

Estimación de paleotemperaturas vía procesos de difusión integrados

Fernando Baltazar-Larios

Departamento de Matemáticas

Facultad de Ciencias, UNAM

fernandobaltazar@ciencias.unam.com

Resumen

Se presenta un análisis de los registros de la concentración de isótopos de oxígeno en los conos de hielo de Groenlandia, los cuales son usados para estudiar las paleotemperaturas. Estos datos son considerados como observaciones a tiempo discreto de la integral de un proceso de difusión. Obtenemos el estimador máximo verosímil de los parámetros de un proceso de difusión cuando la información que se tiene es una muestra discreta de la integral del proceso.

1. Introducción

En este trabajo se presenta una estimación de las paleotemperaturas a través de la concentración de isótopos de oxígeno en los conos de hielo de Groenlandia. Para dicha estimación se utiliza el método de estimación por máxima verosimilitud para procesos de difusión desarrollado en [3] para los casos que no se observa el proceso de difusión directamente, sino se observan integrales del proceso en intervalos de tiempo disjuntos contaminadas con un error de medición. En [3], dicho método es utilizado para estimar los parámetros del proceso CIR en un modelo de volatilidad estocástica. En este trabajo, dicho método es utilizado para estimar los parámetros de un proceso Ornstein-Uhlenbeck a partir de una muestra de observaciones del proceso integrado de registros de concentración isótopos de oxígeno obtenidas de conos de hielo de Groenlandia.

Los procesos de difusión integrados son de gran importancia en las finanzas para modelos de volatilidad, véase por ejemplo [1],[2],[3] y

[4]. Otra importante aplicación de este tipo de procesos se presenta en el campo de la ingeniería y las ciencias, (véase por ejemplo [13]). La inferencia paramétrica para procesos de difusión integrados ha sido estudiada en [6],[14],[17],[18] y [19]. Para inferencia no paramétrica, se puede consultar [11].

Este trabajo se desarrolla de la siguiente forma: En la Sección 2 se presenta la relación que existe entre la concentración de isótopos de oxígeno en los conos de hielo y las paleotemperaturas, se plantea resolver el problema de la estimación de paleotemperaturas por medio de esta relación y el uso de la metodología desarrollada en [3]. La Sección 3 contiene algunos conceptos básicos de procesos de difusión, en particular se presenta el concepto de ecuación diferencial estocástica, se define a un proceso de difusión y a los procesos de difusión integrados. En la sección 4 se presenta el método desarrollado en [3] para la estimación de los parámetros de un proceso de difusión en el caso que la información con la que se cuenta es una muestra discreta de integrales del proceso y se estudia un ejemplo de simulación para la estimación de los parámetros de un proceso Ornstein-Uhlenbeck. Un análisis de los registros de la concentración de isótopos de oxígeno en los conos de hielo de Groenlandia considerandos como una muestra discreta de integrales de un proceso Ornstein-Uhlenbeck es presentado en la sección 5, se hace la estimación de los parámetros correspondientes utilizando el método descrito en la sección 4 y se realiza una estimación de las paleotemperaturas. En la sección 6 se presentan las conclusiones de este trabajo.

2. Paleotemperaturas y conos de hielo

El clima en la Tierra varía en diferentes escalas espaciales y temporales. Para poder entender la previsibilidad y el origen dinámico de los cambios climáticos necesitamos registros de miles de años de la temperatura de la Tierra. Lamentablemente, los registros de la temperatura y otros parámetros meteorológicos se remontan a poco más de un siglo. Este periodo de observación es demasiado corto para obtener toda la gama de la variabilidad natural, que es el fondo sobre lo que se deben evaluar los posibles cambios climáticos causados por el hombre. Las variaciones climáticas datan de mucho más atrás, siendo las más destacadas las grandes edades de hielo, las cuales se registran a través de marcas geológicas, como las morrenas. Otros registros paleoclimáticos más detallados son el espesor de anillos de árboles, corales, sedimentos de lagos y océanos y registros de los conos de hielo. Todos estos indicadores son grabaciones indirectas del clima del pasado. Algunos de

los registros más largos con alta resolución temporal son los registros de los conos de hielo. Del análisis de estos registros es posible deducir parte del comportamiento dinámico del sistema climático.

