



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/174268/>

Version: Submitted Version

Book Section:

Buck, C.E. and Juárez, M.A. (2020) Modelización bayesiana de radiocarbono para principiantes. In: Barceló, J.A. and Morell, B., (eds.) Métodos cronométricos en arqueología, prehistoria y paleontología. Dextra Editorial, Madrid, pp. 297-314. ISBN: 9788417946340.

© 2020 Dextra. This is an author-produced version of a chapter submitted for publication in Métodos Cronométricos En Arqueología, Prehistoria Y Paleontología. Uploaded with the publisher's permission.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Modelización bayesiana de radiocarbono para principiantes

Caitlin E. Buck y Miguel A. Juárez

University of Sheffield

28 de marzo de 2017

1. Antecedentes

Los modelos bayesianos cronológicos son modelos estadísticos que permiten representar, gestionar e interpretar información cronológica, tanto absoluta como relativa, proveniente de uno o más proyectos de investigación arqueológica o paleo-ambiental. Han sido desarrollados durante los últimos treinta años específicamente para las comunidades de investigación arqueología y paleológica, por estadísticos y desarrolladores de software, aprovechando la revolución en nuestra capacidad de implementarlos utilizando métodos basados en simulación (en contraposición al uso de métodos exactos).

Los usuarios de este tipo de software no necesitan conocer los detalles de los modelos estadísticos, ni de las técnicas matemáticas utilizadas, ni de su implementación computacional. Sin embargo, es esencial entender qué es un modelo (estadístico), para poder juzgar las elecciones hechas cuando utilizan un modelo específico para su proyecto y al mismo tiempo comprender suficientemente si aquellas decisiones clave, hechas por los estadísticos y desarrolladores de software, son apropiadas para sus fines.

En este artículo proveeremos al lector con los antecedentes básicos que requieren para llevar a cabo cada una de las tareas anteriores. En la Sección 1.1 presentamos algunos de los conceptos y decisiones clave en el proceso de modelización en general y la Sección 2 se centra en la construcción formal de cronologías, resaltando que tanto los dibujos de sección como las matrices de Harris son ejemplos de modelos cronológicos. Formalizamos la notación matemática en la Sección 3 para definir los modelos estadísticos utilizados regularmente en la datación bayesiana por radiocarbono. Usamos éstos en ejemplos específicos y discutimos el uso de las salidas de software. La Sección 4 se centra en aspectos prácticos del uso de software para calibración bayesiana de radiocarbono y la Sección 5 ofrece una mirada al futuro.

1.1. ¿Qué contiene un buen modelo?

Para efectos de este artículo, definimos un modelo como

una representación de una persona, organismo, estructura o concepto, típicamente más pequeña, más simple y/o más abstracta que la original.

Esta definición resalta la gran variedad de maneras en las que el término modelo es usado en el habla moderna, pero no dice mucho acerca de lo que un buen modelo debe contener para algún propósito específico. Para acercarnos a esto, comenzamos considerando un problema de modelización simple y acaso trivial: ¿qué debe de contener un buen modelo de elefante?



Figura 1. Dos modelos de elefante: ¿cuál es mejor?

Analicemos los modelos de elefante en la Figura 1. Claramente, ninguno de los dos es preciso anatómicamente, pero ¿es uno de los dos modelos mejor que el otro? Nuestra aseveración es que la respuesta depende del objetivo para el cual éste es utilizado. Si el modelo es para entretener a una niña de 3 años alguna tarde lluviosa, tal vez el mejor modelo es el de la izquierda: no tiene partes afiladas o frágiles, así que es seguro para que juegue con él sin supervisión y además es fácil que reconozca al elefante solo por la trompa larga. Si, en cambio, el modelo es para que una niña de 10 años aprenda los rasgos esenciales de un elefante, el modelo de la derecha es ciertamente mejor: tanto las patas como la cabeza son más fidedignas y también tiene cola, orejas y colmillos, a diferencia del otro modelo. Si quisiésemos ir más allá de estos componentes básicos para enseñar a una persona más adulta las diferencias entre, digamos, un elefante africano y otro asiático o acerca de la fisiología de los elefantes en relación con otros grandes mamíferos, ninguno de estos dos modelos sería adecuado y tendríamos que encontrar otro con mayor detalle anatómico.

Este tipo de analogías son útiles hasta cierto punto y corremos el riesgo de llevarla algo más allá de lo necesario pero, antes de dejarla y comenzar a contemplar modelos cronológicos, vale la pena enfatizar un par de similitudes entre éstos y los modelos de elefante: ambos pueden ser modelos ya existentes o hechos a medida y descriptivos o mecanicistas. Los elefantes en la Figura 1 son existentes y descriptivos, pero si quisiéramos encontrar modelos anatómicamente correctos para fines algo más sofisticados, éstos tendrían que ser hechos a medida y mecanicistas.

Análogamente, para una variedad de fines, puede ser que lo necesario para llevar a cabo nuestra inferencia arqueológica o paleológica sean modelos cronológicos descriptivos y existentes. En alguna otra ocasión podremos requerir modelos existentes y mecanicistas que, al menos en parte, capturen los mecanismos que dieron lugar a las observaciones cronológicas que estudiamos. En otras, sin embargo, ninguno de los modelos existentes serán apropiados para nuestras necesidades y tengamos que requerir un modelo adecuado de un estadístico y/o desarrollador de software.

2. Construcción formal de una cronología

Modelizar siempre implica elegir. Normalmente comenzamos por seleccionar un medio, paradigma o marco de referencia dentro del cual construir el modelo y llevar a cabo las elecciones concomitantes. La primera elección en modelización estadística es entre las perspectivas frecuentista y bayesiana. La estadística frecuentista (o clásica) está basada en la interpretación de probabilidad como la frecuencia relativa de sucesos homogéneos en un gran número (teóricamente infinito) de experimentos repetibles. La perspectiva bayesiana —en honor a Thomas Bayes (1701–1761)— afirma que la evidencia acerca del estado (desconocido) del mundo se expresa en términos de “estados de creencia o conocimiento”, representados como afirmaciones de probabilidad subjetiva. Debido a su focalización en estos grados de conocimiento, el paradigma bayesiano es idóneo para representar y gestionar las opiniones de los expertos y su conocimiento previo conjuntamente con datos científicos, por lo que es particularmente atractivo para científicos arqueólogos y paleo-ambientalistas en busca de un método para combinar coherentemente información de fuentes diversas.

Buck et al. (1996) propugnan en detalle por el uso de el paradigma bayesiano en arqueología y no resumiremos sus argumentos aquí. En lugar de ello, evaluaremos las circunstancias en las que nos hallamos al construir modelos cronológicos y propondremos que, excepto para algunos de los problemas más simples, el uso de la estadística bayesiana es enteramente natural.

