Date Evaluation: The Use of Bayesian Methodologies

Fiona Petchey
(University of Waikato, New Zealand; ARC Centre of Excellence for Australian Biodiversity and Heritage, Australia)

Radiocarbon dating is a widely used analytical technique to determine the age of objects and archaeological remains, but the interpretation of radiocarbon ages (^{14}C) is far from simple. A ^{14}C age, reported as a mean and standard error, is a probability distribution that scatters around the true age of the sample. This ^{14}C age is converted to a calendar date using a calibration curve independently developed from materials of known age (e.g. Intcal20; Reimer et al. 2020). These calibration curves also incorporate statistical scatter, and the converted calendar age is non-normally distributed (Figure 1). These factors complicate the interpretation and may lead to wrong conclusions when assessed by eye alone.

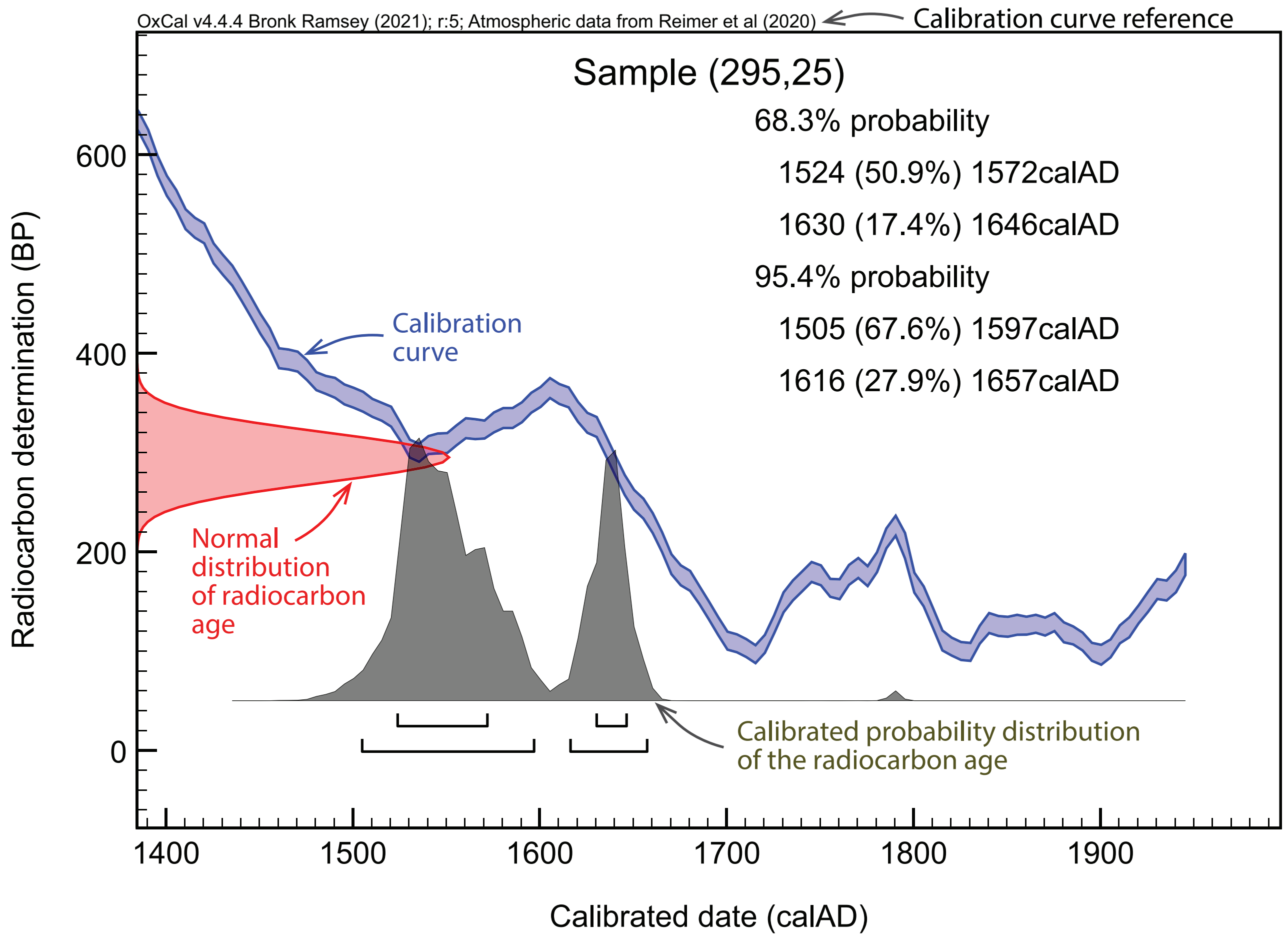


Figure 1. OxCal-generated plot of a calibrated radiocarbon age.