Durante la última glaciación y las etapas glaciales anteriores, las capas de hielo expandidas en el hemisferio norte experimentaron aumentos periódicos que produjeron extensos icebergs en el Atlántico Norte, los cuales, a su vez, pueden haber iniciado oscilaciones bruscas en el clima de los llamados ciclos Dansgaard-Oeschger y Eventos-Heinrich (véase [7]).

Una de las características del hielo del glaciar es su tendencia a mantener un registro detallado de los cambios ambientales del pasado mediante la acumulación sucesiva de incrementos anuales de la nieve que bajo presión de sobrecarga se convierte en hielo sólido. Cada incremento anual de la nieve depositada sobre una superficie de capa de hielo atrapa los gases atmosféricos y las impurezas correspondientes a medida que se convierte en hielo sólido, conservando una muestra de los componentes físicos de la columna de aire sobre la superficie del hielo. Los altibajos de estos componentes físicos pueden instigar cambios significativos en el nivel eustático del mar, el albedo¹ y la forma de la superficie sobre la cual se generan.

Desde 1960, varias muestras de hielo han sido extraídas de las capas de hielo polares, incluidos los de Camp Century, Tinte-3, Creta, Milcent y Summit en Groenlandia y los de Byrd, Vostok y Dome de la Antártida. Numerosas características han sido medidas en muestras de hielo para hacer inferencias sobre la historia de la Tierra, algunas de las principales son la razón de isótopos de oxígeno, la concentración de dióxido de carbono, la concentración de metano, la conductividad eléctrica y la concentración de partículas de polvo. Como las variaciones de isótopos de oxígeno que abarca el último ciclo glacial y el Holoceno observadas en los registros de los conos de hielo de Groenlandia y la Antártida muestran fuertes similitudes, esto sugiere que la influencia dominante en las variaciones de isótopos de oxígeno se refleja en los registros de la capa de hielo como el cambio climático regional.

Probablemente, la medida más importante utilizado en el estudio del clima ha sido la proporción de oxígeno 18 a oxígeno 16 ($^{18}O/^{16}O$). La razón isotópica ($^{18}O/^{16}O$) en el hielo, medido como un promedio en cada muestra de hielo, es un indicador de la paleotemperatura promedio durante el intervalo de tiempo en el que cayó la nieve que formó el cono de hielo. Se ha demostrado empíricamente que la diferencia, $\delta^{18}O$, entre el contenido de ^{18}O en el hielo y el contenido de ^{16}O en el agua

¹ Es la reflectividad de la superficie terrestre y se refiere a la energía reflejada desde la Tierra al universo.

del océano actual, tiene una relación aproximadamente lineal constante con la temperatura del aire.

Una relación directa entre las proporciones de isótopos estables en la precipitación de la nieve y la temperatura del vapor de agua en el momento de la precipitación está implícita, suponiendo que no hay cambios en la temperatura y la humedad en las áreas de origen durante todo el período de observación. Actualmente, en la capa de hielo de Groenlandia, la media anual de $\delta^{18}O$ está estrechamente relacionada con la temperatura media anual (T) de la superficie, en grados Celsius, por la ecuación

$$\delta^{18}O = 0,67 \times T - 13,7\%. \quad (1)$$

Esta relación, sin embargo, no se puede mantener durante los cambios temporales en valores de isótopos medidos en los conos de hielo de Groenlandia. La relación temporal entre $\delta^{18}O$ y temperatura de la superficie contemporánea (T) está dada por la siguiente ecuación:

$$T = \alpha + \beta\delta^{18}O + \gamma(\delta^{18}O)^2 \quad (2)$$

donde α, β, γ son determinadas por un ajuste de mínimos cuadrados del perfil de temperatura (véase [12]).

Los conos de hielo de Groenlandia y la Antártida se cortan en trozos cada uno de los cuales representa un intervalo de tiempo en el pasado, al aumentar el tiempo en función de la profundidad.