2.1. Modelos cronológicos adecuados

Así como en el problema de modelizar un elefante, para seleccionar el tipo y escala apropiados para el modelo en nuestro proyecto de investigación paleo o arqueológico es imprescindible saber precisamente qué es lo que estamos modelando y por qué. Si, por ejemplo, estamos tratando de datar un solo evento en el registro arqueológico, como la muerte de un espécimen humano cuyo esqueleto articulado ha sido encontrado en una sepultura bien sellada, un modelo frecuencialista puede ser adecuado. Podríamos, en teoría, datar el mismo evento repetidas veces enviando varias muestras del esqueleto al laboratorio y resumir los resultados usando técnicas frecuencialistas. Si, en cambio, quisiéramos datar una secuencia de eventos (e.g. la estratigrafía de un sitio arqueológico o un núcleo de sedimento), será necesario modelizar el conocimiento de la cronología relativa y muy probablemente la opinión o el conocimiento previo de algún experto, en cuyo caso seguramente tendemos que recurrir al paradigma bayesiano. Si queremos construir una cronología para el sitio entero, tendremos que comparar y combinar secuencias múltiples con información de varios expertos, provenientes de diferentes sitios y trabajando en tiempos distintos. Así que el modelo en cuestión tendría que ser modular y con seguridad incorporar expresiones de probabilidad individual de cada uno de ellos. En estas circunstancias la única opción viable, de la que tenemos conocimiento, capaz de proveer una interpretación simultánea, robusta y coherente de todas estas fuentes de información interrelacionadas, es la estadística bayesiana.

2.2. Modelos cronológicos como pictogramas

La principal motivación de casi todas las cronologías encuadradas en el paradigma bayesiano es, al menos parcialmente, la de combinar dataciones científicas con conocimiento inicial absoluto y/o relativo. La información más común de este tipo, proveniente de excavaciones arqueológicas, es estratigráfica, normalmente almacenada en las memorias de investigación como bosquejos o mapas: **DIBUJOS SECCIONALES**, diagramas de fase, matrices de Harris o bocetos de edad-profundidad. Todos los anteriores son modelos pictóricos de cronologías relativas y un punto clave de partida en la modelización de cronologías. La Figura 2 es un ejemplo de dibujo seccional, de un sitio arqueológico ficticio, con el que ilustramos la manera en que esta clase de información puede ser usada para comenzar la modelización cronológica.

El paso inicial para formar un modelo cronológico es simplificar el bosquejo estratigráfico, centrando la atención en su información temporal. El boceto en la figura 3a muestra esta simplificación, usando la Figura 2. En el panel 3b se muestra una simplificación, tomando en cuenta únicamente aquellos contextos que contienen muestras que pudieran ser enviadas para datación cronométrica.

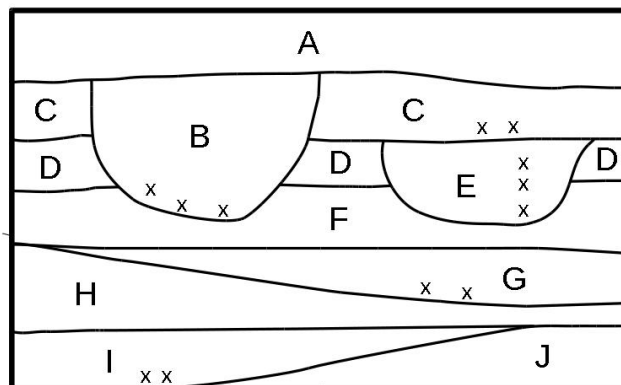


Figura 2. Un ejemplo ilustrativo de un dibujo seccional de parte de un sitio arqueológico ficticio. La localización de las muestras apropiadas para la datación cronométrica se identifican con las cruces.

Presentar los dos bocetos adyacentes en la Figura 3 nos permite resaltar una de las elecciones cruciales que todo creador de modelos cronológicos debe hacer: cuáles contextos y muestras incluir. La mayoría de los modeladores concuerdan en que, permaneciendo el resto constante, debemos de seguir el principio de Occam y mantener el modelo tan sencillo como sea posible. Esta parece una buena política a seguir, aunque las cosas raramente permanecen constantes en la práctica. Lo que queremos decir con esto, es que el excluir o no contextos y muestras en nuestro modelo cronológico y la manera que en representamos aquellas que incluimos, seguramente tendrán un impacto, aunque sea pequeño, en los resultados.

Debido a ésto y a que hay muchas más elecciones que hacer mientras desarrollamos la implementación, en las secciones posteriores del artículo proveemos algunas reglas generales para aquellos que requieran un uso responsable del software de modelización bayesiana de cronologías. Antes de eso es necesario que a) identifiquemos claramente aquellos componentes clave que deseamos interpretar o gestionar y b) pensar y escribir de manera formal, dando pie a la elaboración de modelos estadísticos. Abordamos a) en la siguiente sección y en la sección 3 nos concentramos en b). En el proceso, hacemos uso de una extensa literatura, en particular Blackwell y Buck (2008); Buck et al. (1996); Naylor y Smith (1988) y las referencias ahí citadas.

2.3. Componentes clave de un modelo estadístico de cronologías

Hay, en líneas generales, dos tipos de información a representar en un modelo estadístico para la construcción de cronologías: relativa y absoluta. La información cronológica relativa se refiere normalmente al ordenamiento (a priori) de los eventos, mientras que la información

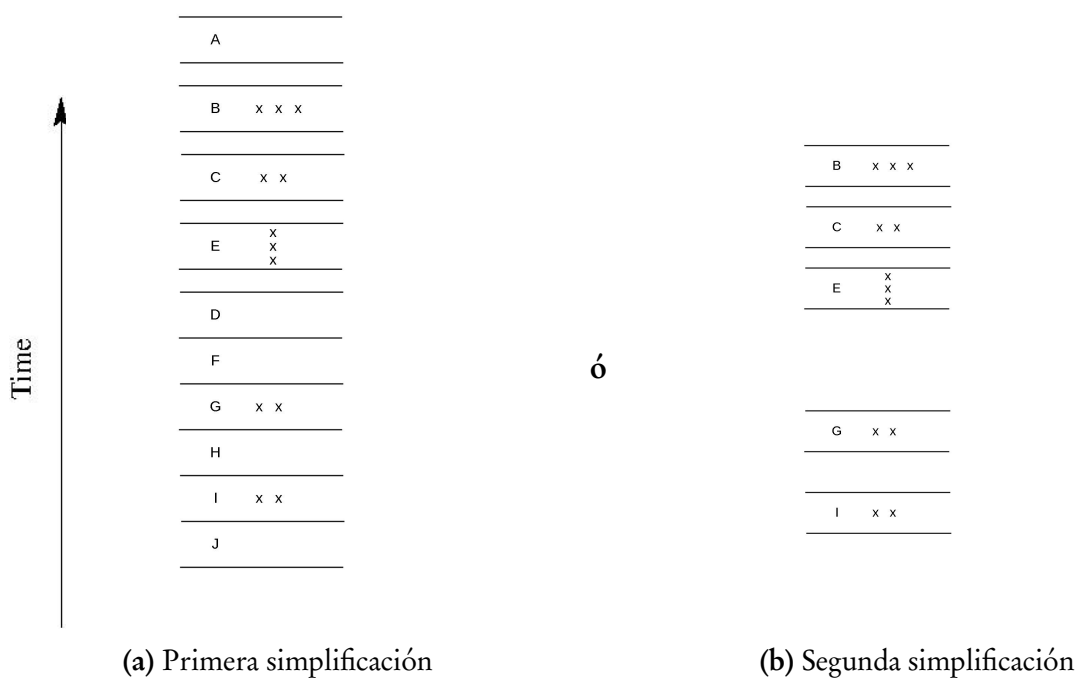


Figura 3. Bocetos de la información cronológica relativa en los perfiles estratigráficos mostrados en la Figura 2. Las líneas horizontales representan las fronteras del contexto arqueológico y las cruces indican la localización dentro de la secuencia estratigráfica de las muestras apropiadas para datación cronométrica. El perfil en (3a) muestra todos los contextos en el bosquejo original; el perfil en (3b) muestra solo los contextos con muestras apropiadas para datación absoluta.

cronológica absoluta normalmente surge de registros históricos o de métodos científicos de datación.