However, archaeologists are rarely interested in the age of a single object but are interested in answering questions such as when the activity started and ended and how long did an activity continue (Hamilton & Krus 2018). It is now common practice to use Bayesian statistical methodologies to develop chronological models that can provide answers to these questions. The core mathematics of Bayesian applications is Bayes's theorem:

Standardized likelihood (e.g. the dates) × Prior beliefs (e.g. the archaeological data and material-related uncertainties) = Posterior belief (an answer).

The use of Bayesian modelling allows the statistical scatter on the ^{14}C dates to be taken into account and integrates the dating with archaeological observations and other sources of information on the relative order of events. Importantly, the analysis calculates how successfully the ^{14}C measurements conform to the prior knowledge and narrow down the width of the calibrated dates according to the stratigraphic model. When correctly applied, Bayesian modelling can improve the resolution enabling the assessment of chronometric change within a single generation (e.g. Bayliss et al. 2007). In addition, the quantitative results allow comparison between sites, typologies, and culture-based models. A wide range of available program interfaces are available for different applications (e.g. Bacon: Blaauw & Christen (2011); BChron: Haslett & Parnell (2008); OxCal: Bronk Ramsey (1995, 2001)). Buck et al. (1991, 1994) provide a detailed discussion of the principles of Bayesian analysis in archaeological chronologies.

Basic Bayesian

Bayesian applications specialised for archaeological chronologies allow the development of models either constrained or unconstrained by stratigraphy. The two main building blocks are *Sequence*, where information on the order of events is available, and *Phase*, where there is evidence that the samples belong to the same event, but the order is unknown. These are separated by *Boundaries* that define when the activity starts and ends and are key to broader comparisons beyond the single-site model (Bronk Ramsey 2001, 2009a, 2009b).

Anderson and Petchey (2020) used the single phase option in OxCal to date the initial evidence of *kūmara* (sweet potato) gardening by Maori across different regions of New Zealand. They selected legacy ^{14}C data based on documented association with archaeological evidence of *kūmara* gardening. Figure 2 shows terrestrial dates associated with *kūmara* gardening in inland regions of the North Island. This model construction informs the computer program that at some point in time in the past horticultural activity started, went on for some unknown duration, and then ended. The resultant model enabled the quantitative comparison of previous assumptions about the introduction and spread of *kūmara* around New Zealand and the wider Pacific, allowing a more nuanced evaluation and demonstrating the presence of gaps in our knowledge that need to be addressed by further research.