Las muestras de hielo obtenidas en la Antártica y Groenlandia han proveído datos paleoclimáticos que cubren más de 800 mil años de la historia del clima. Con el fin de interpretar la señal climática proporcionada por los conos de hielo y para facilitar la comparación con otros registros paleoclimáticos, las escalas de tiempo son cruciales. Debido a sus altas tasas de acumulación, los conos de hielo de Groenlandia son muy adecuados para la obtención de una cronología basada en el conteo de la capa anual del último ciclo glacial. Además, los conos de hielo de Groenlandia reflejan fuertemente los cambios bruscos climáticos del último período glacial, los eventos Dansgaard-Oeschger, y contienen muchos horizontes de referencia que permitan la comparación con otros paleo-archivos.

Por otro lado, de acuerdo al comportamiento aleatorio de la variación de las paleotemperaturas en el tiempo resulta natural modelarlas por medio de un proceso de difusión. Bajo este escenario y teniendo observaciones temporales de los registros de las paleotemperaturas se puede hacer inferencia sobre los parámetros de la ecuación diferencial estocástica cuya solución es el proceso de difusión que modela las paleotemperaturas. Diferentes técnicas se han desarrollado para hacer inferencia cuando se tienen observaciones a tiempo discreto de una difusión,

véase por ejemplo [5]. En [13], los autores modelan las paleotemperaturas con un proceso Ornstein-Uhlenbeck para subintervalos adecuados y suponen que el proceso es estacionario. El problema, como se ha hecho mención, es que es difícil contar con registros suficientes para hacer este tipo de inferencia.

Motivados en el problema de falta de información, en este trabajo, se estudia la evolución en el tiempo de las paleotemperaturas utilizando la relación entre estas y la razón isotópica ($^{18}\text{O}/^{16}\text{O}$) dada por la ecuación (2). La información de la razón isotópica ($^{18}\text{O}/^{16}\text{O}$) que se utiliza es obtenida de muestras de conos de hielo de Groenlandia. Estos datos son considerados observaciones a tiempo discreto de la integral de un proceso de difusión con error de medición. Bajo estas condiciones el problema se reduce a estimar los parámetros de la ecuación diferencial estocástica cuya solución es el proceso de difusión que modela la variación en el tiempo de la razón isotópica ($^{18}\text{O}/^{16}\text{O}$) en los conos de hielo. Para resolver el problema de la estimación de los parámetros de la ecuación diferencial estocástica cuando se cuenta con información de la integral del proceso de difusión se utilizará el modelo propuesto en [3].

Para lograr la estimación de las paleotemperaturas en las siguientes secciones se presentan los conceptos básicos de procesos de difusión y el modelo que se utilizará para la estimación de los parámetros de una ecuación diferencial estocástica.

3. Procesos de difusión integrados

El modelo matemático básico en la teoría de la probabilidad es el conocido como **espacio de probabilidad**, que se representa matemáticamente por la terna $(\Omega, \mathcal{F}, \mathbb{P})$ consistente de un espacio muestral Ω , una σ -álgebra \mathcal{F} y una medida de probabilidad \mathbb{P} .

Definición 3.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, una variable aleatoria es una función $X : \Omega \rightarrow \mathbb{R}$ que satisface

$$X^{-1}(B) = \{\omega \in \Omega | X(\omega) \in B\} \in \mathcal{F}$$

para todo $B \in \mathcal{B}(\mathbb{R})$, donde $\mathcal{B}(\mathbb{R})$ es la σ -álgebra de Borel.

Podemos representar la evolución de procesos que varían en el tiempo y de manera aleatoria por medio de variables aleatorias que representen el estado del sistema en cada instante de tiempo. De esta manera llegamos al concepto de proceso estocástico.

Definición 3.2. Un *proceso estocástico* es una colección de variables aleatorias $X = \{X_t\}_{t \in T}$ parametrizada por un conjunto T , definidas sobre el mismo espacio de probabilidad.

Al conjunto T se le interpreta como el conjunto de tiempos y es llamado espacio parametral. En este trabajo consideramos procesos donde las variables aleatorias toman valores reales y $T = [0, \infty)$. De este modo podemos ver a un proceso estocástico como una función de dos variables

$$X : T \times \Omega \rightarrow \mathbb{R},$$

tal que para cada $t \in T$, la función $\omega \rightarrow X_t(\omega)$ es una variable aleatoria y para cada $\omega \in \Omega$, la función $t \rightarrow X_t(\omega)$ es una **realización** o **trayectoria** del proceso.