El resto del artículo se centra en las fechas absolutas resultantes de la datación por radiocarbono, pero cabe mencionar que se han desarrollado métodos bayesianos para dendrocronologías (Jones, 2013; Litton y Zainodin, 1991; Millard, 2002), datación por luminescencia (Zink, 2015) y por barrido de resonancia electrónica (ESR, por sus siglas en inglés) (Millard, 2006). También centraremos nuestra atención en interpretar y gestionar la información cronológica representada en las Figuras 2 y 3. Existen métodos bayesianos capaces de representar una gama bastante más amplia de características que las necesarias para nuestro propósito, pero nuestro objetivo presente es proveer una introducción más que una revisión; así que referimos al lector interesado a los trabajos citados.

Podemos encontrar dos tipos de eventos cronológicos clave en las Figuras 2 y 3: aquellos relacionados directamente a objetos que pueden ser datados (como las muestras depositadas, identificadas con cruces) y aquellas que no lo son, como la creación de fronteras contextuales. Únicamente es posible obtener estimaciones de fechas absolutas para estas últimas creando modelos que las relacionen con objetos datables, a través de la secuencia estratigráfica. En

resumen, requerimos componentes en nuestro modelo para representar:

- las fechas subyacentes verdaderas que deseamos aprender, y solo algunas de ellas pueden ser relacionadas de manera directa a objetos datables.
- la relación estratigráfica entre las fechas subyacentes de todos los componentes en el registro estratigráfico,
- la relación entre las fechas subyacentes verdaderas y su determinación por radiocarbono, incluyendo las incertidumbres provenientes del laboratorio y la calibración requerida.

Consideramos estos tres puntos en la siguiente sección, donde comenzamos con un modelo simple que incluye únicamente aquellos objetos datables directamente y sus correspondientes relaciones estratigráficas, para después incluir las fronteras contextuales y sus relaciones estratigráficas con los objetos datables. En el proceso, suponemos que las muestras datables en las figuras 2 y 3 (identificadas con las cruces) producen las determinaciones por radiocarbono en el cuadro 1.

Cuadro 1. Determinación por radiocarbono correspondientes a la secuencia estratigráfica en la figura 2, junto con una indicación de si estamos suponiendo que la muestra está estratificada bajo el contexto adecuado.

Contexto	Etiqueta de muestra	Estratificación bajo contexto	Edad media ^{14}C	Error de laboratorio
B	θ_1	No	5700	30
	θ_2	No	5670	30
	θ_3	No	5650	30
C	θ_4	No	5720	30
	θ_5	No	5780	30
E	θ_6	Sí	5900	50
	θ_7	Sí	5870	50
	θ_8	Sí	5850	50
G	θ_9	No	6000	30
	θ_{10}	No	6130	30
I	θ_{11}	No	6200	50
	θ_{12}	No	6250	50

3. Modelos bayesianos para datación por radiocarbono

El tono formal de esta sección puede resultar intimidatorio para algún lector. Si este es el caso, les sugerimos que se concentren en comprender la notación que utilizamos y la forma

general de las ecuaciones. No es esencial entender en detalle las ecuaciones para adquirir una intuición de la estructura de los modelos y el modo en que son construidos; siendo esto lo más importante. En vista de esto, presentamos la figura 4, en espera de que sirva de guía al lector para entender la estructura del modelo conforme es descrita. La parte superior de la figura corresponde a las ideas y conceptos en la sección 3.1 y la inferior a aquellas en la sección 3.2.

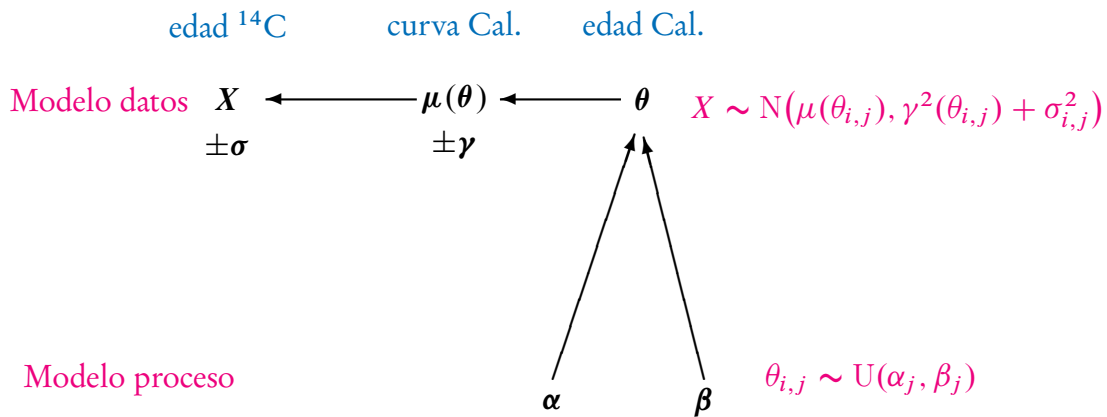


Figura 4. Representación esquemática del modelo estadístico (jerárquico) desarrollado por Naylor y Smith (1988). En él representamos las relaciones entre los parámetros cronológicos (θ y σ) que están relacionados directamente con los datos, x y los utilizados para representar los procesos arqueológicos subyacentes (α y β). Para enfatizar estas relaciones, las flechas representan la dirección causal —que es la opuesta en la que hacemos inferencia.

Ates de entrar en materia, conviene resaltar algunas convenciones en la notación. Utilizamos letras latinas y griegas para denotar observaciones y parámetros, correspondientemente. Fuente normal para escalares (parámetros de dimensión uno) y negritas para vectores (aquellos de dimensión mayor que uno). Dado que los vectores contienen una colección ordenada de valores, normalmente llevaremos a cabo cálculos que los involucren una entrada a la vez. Por ejemplo, si \mathbf{z} es un vector de longitud k (i.e. contiene una secuencia de k escalares), denotamos el producto de todos ellos como $\prod_{m=1}^k z_m$, con \prod representando la operación producto y z_m el m -ésimo elemento de \mathbf{z} .

Usaremos el término distribución de probabilidad (o densidad de probabilidad) para referirnos a aquellas funciones cuyos valores en cualquier punto dado es una declaración probabilística acerca del argumento de esta función. La bien conocida curva en forma de campana de la distribución Normal probablemente viene a la mente. Estas distribuciones pueden ser utilizadas una a la vez para representar parámetros individuales e independientes estocásticamente. Sin embargo, es más común en la práctica trabajar con conjuntos interrelacionados o covariantes de estas funciones. Al hacer esto, tomamos en cuenta varios parámetros de manera simultánea (o conjuntamente), de tal manera que podemos investigar las relaciones

entre ellos. En seguida, cuando creamos gráficos de los resultados, generalmente los hacemos un parámetro a la vez, dado que visualizar distribuciones de probabilidad en dimensiones mayores a dos es notoriamente difícil. A estas representaciones individuales las llamamos distribuciones marginales (contrapuesto a conjunta).