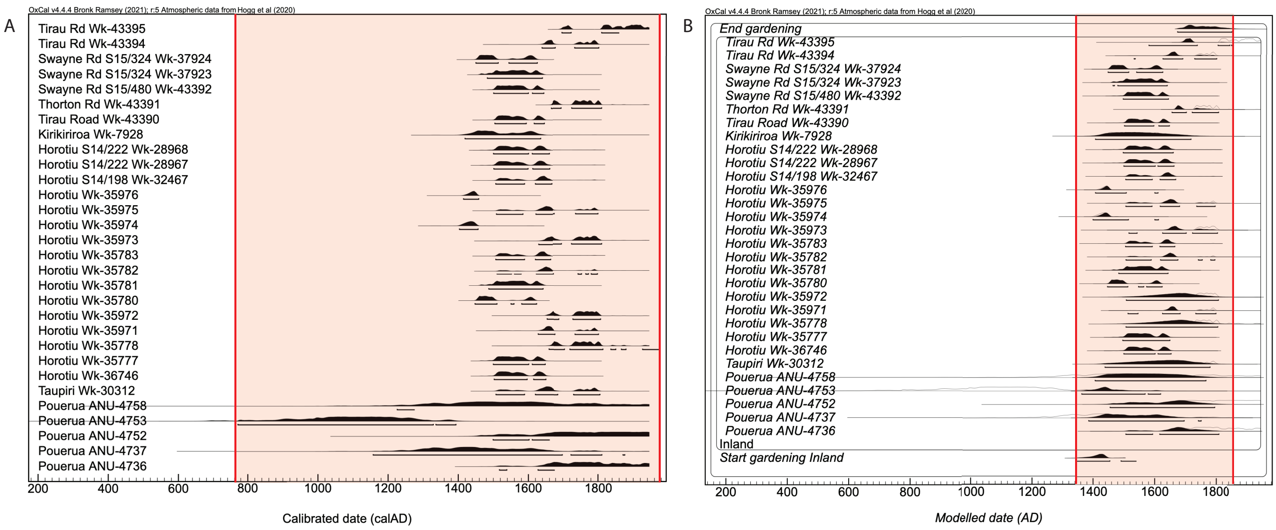


Figure 2. Example of improved date resolution using Bayesian modelling. A: Basic calibrated date series. B: Bayesian modelled single phase model. The lighter shaded probability distributions represent the unmodelled calibrated dates, and the black distributions represent the results after Bayesian modeling.

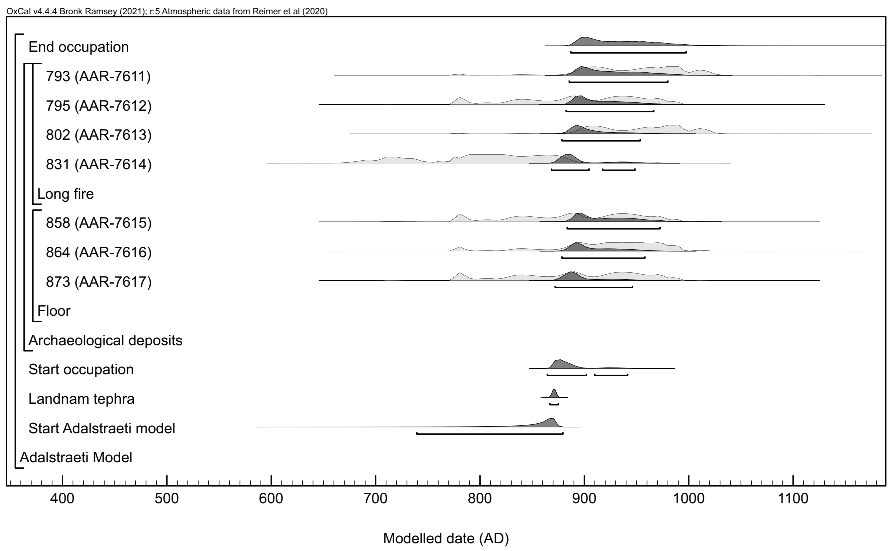
One drawback of this type of model is that the dates are assumed to be a random sample randomly deposited through time (Bronk Ramsey 1998, 2009a). The precision of the posterior probability estimates within these single-phase models depends on the number and the distribution of dates. Therefore, a higher percentage of late or early dates in models results in correspondingly older or younger age ranges that may skew chronologies (Blauuw et al. 2018; Hamilton & Krus 2018).

Informative Models

More structured, informative models place ^{14}C ages within an ordered temporal relationship (a *Sequence*) that mirrors the stratigraphy of an archaeological site. In more complex models, a combination of ordered sequences and unordered phases can be nested together and combined with other forms of dating information (Figure 3).

In this example, Batt et al. (2015) combined stratigraphy, tephrochronology, and multiple ^{14}C ages to assess the chronology of the site of Aðalstræti, a Viking-period hall in Reykjavík, Iceland. The chronological model for Aðalstræti uses seven barley seed ages and a calendar date for the Landnám tephra (871 ± 2 CE), which underlies the hall. Thus the tephra provides a *terminus post quem* for the site and initial occupation of the hall. The boundary command labelled 'Start occupation' occurs between this tephra and the floor and hearth sequences. The barley ^{14}C ages are arranged into two *Sequences* within the same *Phase* (Figure 3). This is because the four stratigraphically related samples within the hearth could not be linked directly to the three stratigraphically ordered samples from the floor layers.