3.0.1. Proceso de Wiener (Movimiento Browniano)

En 1828 el botánico Robert Brown observó que los granos de polen suspendidos en una cierta substancia presentan movimientos irregulares (véase [9]). Con la contribución del trabajo presentado por Albert Einstein en 1905 (véase [15]), se explicó que este movimiento es producto de las múltiples colisiones aleatorias de las moléculas de la substancia con los granos de polen. Se puede decir que un movimiento de este tipo tiene las siguientes características:

1. Es un movimiento continuo.
2. Tiene desplazamientos independientes en intervalos de tiempo disjuntos.
3. Los desplazamientos pueden modelarse por variables aleatorias gaussianas (teorema central del límite, véase [16]).

En 1923 el matemático Norbert Wiener demostró que existe un proceso estocástico con estas propiedades.

Definición 3.3. Un *proceso de Wiener (movimiento browniano)* estándar unidimensional es un proceso estocástico $W = \{W_t\}_{t \geq 0}$ tal que

1. $W_0 = 0$ seguramente (esto es, con probabilidad uno).
2. Tiene trayectorias continuas.
3. Tiene incrementos independientes.
4. La variable aleatoria $W_t - W_s$ tiene distribución $N(0, t - s)$ para $0 \leq s < t$.

3.1 Procesos de difusión

Es común modelar la evolución en el tiempo de un sistema por ecuaciones diferenciales, las cuales usualmente describen la razón de cambio del sistema. Si denotamos por X_t el estado del sistema al tiempo t para $t \geq 0$, podemos caracterizar la evolución del sistema en el tiempo por la siguiente ecuación,

$$\frac{d(X_t)}{dt} = a(t)X_t, \quad (3)$$

donde $X_0 = x_0$ es el estado inicial del sistema y $a(t)$ representa la función de cambio al tiempo t . En muchos casos esta función no es completamente conocida, pero es razonable asumir que está sujeta a algún tipo de efecto aleatorio, entonces tenemos que

$$a(t) = b(t) + \text{ruido}, \quad (4)$$

y sabemos que el ruido está regido por una distribución de probabilidad. Entonces una forma (heurística) más general de escribir a la ecuación (4) es:

$$\frac{d(X_t)}{dt} = b(t, X_t) + \sigma(t, X_t) \times \text{ruido}, \quad (5)$$

donde b, σ son funciones reales conocidas. Para modelar el término de ruido se utilizan los incrementos de un proceso de Wiener. Si la evolución del sistema es observado en el intervalo de tiempo $[0, T]$ y consideramos una partición de este intervalo, $0 = t_0 < t_1 < \dots < t_n = T$, utilizando la ecuación (5) podemos describir la evolución del proceso $X = \{X_t\}_{t=0}^T$ por:

$$X_t = X_0 + \sum_{i=1}^{n-1} b(t_i, X_i) \Delta t_i + \sum_{i=1}^{n-1} \sigma(t_i, X_i) \Delta W_i, \quad (6)$$

donde $X_i = X_{t_i}$, $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ y $\Delta t_i = t_{i+1} - t_i$. Dado que esta ecuación depende de la partición del intervalo, el modelo resulta ser más adecuado en el límite $\Delta t_i \rightarrow 0$. Esto nos lleva a la siguiente ecuación

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \quad (7)$$

para toda $t \geq 0$ y $X_0 = x_0$. A ecuaciones de este tipo se les conoce como ecuaciones diferenciales estocásticas.

Definición 3.4. Un *proceso de difusión* $X = \{X_t\}_{t \geq 0}$ es la solución a una ecuación de la forma (7), donde los coeficientes $b(t, x)$ y $\sigma(t, x)$ son funciones reales de t y x , y se les conoce como coeficiente de deriva y difusión respectivamente.