3.1. El modelo bayesiano básico de radiocarbono

Una determinación por radiocarbono consiste de dos partes: una edad estimada por radiocarbono, x antes del presente (BP, por sus siglas en inglés) y un error asociado de laboratorio, σ . Deseamos utilizar esta determinación para aprender acerca de la (verdadera, subyacente) fecha calendárica (¿?¿?), θ (Cal BP), en la cual la muestra datada dejó de metabolizar. Frecuentemente contamos con información previa (histórica, estratigráfica, etc.) acerca de la fecha calendárica, la que representaremos a través de la distribución previa, $p(\theta)$.

Si queremos utilizar tanto los datos como la información previa para aprender acerca de la fecha calendárica, es necesario formalizar el vínculo entre la información disponible, i.e. $x \pm \sigma$ y $p(\theta)$, y la cantidad que nos interesa, i.e. $p(\theta | x, \sigma)$ (la barra vertical en esta notación se lee como “dado”, en este caso, la distribución de probabilidad de θ —la cantidad de desconocida— dado que conocemos x y σ). Debido a que la proporción de átomos de carbono radioactivo en la atmósfera terrestre no ha permanecido constante a través del tiempo, tendremos que usar una curva de calibración para asociar fechas de radiocarbono con calendáricas. Usamos $\mu(\theta)$ para denotar esta curva de calibración, así que, rigurosamente, deseamos aprender acerca de $p(\theta | x, \sigma, \mu(\theta))$.

Es un hecho que no obtendremos exactamente la misma determinación por radiocarbono si enviamos varias partes de la misma muestra orgánica al mismo laboratorio; es por eso que tratamos esta determinación, x BP, como una instancia de la cantidad aleatoria X BP, asociada a la fecha calendárica θ cal BP. Podemos interpretar X como el valor verdadero de la curva de calibración, θ , con cierta incertidumbre añadida, ε . Si denotamos el valor verdadero de la curva de calibración por $\mu(\theta)$ BP, tenemos $X = \mu(\theta) + \varepsilon$, con el supuesto convencional de que la distribución del error es Normal, $\varepsilon \sim N(0, \sigma^2)$.

Fijando la atención en el contexto E de la figura 2, encontramos tres determinaciones de una secuencia estratigráfica (ver cuadro 1), por lo que $\mathbf{x} = (5900, 5870, 5850)$ y $\boldsymbol{\sigma} = (50, 50, 50)$, con las fechas subyacentes verdaderas $\boldsymbol{\theta} = (\theta_6, \theta_7, \theta_8)$. La información estratigráfica nos permite estar seguros (a priori) de que las tres fechas calendáricas están estrictamente ordenadas, $\theta_6 > \theta_7 > \theta_8$. Si usamos, entonces, uno de los modelos estándar para radiocarbono

(aunque no lo introducimos aquí, se discute en detalle en Buck et al., 1996, Cap. 9),

$$p(\mathbf{x} | \boldsymbol{\theta}) \propto \prod_{i=m}^n \exp \left[-\frac{(x_i - \mu(\theta_i))^2}{2(\sigma_i^2 + \gamma^2(\theta_i))} \right],$$

con $\gamma(\theta_i)$ la desviación estándar de la curva de calibración (acordada internacionalmente) en θ_i y, en este ejemplo, $m = 6$ y $n = 8$. Hacemos notar que en la formulación de Naylor y Smith (1988) y Buck et al. (1996) no se incluye el parámetro $\gamma(\theta_i)$, debido a que los autores lo suponen muy pequeño en comparación a σ_i . Un supuesto razonable en ese momento, pero no más, debido a que la incertidumbre de laboratorio se ha reducido considerablemente desde entonces.

La información estratigráfica previa puede ser formalizada de la siguiente manera,

$$p(\boldsymbol{\theta}) \propto \prod_{i=m}^n I(\theta_i)$$

donde

$$I(\theta_i) = \begin{cases} 1 & \theta_{i-1} > \theta_i > \theta_{i+1} \\ 0 & \text{de otro modo} \end{cases},$$

por lo que la actualización bayesiana (i.e. la distribución posterior) es

$$p(\boldsymbol{\theta} | \mathbf{x}) \propto p(\mathbf{x} | \boldsymbol{\theta}) \times p(\boldsymbol{\theta}) \propto \left[\prod_{i=m}^n \exp \left[-\frac{(x_i - \mu(\theta_i))^2}{2(\sigma_i^2 + \gamma^2(\theta_i))} \right] I(\theta_i) \right]. \quad (1)$$

Con los recursos computacionales actuales y una implementación basada en métodos por simulaciones (descritos en Buck et al., 1996, y discutidos más abajo), es inmediato aplicar esta ecuación incluso a secuencias estratigráficas muy largas, gracias a paquetes informáticos existentes para la calibración por radiocarbono como OxCal (Bronk Ramsey, 2009) y BCal (Buck et al., 1999). Como ilustración, usamos Bcal con la estimación de la curva de calibración de radiocarbono IntCal13, acordada internacionalmente (Reimer et al., 2013), para calibrar las tres determinaciones estratificadas del contexto E en la figura 2. Presentamos los resultados en la figura 5, con las distribuciones de probabilidad obtenidas al calibrar individualmente identificadas en azul y en rojo las obtenidas usando la aproximación bayesiana en (1).

Este ejemplo de modelo bayesiano simple es bastante ilustrativo, formalizando la inclusión de la información estratigráfica, que hubiese sido ignorada o incluida de manera ad hoc. Al hacerlo formalmente, las fechas estimadas son típicamente más precisas que las obtenidas al calibrar individualmente. Para el contexto E, el incremento en precisión es bastante modesto,

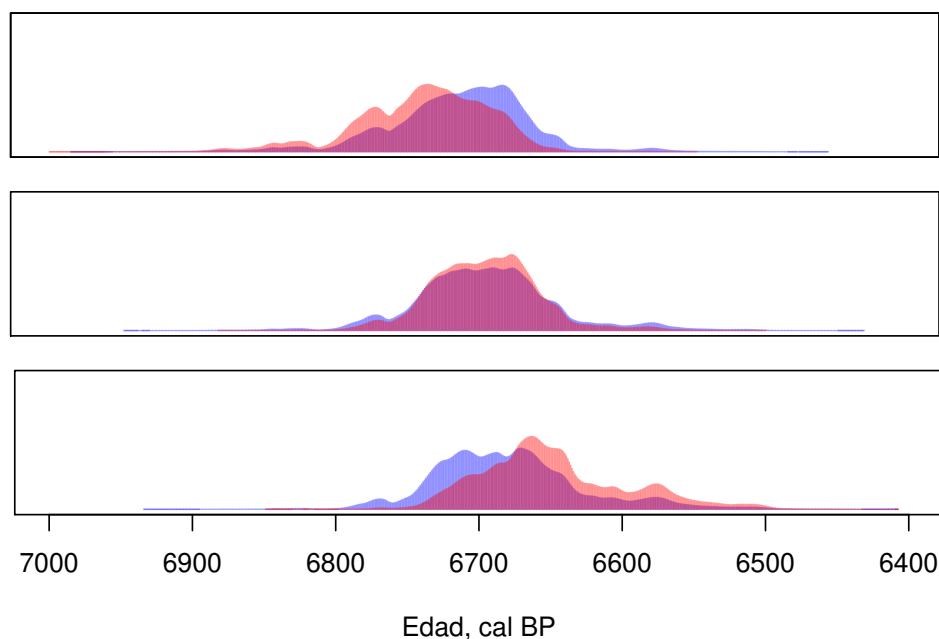


Figura 5. El resultado de calibrar las tres determinaciones por radiocarbono, 5900 ± 50 BP, 5870 ± 50 BP and 5850 ± 50 BP. Azul si la información estratigráfica no es incluida. Rojo cuando sí lo es usando la ecuación (1).

debido principalmente a que contamos con solo tres determinaciones bastante alejadas temporalmente; con una cantidad mayor de datos en intervalos similares el incremento puede ser sustancial. Sin embargo, el modelo (1) no es lo suficientemente sofisticado como para incluir todas las características de la estratigrafía en las figuras 2 o 3, debido a que no ofrece manera de representar eventos para los que no se cuenta con evidencia datable directa. Particularmente, no tenemos forma de representar el historial de deposiciones, registrado a través de las relaciones entre los contextos arqueológicos.