A



B

```
Plot()
{
 Sequence("Adalstraeti Model")
 {
  Boundary("Start Adalstraeti model");
  C_Date("Landnam tephra", 871, 2);
  Boundary("Start occupation");
  Phase("Archaeological deposits")
  {
   Sequence("Floor")
   {
    R_Date("873 (AAR-7617)", 1152, 36);
    R_Date("864 (AAR-7616)", 1129, 35);
    R_Date("858 (AAR-7615)", 1153, 36);
   };
   Sequence("Long fire")
   {
    R_Date("831 (AAR-7614)", 1218, 40);
    R_Date("802 (AAR-7613)", 1087, 35);
    R_Date("795 (AAR-7612)", 1150, 36);
    R_Date("793 (AAR-7611)", 1092, 39);
   };
  };
  Boundary("End occupation");
 };
};
```

Figure 3. A: Bayesian chronological model for Aðalstræti Hall. B: OxCal script for Aðalstræti Hall showing the use of *Boundary*, *Sequence*, and *Phase* commands. Radiocarbon ages are defined in the model code by the term R_Date (*syntax* = R_Date *'name', date, error)*. The calendar (tephra) date is defined by C_date.

This model suggests that the hall was occupied by 890 CE, indicating that construction was within a few years of the deposition of the Landnám tephra, while the constrained dates of the floor and hearth deposits suggest that the hall was in use for less than 150 years. However, the authors recommended caution with this interpretation because the ^{14}C samples came from hearths and floor surfaces that would have been repeatedly swept out, leaving gaps not reflected in the model. Therefore, these samples likely have not captured the earliest and final phases of use.

Dealing with Outliers

Even after defining realistic models and selecting and submitting secure samples for dating, it is likely that some of the dates will not conform to prior expectations. Disassociation between the event and date is the most common form of anomaly between a ^{14}C age and expected age, and is not unexpected given the complexity of achieving such a goal with dynamic taphonomic contexts. Outlier analysis in OxCal (Bronk Ramsey 2009b) is one way to test whether the radiocarbon data agree with the prior information. This methodology provides a probabilistic measure of the degree to which samples appear to be outliers, and then calculates an offset relative to the context within which it is found and downweighs the influence of those ages on the model (Bronk Ramsey et al. 2010). In addition, specific outlier models can be chosen for specific types of outliers (Bronk Ramsey 2009b), enabling the inclusion of lower precision ^{14}C results that are functionally related to the context (e.g. samples composed of old-growth wood) and samples of insecure context.