En la definición anterior $b(t, x)$ y $\sigma(t, x)$ pueden ser interpretados como la media y varianza infinitesimal de los incrementos del proceso X . Al proceso de difusión se le interpreta como el estado del sistema que evoluciona de manera determinista gobernado por la parte no aleatoria de la ecuación pero perturbado por un ruido aditivo dado por la integral estocástica.

Para que una ecuación de la forma (7) tenga alguna solución se deben imponer condiciones a sus coeficientes. Existen resultados de existencia y unicidad para ecuaciones diferenciales estocásticas que establecen condiciones de regularidad para sus coeficientes (véase [20]).

3.2 Procesos de difusión integrados

Consideremos el proceso de difusión unidimensional $X = \{X_t\}_{t \geq 0}$ dado por la ecuación diferencial estocástica (7). Supongamos que el intervalo de observación de las trayectorias de la difusión es $[0, T]$, pero, solo una integral del proceso con respecto a alguna función de peso está disponible. Más específicamente, supongamos que el intervalo de observación $[0, T]$ se subdivide en n intervalos más pequeños de longitud $\Delta = T/n$ y sea ν una medida de probabilidad en el intervalo $[0, \Delta]$. Entonces, consideramos observaciones de la forma

$$Y_i = \int_0^\Delta X_{(i-1)\Delta+s} d\nu(s) \quad (8)$$

para $i = 1, \dots, n$.

Suponiendo que ν tiene una densidad π con respecto a la medida de Lebesgue en $[0, \Delta]$, podemos escribir

$$Y_i = \int_{(i-1)\Delta}^{i\Delta} X_s \pi(s - (i-1)\Delta) ds \quad (9)$$

para $i = 1, \dots, n$.

Si las observaciones se obtienen mediante la integración uniforme sobre el eje de tiempo, entonces ν es simplemente la distribución uniforme en $[0, \Delta]$ con $\pi = 1/\Delta$ y podemos escribir a la ecuación (9) como

$$Y_i = \frac{1}{\Delta} \int_{(i-1)\Delta}^{i\Delta} X_s ds$$

para $i = 1, \dots, n$.

Dado que el proceso X es invariante bajo traslaciones de tiempo por estacionalidad, el proceso $Y = \{Y_i\}_{i=1}^n$ es estacionario.

4. Estimador máximo verosímil para procesos de difusión integrados

En esta sección se presenta el algoritmo desarrollado en [3] para la estimación de los parámetros de un proceso de difusión cuando no se observa de manera directa, pero se cuenta con observaciones de integrales de dicho proceso en intervalos de tiempo disjuntos y estas observaciones están contaminadas por errores de medición.

4.1 Modelo

Se considera la estimación por máxima verosimilitud para un proceso de difusión $X = \{X_t\}_{t \geq 0}$ que es solución a la ecuación diferencial estocástica

$$dX_t = b(t, X_t)dt + \sigma(t, X_t; \psi)dW_t, \quad (10)$$

donde $W = \{W_t\}$ es un proceso de Wiener estándar y los coeficientes dependen de un parámetro p -dimensional ψ desconocido que pertenece al conjunto parametral $\Psi \subset \mathbb{R}^p$. Suponemos que X es un proceso ergódico y estacionario con distribución invariante $\nu_\psi(x)$ (independiente de W). Suponemos que la ecuación (10) tiene una única solución débil.

Los datos que se tienen para hacer la estimación de los parámetros son integrales de X en intervalos $[t_{i-1}, t_i]$ observados con error de medición, es decir

$$Y_i = \int_{t_{i-1}}^{t_i} X_s ds + Z_i, \quad i = 1, \dots, n, \quad (11)$$

donde $Z_i \sim N(0, \tau^2)$, $i = 1, \dots, n$ son mutuamente independientes e independientes de X . Supondremos que $t_0 = 0$ y el intervalo total de observación es $[0, t_n]$. Como la varianza de la medición de error, τ^2 , es otro parámetro desconocido, entonces se necesita estimar el parámetro $p + 1$ -dimensional $\theta = (\psi, \tau^2)$.