3.2. Modelizando deposiciones

En Naylor y Smith (1988) se introducen en el modelo dos parámetros, α y β (ambos cal BP), para representar los límites entre contextos o fases deposicionales, con $\alpha_j > \beta_j$ (i.e. α_j y β_j son las fechas límite del el j -ésimo contexto). Normalmente contamos con alguna información cronológica relativa que vincula a estos parámetros con muestras datables, pero no con evidencia científica directa acerca de su datación. Esta situación es común, ya que es poco factible encontrar, en una investigación arqueológica o paleoambiental, material datable directamente asociado con todas las localizaciones en la secuencia estratigráfica.

El modelo (1) corregido toma en cuenta aquellos contextos para lo que no se cuenta con

una estratigrafía interna y por ende no hay un orden previo en las fechas de las muestras,

$$p(\mathbf{x} | \boldsymbol{\theta}) \propto \prod_{j=1}^J \left[(\alpha_j - \beta_j)^{-n_j} \prod_{i=1}^{n_j} z_{i,j} \mathbb{I}_B(\theta_{i,j}) \right],$$

con J , el número de contextos o fases en el modelo, n_j el número de muestras datables en el j -ésimo contexto o fase,

$$z_{i,j} = \exp \left[-\frac{(x_{i,j} - \mu(\theta_{i,j}))^2}{2(\sigma_{i,j}^2 + \gamma(\theta_{i,j})^2)} \right]$$

e

$$\mathbb{I}_B(\theta_{i,j}) = \begin{cases} 1 & \beta_j \geq \theta_{i,j} \geq \alpha_j \\ 0 & \text{en otro caso} \end{cases}.$$

Suponiendo una tasa de deposición constante para el material dentro de cada contexto, pero permitiendo que éstas varíen entre contextos, Naylor y Smith (1988) proponen

$$p(\boldsymbol{\theta}) \propto \mathbb{I}_A(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \begin{cases} 1 & \boldsymbol{\alpha}, \boldsymbol{\beta} \in A \\ 0 & \text{de otro modo} \end{cases}$$

como modelo para el conocimiento previo, donde A es el conjunto de valores para los que $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ consistentes con la información previa acerca de la cronología (e.g. estratigrafía). Consecuentemente, las fechas calibradas se obtienen de

$$p(\boldsymbol{\theta} | \mathbf{x}) \propto \mathbb{I}_A(\boldsymbol{\alpha}, \boldsymbol{\beta}) \times \prod_{j=1}^J \left[(\alpha_j - \beta_j)^{-n_j} \prod_{i=1}^{n_j} z_{i,j} \mathbb{I}_B(\theta_{i,j}) \right]. \quad (2)$$

En los casos donde se tiene un orden directo de las muestras dentro del contexto (como en el contexto E), la inclusión de los componentes pertinentes de (1) es inmediato al implementar (2) en un software. Es igualmente trivial incluir aquellos contextos o fases para los que se cuenta con información estratigráfica sin ninguna evidencia directa por datación (i.e. aquellos binomios α_j, β_j relacionados únicamente con los parámetros de otros límites, pero ninguna con las muestras susceptibles de datación directa).

OxCal (Bronk Ramsey, 2009) y BCal (Buck et al., 1999) proveen herramientas para implementar este tipo de modelos; aquí empleamos el segundo para formalizar los modelos descritos en la figura 3. En el proceso, encajamos el contexto E dentro de un modelo que abarca la secuencia estratigráfica entera mostrada en la figura 2, aunque, con fines ilustrativos,

nos centraremos en los resultados marginales de las fechas de las muestras dentro de este contexto. El resumen gráfico en la figura 6 sugiere que el uso de un modelo con mayor detalle estratigráfico (y por tanto, más parámetros), permite (aparentemente) aprender las fechas de las muestras en el contexto E con una mayor precisión.

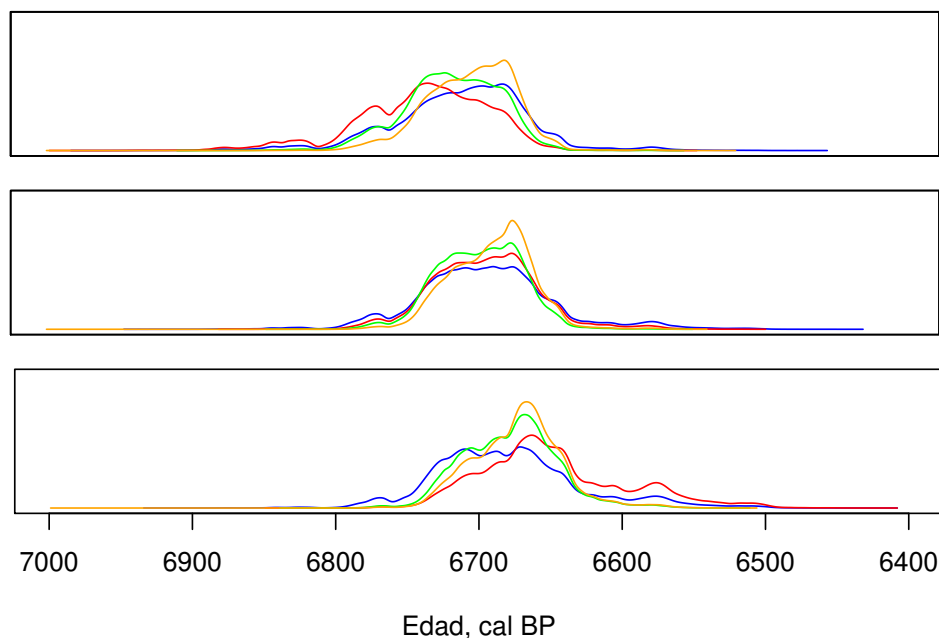


Figura 6. Calibración resultante de las tres determinaciones por radiocarbono en en contexto E. En **azul** cuando toda la información estratigráfica es ignorada; **rojo** cuando se incluye solo el orden dentro del contexto E (usando la ecuación (1)); **verde** cuando se incluye el orden y usamos un modelo para las relación contextual ilustrada en la figura 3b (usando la ecuación (2)); **naranja** cuando el orden dentro del contexto se incluye y la relación contextual en la figura 3a es incluida (usando la ecuación (2)). Téngase en cuenta que, a) por claridad, mostramos solo las distribuciones posteriores marginales para cada caso y b) estas distribuciones son cada vez más apuntadas (i.e. la estimación es más precisa), conforme mayor detalle estratigráfico se incluye en el modelo.