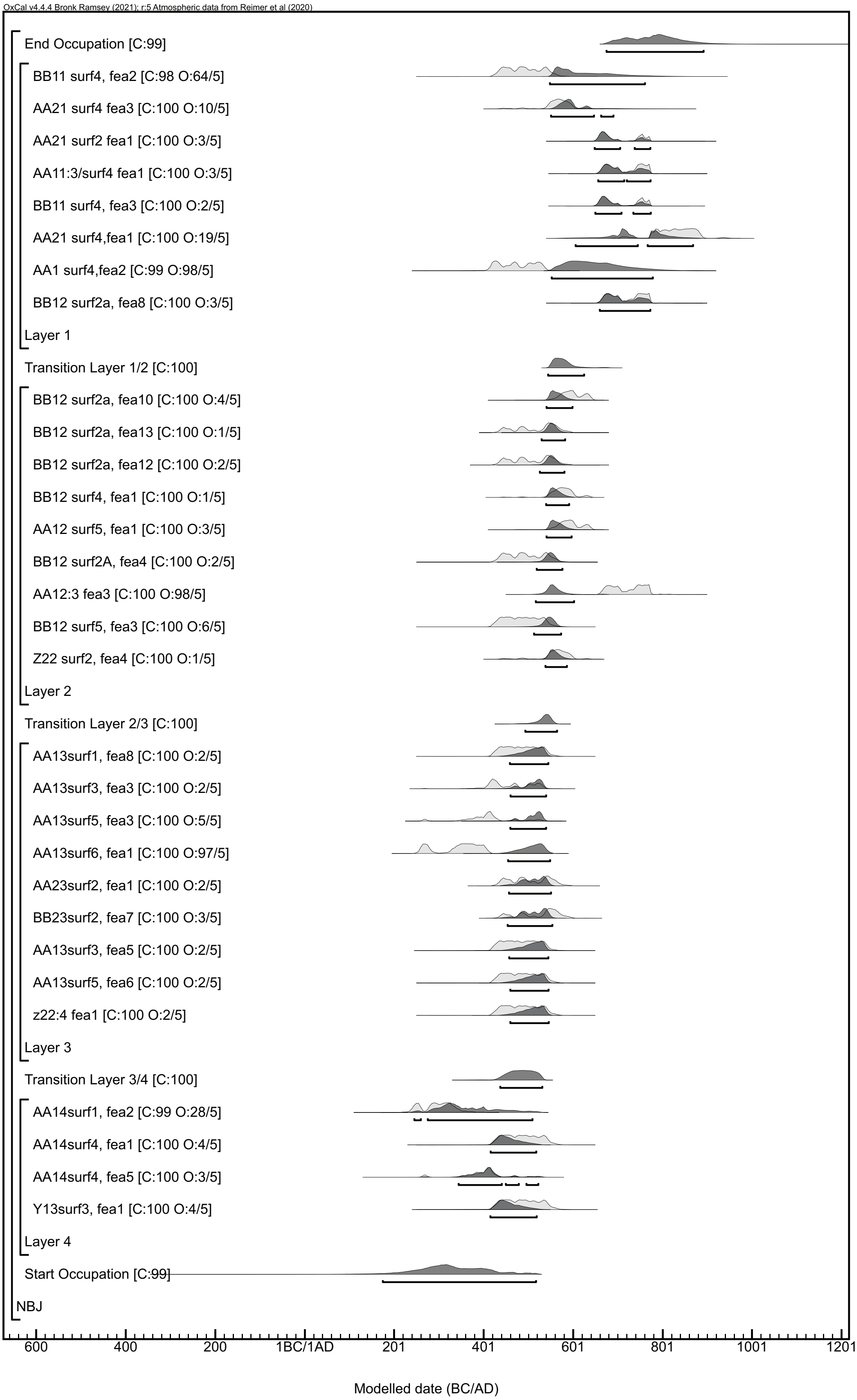


Figure 4. Bayesian age model for the Non Ban Jak sequence. The notation [e.g. O:2/5] indicates a 2% posterior probability of being an outlier in the model. Values >5 are considered outliers. C= Convergence values (see text for discussion) (figure after Higham et al. 2020).

The Bayesian model developed for Non Ban Jak shows the influence of outlier analysis on the modelled chronology and site interpretation (Figure 4). This model consists of four layers (*Phases*) arranged in *Sequence*, each separated by a *Boundary* command. Within this model, the internal consistency of the calibrated dates was tested using a General t-Type model with a prior outlier probability of 5%. This allows outliers to be either too young or too old and is the most common type of outlier model for archaeological applications (Bronk Ramsey 2009b). In this example, the model results indicate the presence of four major outliers (those with posterior outlier probabilities >60%). These outliers include three of >97%, which are not included in the model results 97% of the time, and two with outlier probabilities of between c. 20 and 30%. Three of these outliers come from Layer 1, and two are of similar date to material in Layer 2, suggesting intermixing between the two layers. Despite these outliers, the convergence values generated by OxCal were uniformly high (>98) and therefore indicate that the overall model is robust (low values indicate many different incompatible solutions to the model; Bronk Ramsey 2009a).

