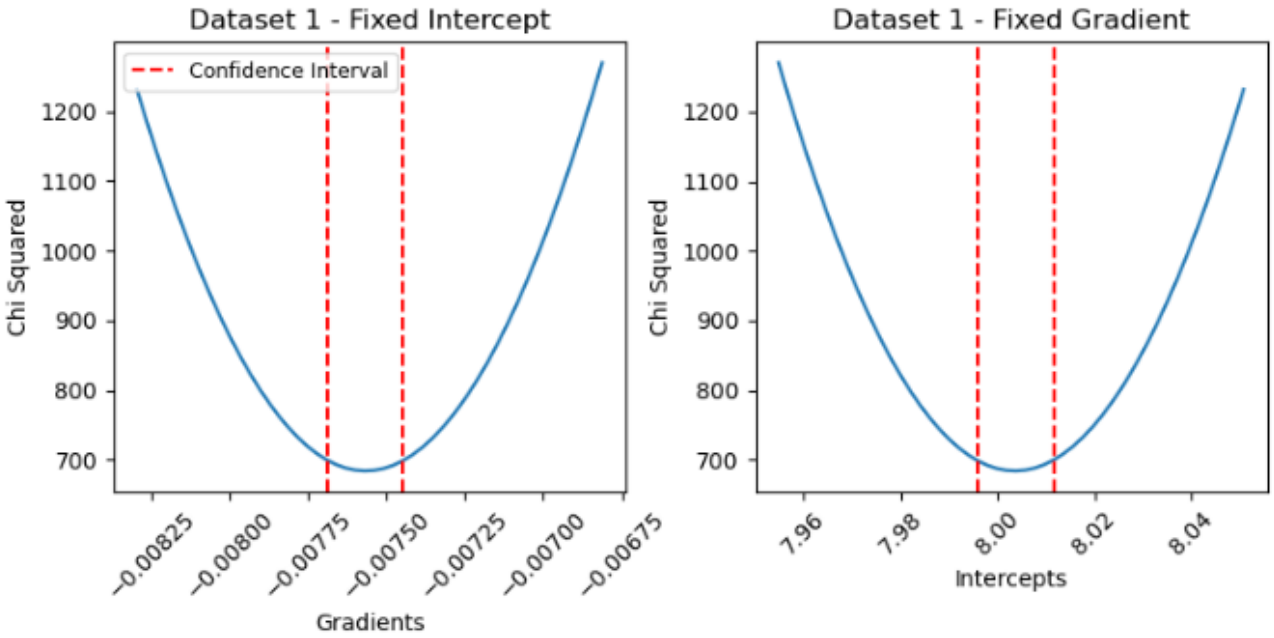
Full table of results for single and dual isotope fits to data from multiple isotope sample.

Model	Dataset	λ (Median and 16 – 84% CI) (s^{-1})	A_0 (Median and 16 – 84% CI) (Counts/s)	Half Life (s)
Single Isotope	1	0.00753 ± 0.00011	2983 ± 24	92.07 ± 1.30
	2	0.00764 ± 0.00013	745 ± 9	90.76 ± 1.50
	3	0.00748 ± 0.00013	231 ± 6	92.69 ± 1.57
	4	0.00748 ± 0.00013	82 ± 3	92.69 ± 1.57
Dual Isotope: Isotope 1	1	0.03548 ± [0.00357, 0.00409]	1159 ± [95, 103]	19.53 ± [1.97, 2.25]
	2	0.04419 ± [0.00972, 0.00872]	257 ± [30, 46]	15.68 ± [3.45, 3.09
	3	0.06490 ± [0.01952, 0.02364]	54 ± [11, 13]	10.68 ± [3.21, 3.89]
	4	0.05876 ± [0.03029, 0.02514]	23 ± [6, 13]	11.79 ± [6.08, 5.04]
Dual Isotope: Isotope 2	1	0.00439 ± [0.00041, 0.00038]	2167 ± [115, 107]	157.91 ± [14.89, 13.81]
	2	0.00510 ± [0.00081, 0.00048]	583 ± [55, 31]	135.79 ± [21.49, 12.71]
	3	0.00619 ± [0.00065, 0.00045]	204 ± [13, 8]	111.83 ± [11.87, 8.16]
	4	0.00585 ± [0.00212, 0.00077]	69 ± [17, 5]	118.44 ± [43.07, 15.53]

Appendix 3

χ^2 grid and 1D parameter plots for dataset 1 from multiple isotope source, modelled assuming a single isotope was present.





Appendix 4

Trace plots and histograms for dataset 1 MCMC chain with multiple isotope sample, modelled assuming two isotopes were present.