En primera instancia, dado que incrementar la precisión es uno de los objetivos al estimar una cronología, eso parece atractivo; debemos, sin embargo, ser cautos. Seguramente hay un límite después del cual añadir contextos al modelo, sin incluir ninguna evidencia de datación adicional, conlleva una percepción falsa de información extra. ¿Provee el saber que hay otros dos contextos entre el G y el E alguna información cronológica? (ninguna de estas tienen evidencia absoluta de datación asociada) y si este fuese el caso, ¿nuestro modelo de la información en la figura 4 la captura apropiadamente? Discutimos ésta y otras cuestiones relacionadas a la implementación en la siguiente sección pero, antes de ir a ello, revisemos algunas de las estimaciones de las fechas de los límites entre contextos provenientes de nuestro modelo.

Seguimos las sugerencias convencionales acerca de selección de modelos (que menciona-

mos antes y discutimos más adelante) y preferimos modelos sencillos (menos parámetros) a complejos (más parámetros); por ende en la figura 7 mostramos las estimaciones de las fechas entre límites de contextos de la figura 3b. Es posible mostrar resultados similares usando la información en la figura 3a, en caso de que hubiese una razón arqueológica para incluirla. Lo que podemos ver claramente en la figura 7 es el efecto del orden cronológico en las fechas límites, impuesto por la distribución previa, que a su vez ha sido determinado por las relaciones contextuales en la figura 3. También observamos que es posible estimar estas fechas con un margen de 500 años e incluso de 100, en el caso del contexto B, aun a pesar de no contar con evidencia de datación directa para ninguno de estos parámetros; claramente una herramienta de estimación arqueológica poderosa, incluso para esta simple secuencia estratigráfica. Antes de usar estos modelos para asistir en la interpretación en proyectos arqueológicos reales, discutimos algunas de las decisiones tomadas al modelizar o hacer inferencia, en particular a las preguntas que nos planteamos en el párrafo anterior.

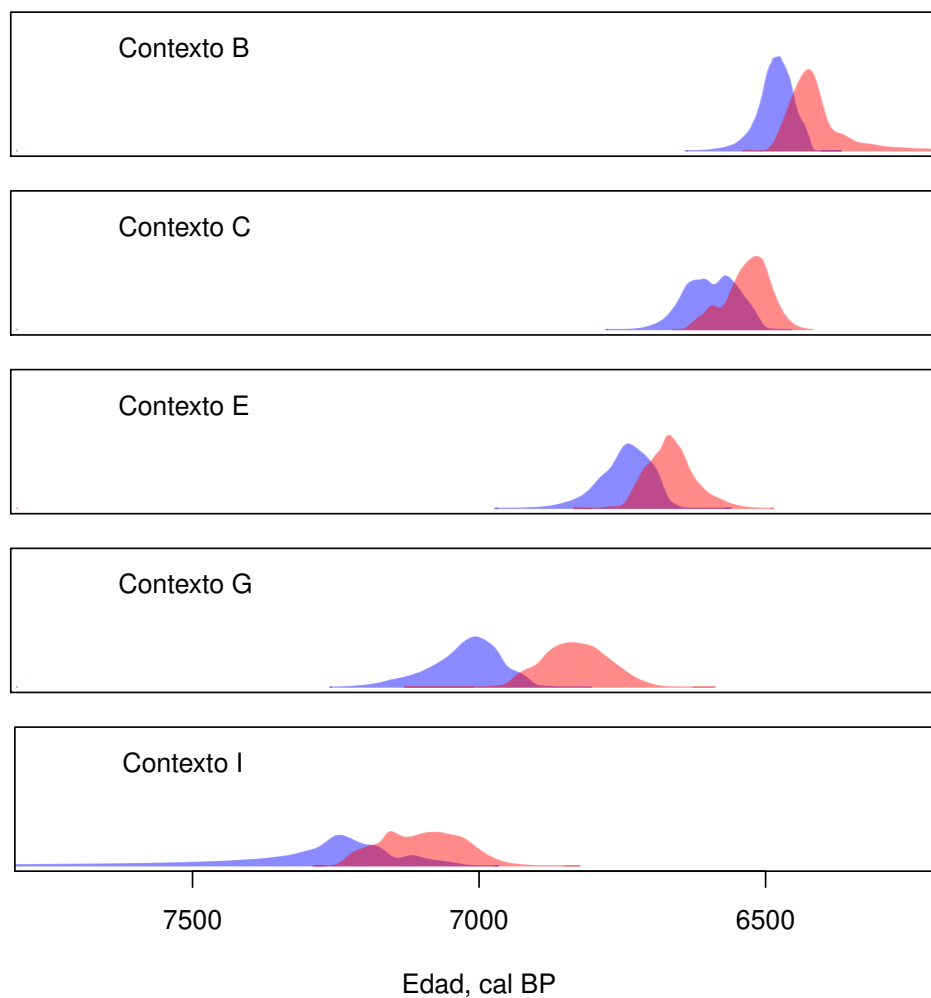


Figura 7. Estimación de fechas calibradas de los límites entre contextos provenientes de la figura 3b. Fechas iniciales en rojo y finales en azul.

4. Decisiones en la modelización e implementación

4.1. Selección de modelos

La implementación de los modelos (1) y (2) que originalmente usaron Naylor y Smith (1988) está basada en integración numérica, una técnica computacionalmente ineficiente y que no escala apropiadamente con el número de parámetros. Si se considera, además, que la potencia de las computadoras en ese entonces era muy limitada, resulta que solo algunos modelos arqueológicos realmente simples podían ser implementados. Como consecuencia natural, la adopción de los modelos bayesianos de cronologías se materializó hasta varios años después, cuando los programadores comenzaron a desarrollar implementaciones basadas en métodos de simulación, conocidos como cadenas de Markov Monte Carlo (MCMC, por sus siglas en inglés). Estos algoritmos son considerablemente más sencillos de programar y además pueden ser modificados fácilmente en caso de requerir componentes adicionales. Estas características los vuelven atractivos por la modularidad de la clase de modelos que consideramos y propicios para su implementación en algoritmos generales como los de OxCal y BCal.

A consecuencia del incremento en el poder de computación y de estas implementaciones, mucho más flexibles, hay disponible una gama de alternativas de software para la modelización de cronologías. A su vez esto ha devenido en un rango amplio de modelos, haciendo la selección de uno de ellos un problema rutinario. Actualmente OxCal (Bronk Ramsey, 2009) es la única implementación que ofrece herramientas formales para la selección de modelos; en breve, la selección se basa en comparar las distribuciones previa y posterior y seleccionar aquel modelo que produzca posteriores más cercanas a las previas.

Esta aproximación es arbitraria y en realidad solo válida para comparar modelos de estructura y tamaño similares y, por tanto, no aptos para enfrentar el problema expuesto al final de la sección 3.2. Para ello requerimos de alguna técnica apropiada para selección de modelos, siendo estas notoriamente difíciles de construir, particularmente cuando el número de parámetros en las diferentes representaciones tiene una variación sustancial, algo común cuando se trabaja con largas secuencias arqueológicas. Entre otras, es por estas razones que la selección formal de modelos es un arte necesario para modelizador bayesiano y, como todo arte, realmente se adquiere a través de la experiencia.