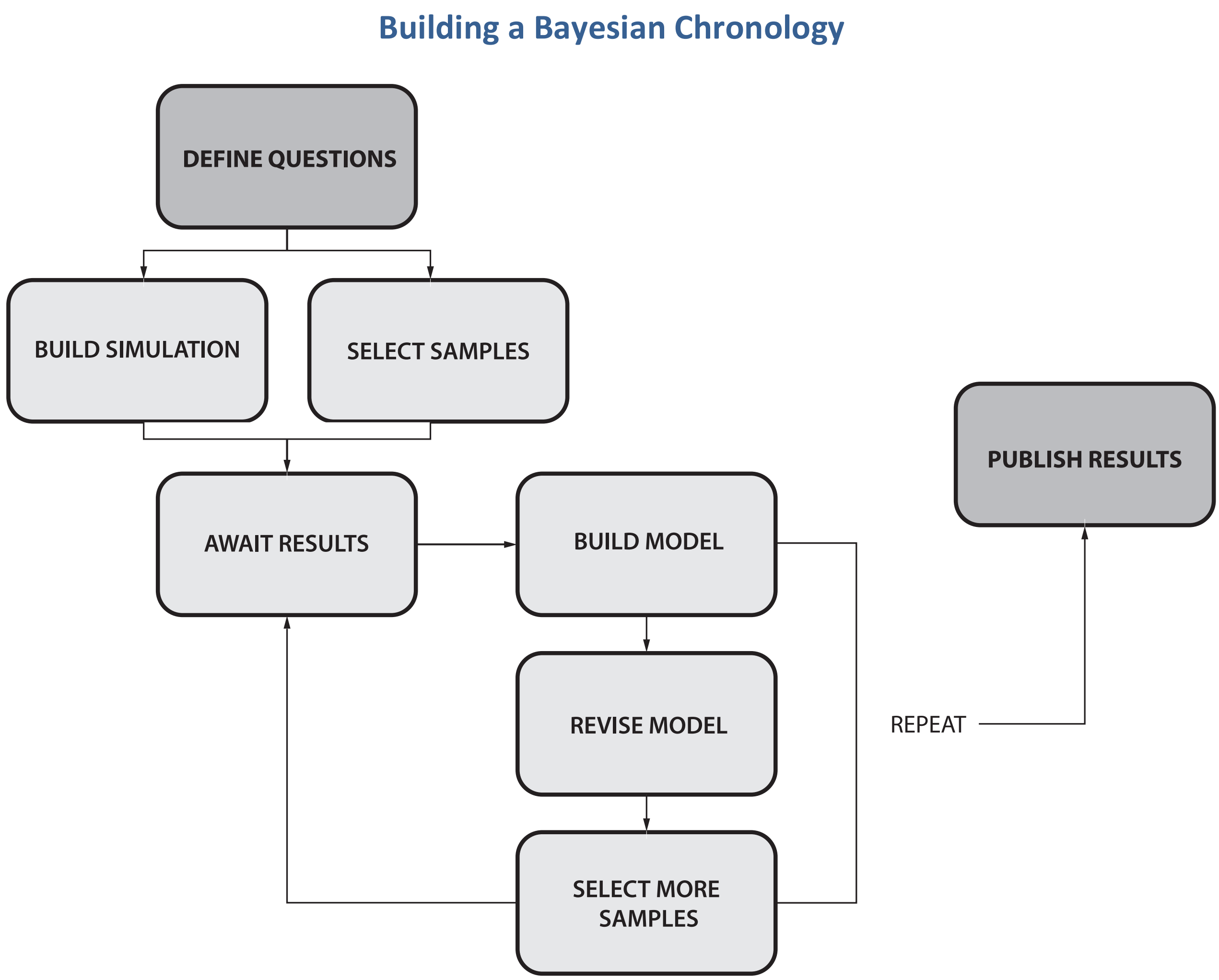


Figure 5. Flow diagram showing the stages of building a Bayesian chronology (after Bayliss 2009).

The examples of Bayesian chronologies presented here all started with careful sample selection and interpretation of how the ^{14}C age relates to the activity under investigation (Figure 5). When developing a model, the specific parameters should be constructed following a progressive dating programme that infills and tests various aspects of the model even before the dating begins (e.g. using R_simulate in OxCal (Bronk Ramsey 2009b)). Bayliss (2015) discusses in detail guidelines for constructing Bayesian chronological models. However, it is important to remember that, ultimately, all models are wrong; the age ranges are a product of a statistical model imposed on the data and an interpretation of the archaeological evidence. Regardless, recognition of this fact, combined with care and attention to detail, will ensure that these models are valuable tools for building robust archaeological chronologies.