Los autores tenemos algo de experiencia con este tipo de modelización y, en general y con el principio de Occam en mente, preferimos aquellos modelos con la menor cantidad de parámetros. Así que, en el caso del problema de la sección anterior, preferimos el modelo representado por el panel derecho en la figura 4. Esto es porque el modelo más sencillo permite la representación del historial clave de las deposiciones para las muestras datables, pero no incluye explícitamente aquellos contextos sin evidencia datable directa que, por tanto,

contribuyen relativamente poca información a la inferencia.

Si nuestro interés fuese datar un evento arqueológico para el cual no hubiese un contexto con muestras datables, la decisión sería relativamente sencilla. Supongamos, por ejemplo, que el contexto F es una capa de destrucción entre dos periodos de actividad humana; podríamos entonces incluirlo en el modelo, aceptando que la cantidad total de información cronológica disponible no se ha incrementado apenas, pero posibilitando el intentar datar el contexto de interés.

4.2. Uso responsable del software

Nos parece claro, a partir de nuestra discusión acerca de la implementación de modelos de cronologías, que ésta no es una tarea para novicios y por tanto suponemos que el lector hará uso de algún software existente, como OxCal y Bcal, que son herramientas muy potentes, cuando son usadas cuidadosa y responsablemente. Con esto nos referimos a la necesidad de entender, al menos de manera intuitiva, los modelos usados y cómo son implementados en el software. Aquí presentamos una introducción a algunos de estos aspectos, pero es muy probable que el usuario necesite familiarizarse con la literatura provista por el desarrollador y con los manuales de usuario.

Estas responsabilidades se extienden al momento de reportar nuestra modelización. Por ejemplo, es fundamental informar al lector no solo acerca del modelo (o software) que se ha utilizado, sino también del por qué lo hemos seleccionado para esta aplicación en particular. Las razones pueden ser, como en nuestros ejemplos anteriores, relacionadas a la interpretación que el autor hace de la arqueología, o a algunas restricciones prácticas o teóricas. La importancia de resaltar estos detalles estriba en la naturaleza de la obtención de muestras arqueológicas, normalmente anterior a la especificación del modelo. Por tanto, la única manera que el lector tiene de saber que el autor no ha utilizado la misma información más de una vez, es que detalle su razonamiento al seleccionar modelo y previas y, además, que ninguno de estos dos esté basado en los datos científicos que se usarán para la estimación.

El usuario debe también de conocer el software de tal manera que pueda tomar decisiones basadas en principios y reportarlas claramente. Esto es particularmente acuciante dado que el software basado en MCMC usa simulación y, por tanto, cada vez que el modelo es ajustado, obtendremos resultados (ligeramente) distintos. Algunos conceptos claves incluyen:

- Descartado (burn-in, en inglés). En general, las primeras muestras producidas por un algoritmo MCMC no son representativas de la distribución posterior, generalmente porque el punto inicial en el algoritmo es arbitrario y potencialmente alejado de la región

posterior de interés. Por esta razón es necesario descartar un número (indeterminado) de realizaciones iniciales de cada una de las cadenas.

- **Convergencia.** En teoría, las realizaciones obtenidas después del descartado son una muestra (no independiente) de la distribución posterior y pueden ser usadas para inferencia.
- **Adelgazamiento y tamaño efectivo de muestra.** Las muestras MCMC no son independientes e incluso pueden ser fuertemente (auto)correlacionadas en modelos complejos (como los de secuencias estratigráficas). En estos casos, el tomar una submuestra (típicamente de manera sistemática) de la cadena MCMC permite aliviar el efecto esta correlación y por tanto tener una cantidad de información similar en esta muestra de menor tamaño que en la cadena completa.

La mayoría de los proveedores de software incluyen herramientas de determinación automática de convergencia, así como los grados de descarte y adelgazamiento. El usuario debe, no obstante, entender, al menos intuitivamente, las pruebas de convergencia que se han llevado a cabo y registrar las decisiones asociadas, de manera que los lectores interesados sean capaces de replicar sus resultados y en dado caso extenderlos. Los manuales de OxCal y BCal contienen recomendaciones al respecto.

Debido a la naturaleza de los métodos Monte Carlo, es necesario constatar la reproducibilidad de los resultados MCMC repitiendo el análisis con varias cadenas, cada una con diferentes estados iniciales. Lo que queremos verificar, al menos para aquellos parámetros clave, es que las distribuciones posteriores de cada una de las cadenas son la misma, con una precisión *adecuada*. Qué tan adecuada, es relativo; por ejemplo, necesitaremos distintos grados de precisión para distintos periodos del paleolítico.

El uso de la estadística bayesiana requiere un buen número de decisiones subjetivas y por tanto es aconsejable que el usuario lleve a cabo un análisis de sensibilidad acerca de éstas. Los usuarios deben, en cada aplicación, explorar y reportar el efecto que cambios, incluso sutiles, en la forma y los parámetros de las distribuciones previas tienen en la distribución posterior, tal y como lo hicimos al considerar las dos formulaciones de la información estratigráfica de la figura 2. Un análisis de sensibilidad es parte integral de toda implementación bayesiana, ya que sin él no tenemos idea de que tan robustos son los resultados obtenidos. De hecho, en caso de que éstos en realidad fueran poco robustos, como es no poco común para modelos complejos, es muy probable que tengamos que llevar a cabo una exploración más extensa para justificar nuestra selección de previas.

5. Algunas opciones presentes y futuras

Hay disponibles una multitud de software existente para los usuarios de modelos bayesianos de cronologías que ofrecen diferentes opciones de modelización. Aquí solo resaltaremos unas cuantas de estas alternativas que creemos notables:

- Una gama de modelos de deposiciones para e.g. la colonización gradual de un sitio (Jones, 2013; Lee y Bronk Ramsey, 2012) y la acumulación de sedimentos o turba (¿¿hulla??) (Blaauw y Christen, 2011, 2005; Bronk Ramsey, 2008; Christen et al., 1995; Haslett y Parnell, 2008).
- La detección de observaciones sobresalientes en la datación por radiocarbono (Bronk Ramsey, 2008; Christen, 1994).
- Maneras de extender los modelos que permitan la inclusión de una gran variedad de información previa absoluta. Por ejemplo, aquella obtenida a través de documentos históricos, textos antiguos o alguna otra fuente científica de datación, diferente de por radiocarbono.
- La selección automática de muestras durante (¿??un programa progresivo) de datación por radiocarbono, para utilizar el presupuesto de una manera efectiva (Buck y Christen, 1998; Christen y Buck, 1998).

Los primeros tres puntos son de uso rutinario y han sido implementados en software gratuito, en particular OxCal (Bronk Ramsey, 2009). Este no es el caso para el cuarto punto, lo que parece un tanto extraño, dado el gran número de proyectos de datación que se llevan a cabo actualmente desde la perspectiva bayesiana. Su adopción, por supuesto, será limitada hasta que existan implementaciones en software gratuito, así que la aparición de estos últimos será muy bienvenida. OxCal incluye una función, `R_simulate`, que facilita al usuario la generación de dataciones por radiocarbono adicionales para muestras potenciales dentro del modelo correspondiente, y que puede ser considerada como un paso en esta dirección; sin embargo, una selección automática de muestras que conduzca al programa de datación más rentable aún está por venir.

Otras características deseables que no existen aún en software disponibles.