References

Anderson, A. & Petchey, F. (2020). The transfer of kūmara (*Ipomoea batatas*) from East to South Polynesia and its dispersal in New Zealand. *Journal of the Polynesian Society* 129(4): 351–382. https://doi.org/10.15286/jps.129.4.351-382

Batt, C. M., Schmid, M. M. E. & Vésteinsson, O. (2015). Constructing chronologies in Viking Age Iceland: Increasing dating resolution using Bayesian approaches. *Journal of Archaeological Science* 62: 161–174. doi: https://doi.org/10.1016/j.jas.2015.07.002

Bayliss, A. (2009). Rolling out revolution: Using radiocarbon dating in archaeology. *Radiocarbon* 51(1): 123–147. https://doi.org/10.1017/S0033822200033750

Bayliss, A. (2015). Quality in Bayesian chronological models in archaeology. *World Archaeology* 47(4): 677–700. https://doi.org/10.1080/00438243.2015.1067640

Bayliss, A., Bronk Ramsey, C., van der Plicht, J. & Whittle, A. (2007). Bradshaw and Bayes: Towards a timetable for the Neolithic. *Cambridge Archaeological Journal* 17(1): 1–28. https://doi.org/10.1017/S0959774307000145

Blaauw, M. & Christen, J. A. (2011). Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis* 6(3): 457–474. https://doi.org/10.1214/11-BA618

Blaauw, M., Christen, J. A., Bennett, K. D. & 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. https://doi.org/10.1016/j.quascirev.2018.03.032

Bronk Ramsey, C. (1995). Radiocarbon calibration and analysis of stratigraphy: The OxCal program. *Radiocarbon* 37(2): 425–430. https://doi.org/10.1017/S0033822200030903

Bronk Ramsey, C. (1998). Probability and dating. *Radiocarbon* 40: 461–474. https://doi.org/10.1017/S0033822200018348

Bronk Ramsey, C. (2001). Development of the radiocarbon program OxCal. *Radiocarbon* 43(2A): 355–363. https://doi.org/10.1017/S0033822200038212

Bronk Ramsey, C. (2009a). Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1): 337–360. https://doi.org/10.1017/S0033822200033865

Bronk Ramsey, C. (2009b). Dealing with outliers and offsets in radiocarbon dating. *Radiocarbon* 51: 1023–1045. https://doi.org/10.1017/S0033822200034093

Bronk Ramsey, C., Dee, M., Lee, S., Nakagawa, T. & Staff, R. (2010). Developments in the calibration and modelling of radiocarbon dates. *Radiocarbon* 52(3): 953–961. https://doi.org/10.1017/S0033822200046063

Buck, C. E., Kenworthy, J. B., Litton, C. D. & Smith, A. F. M. (1991). Combining archaeological and radiocarbon information: A Bayesian approach to calibration. *Antiquity* 65(249): 808–821. https://doi.org/10.1017/S0003598X00080534

Buck, C. E., Litton, C. D. & Scott, E. M. (1994). Making the most of radiocarbon dating: Some statistical considerations. *Antiquity* 68(259): 252–263. https://doi.org/10.1017/S0003598X00046561

Haslett, J. & Parnell, A. (2008). A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society. Series C: Applied Statistics* 57: 399–418. https://doi.org/10.1111/j.1467-9876.2008.00623.x

Hamilton W. D. & Krus A. M. (2018). The myths and realities of Bayesian chronological modeling revealed. *American Antiquity* 83(2): 187–203. https://doi.org/10.1017/aaq.2017.57

Higham, C. F. W., Higham, T. F. G. & Petchey, F. (2020). Introduction and chronology. In Higham, C. F. W. & Kijngam, A. (Eds.), *The Origins of the Civilization of Angkor: Volume VII. The Excavation of Non Ban Jak*, pp. 1–6. Bangkok: Fine Arts Department.

Reimer, P. J., Austin, W. E. N., Bard, E., Bayliss, A. Blackwell, P. G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Hajdas, I., Heaton, T. J., Hogg, A. G., Hughen, K. A., Kromer, B., Manning, S. W., Muscheler, R., Palmer, J. G., Pearson, C., van der Plicht, J., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Turney, C. S. M., Wacker, L., Adolphi, F., Büntgen, U., Capano, M., Fahrni, S. M., Fogtmann-Schulz, A., Friedrich, R., Köhler, P., Kudsk, S., Miyake, F., Olsen, J., Reinig, F., Sakamoto, M. Sookdeo, A. & Talamo, S. (2020). The INTCAL20 northern hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* 62(4): 725–757. https://doi.org/10.1017/RDC.2020.41