- Extensiones de los modelos bayesianos actuales que permitan la inclusión de componentes espaciales, que sirvan como andamiaje para la modelización espacio-temporal. Este tipo de modelos son atractivos, por ejemplo, para proyectos relacionados a la propagación de ideas o de humanos a través del tiempo y el espacio. Ajustar este tipo de

modelos, habrá que notar, tiende a ser bastante costoso (computacionalmente), debido a que requieren el análisis simultáneo de datos provenientes de un (escenario entero). Es por esto que en la práctica hay dos alternativas usadas comúnmente; la más sencilla es dividir espacialmente el escenario y analizar cada uno de estos grupos de manera separada (como hacen Blackwell y Buck, 2003, por ejemplo). Una alternativa más sofisticada es incluir un modelo mecanicista (normalmente determinista) para el proceso espacial, para después contrastar los patrones espaciales resultantes con la evidencia cronológica usando métodos estadísticos formales (como en Baggaley et al., 2012a,b, por ejemplo).

- Herramientas que produzcan bocetos como el de la figura 4 a partir de bases de datos arqueológicas, para, a su vez, generar modelos cronológicos como los que explicamos aquí. Dye y Buck (2015) presentan una prueba de concepto al respecto, utilizando técnicas de teoría de grafos y una adaptación de las matrices de Harris; sin embargo, aún hay bastante que recorrer antes de que el uso este tipo de herramientas pueda ser rutinario.
- Mejoras a los modelos de edad-profundidad que tomen en cuenta la geomorfología. En la mayoría de los modelos existentes se supone que los núcleos de sedimento lacustre provienen de cuencas cilíndricas. En realidad no conocemos la morfología de los lagos de donde proviene la mayoría de las muestras, pero estamos bastante seguros de que no son cilíndricos; recientemente, Bennett y Buck (2016) muestran que la morfología de la cuenca potencialmente tiene un impacto importante en la relación entre edad y profundidad. Es por esto que es necesario a) desarrollar un método económico de determinar la morfología de una cuenca y b) evolucionar los modelos bayesianos actuales para incluir esta información.

Extender los modelos en estos sentidos llevaría la construcción de cronologías bastante más allá de los modelos, potentes pero simples, que hemos delineado aquí. Algunas de estas pudieran ser añadidas con cierta facilidad a alguna de las alternativas existentes, como OxCal y BCal, aunque otras, en particular las espacio-temporales, requieren de parámetros con estructuras enteramente diferentes a las que hemos usado hasta ahora. Por tanto y por ahora, quien esté interesado en usar este tipo de modelos, se verá en la necesidad de no solo adquirir las habilidades necesarias para crear este tipo de modelos, sino también para implementarlos. Actualmente no contamos con un exceso de gente con estas habilidades dentro de la comunidad de cronologías y deseamos que algún lector entusiasta esté dispuesto a aprenderlas, si es que no las tiene aún, o compartirlas en caso contrario. Nos encantará saber si estás interesado en este tipo de trabajo y te exhortamos a que nos ayudes a que los siguientes treinta años de modelización de cronologías sean tan fructíferos como los anteriores.

Referencias

- Baggaley, A.W., Boys, R.J., Golightly, A., Sarson, G.R. y Shukurov, A. (2012a). Inference for population dynamics in the Neolithic period, *Annals of Applied Statistics*, **6**, 1352–1376.
- Baggaley, A.W., Sarson, G.R., Shukurov, A., Boys, R.J. y Golightly, A. (2012b). Bayesian inference for a wavefront model of the Neolithisation of Europe, *Physical Review E*, **86**.
- Bennett, K.D. y Buck, C.E. (2016). Interpretation of lake sediment accumulation rates, *Holocene*, **26**(7), 1092–1102.
- Blaauw, M. y Christen, J. (2011). Flexible paleoclimate age-depth models using an autoregressive gamma process, *Bayesian Analysis*, **6**, 457–474.
- Blaauw, M. y Christen, J.A. (2005). Radiocarbon peat chronologies and environmental change, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**(4), 805–816, URL <http://dx.doi.org/10.1111/j.1467-9876.2005.00516.x>.
- Blackwell, P.G. y Buck, C.E. (2003). The late glacial human reoccupation of north western Europe: new approaches to space-time modelling, *Antiquity*, **77**(296), 232–240.
- Blackwell, P.G. y Buck, C.E. (2008). Estimating radiocarbon calibration curves, *Bayesian Analysis*, **3**, 225–248.
- Bronk Ramsey, C. (2008). Deposition models for chronological records, *Quaternary Science Reviews*, **27**(1–2), 42–60.
- Bronk Ramsey, C. (2009). Bayesian analysis of radiocarbon dates, *Radiocarbon*, **51**(1), 337–360.
- Buck, C.E., Cavanagh, W.G. y Litton, C.D. (1996). *The Bayesian Approach to Interpreting Archaeological Data*, Chichester: Wiley.
- Buck, C.E. y Christen, J.A. (1998). A novel approach to selecting samples for radiocarbon dating, *Journal of Archaeological Science*, **25**, 303–310.
- Buck, C.E., Christen, J.A. y James, G.N. (1999). BCal: an on-line Bayesian radiocarbon calibration tool, *Internet Archaeology*, **7**, URL <http://intarch.ac.uk/journal/issue7/buck/>.
- Christen, J.A. (1994). Summarizing a set of radiocarbon determinations: a robust approach, *Applied Statistics*, **43**(3), 489–503.
- Christen, J.A. y Buck, C.E. (1998). Sample selection in radiocarbon dating, *Applied Statistics*, **47**, 543–557.

- Christen, J.A., Clymo, R.S. y Litton, C.D. (1995). A Bayesian approach to the use of ^{14}C dates in the estimation of the age of peat, *Radiocarbon*, **37**(2), 431–442.
- Dye, T.S. y Buck, C.E. (2015). Archaeological sequence diagrams and Bayesian chronological models, *Journal of Archaeological Science*, **83**, 84–93.
- Haslett, J. y Parnell, A. (2008). A simple monotone process with application to radiocarbon-dated depth chronologies, *Applied Statistics*, **57**, 399–418.
- Jones, E. (2013). *Practical Bayesian Dendrochronology*, Tesis Doctoral, University of Sheffield.
- Lee, S. y Bronk Ramsey, C. (2012). Development and application of the trapezoidal model for archaeological chronologies, *Radiocarbon*, **54**(1), 107–122.
- Litton, C.D. y Zainodin, H.J. (1991). Statistical models of dendrochronology, *Journal of Archaeological Science*, **18**, 429–440.
- Millard, A. (2002). Bayesian approach to sapwood estimates and felling dates in dendrochronology, *Archaeometry*, **44**(1), 137–143, URL <http://dx.doi.org/10.1111/1475-4754.00048>.
- Millard, A. (2006). Bayesian analysis of ESR dates, with application to Border Cave, *Quaternary Geochronology*, **1**(2), 159–166, URL <http://dx.doi.org/10.1016/j.quageo.2006.03.002>.
- Naylor, J.C. y Smith, A.F.M. (1988). An archaeological inference problem, *Journal of the American Statistical Association*, **83**(403), 588–595.
- Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Haffidason, H., Hajdas, I., Hatté, C., Heaton, T.J., Hoffmann, D.L., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Niu, M., Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Staff, R.A., Turney, C.S.M. y van der Plicht, J. (2013). IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP, *Radiocarbon*, **55**(4), 1869–1887.
- Zink, A. (2015). Bayesian analysis of luminescence measurements, *Radiation Measurements*, **81**, 71–77, URL <http://dx.doi.org/10.1016/j.radmeas.2015.04.009>.