Condicionado en la realización de X , las observaciones $Y_i, i = 1, \dots, n$ son independientes y con distribución normal

$$Y_i | X_t : t \in [0, t_n] \sim N \left(\int_{t_{i-1}}^{t_i} X_s ds, \tau^2 \right). \quad (12)$$

Por otro lado, suponemos que los coeficientes de la ecuación (10) satisfacen las condiciones usuales de regularidad para todo $\psi \in \Psi$, véase [3].

4.2 Función de verosimilitud y algoritmo EM

Se consideran los datos $Y = (Y_1, \dots, Y_n)$ como información incompleta de la realización completa $X_t, t \in [0, t_n]$ y la medición de los errores Z_1, \dots, Z_n . Entonces, se puede hacer estimación con base en la función de verosimilitud utilizando el algoritmo EM o métodos de Monte Carlo vía Cadenas de Markov (véase [8]). En [3] se utiliza el algoritmo EM. Para realizar la estimación es necesario encontrar la función de verosimilitud de los datos completos y la esperanza condicional dadas las observaciones $Y = (Y_1, \dots, Y_n)$.

Como la observación completa de la realización de una difusión en el intervalo de tiempo $[0, t_n]$ es elemento del espacio C de funciones continuas de $[0, t_n]$ a \mathbb{R} , consideramos la σ -álgebra \mathcal{C} generada por los cilindros y las medidas de probabilidad inducidas en (C, \mathcal{C}) por las soluciones de la ecuación (10). Para obtener la función de verosimilitud se usa la transformación estándar

$$h(x; \psi) = \int_{x^*}^x \frac{1}{\sigma(u; \psi)} du. \quad (13)$$

donde x^* es un elemento arbitrario del espacio de estados de X . Por la fórmula de Itô se tiene que

$$U_t = h(X_t; \psi)$$

satisface la ecuación diferencial estocástica

$$dU_t = \mu(U_t; \psi)dt + dW_t, \quad (14)$$

con deriva dada por

$$\mu(u; \psi) = \frac{b(h^{-1}(u; \psi); \psi)}{\sigma(h^{-1}(u; \psi); \psi)} - \frac{\sigma'(h^{-1}(u; \psi); \psi)}{2},$$

donde σ' es la derivada de σ con respecto a x . En la ecuación (14) el coeficiente de difusión no depende de los parámetros y por tanto la medida de probabilidad inducida en (C, \mathcal{C}) por la solución de la ecuación (14) es equivalente y la función de verosimilitud puede ser encontrada.

Expresando a las observaciones Y_i en términos del proceso U tenemos que

$$Y_i = \int_{t_{i-1}}^{t_i} h^{-1}(U_s; \psi) ds + Z_i, \quad i = 1, \dots, n.$$

Ahora consideramos que la información completa está dada por U_t , $t \in [0, t_n]$ y $Y = (Y_1, \dots, Y_n)$. Dado que condicionada en las trayectorias de U las observaciones Y son independientes, la verosimilitud de Y condicionada en la realización de U en $[0, t_n]$ es

$$L(Y_1, \dots, Y_n | U_t, t \in [0, t_n]) = \prod_{i=1}^n \phi(Y_i; \int_{t_{i-1}}^{t_i} h^{-1}(U_s; \psi) ds, \tau^2), \quad (15)$$

donde $\phi(u; a_1, a_2)$ denota la densidad de una distribución normal con media a_1 y varianza a_2 evaluada en u .

Sea P_ψ la medida de probabilidad inducida por $U = \{U_t\}_{t \in [0, t_n]}$ y Q la medida de Wiener en (C, \mathcal{C}) . Suponiendo que el coeficiente μ satisface las condiciones que garantizan que se puede aplicar el teorema de

Girsanov tenemos que la derivada de Radon-Nykodym es

$$\frac{dP_\psi}{dQ}(B) = \exp \left\{ a(B_{t_n}) - a(B_0) - \frac{1}{2} \int_0^{t_n} [\mu(B_t; \psi)^2 + \mu'(B_t; \psi)] dt \right\}, \quad (16)$$

donde μ' denota la derivada de $\mu(u; \psi)$ con respecto a u y $a(x; \psi) = \int^x \mu(u; \psi) du$.

Combinando las ecuaciones (16) y (15) se tiene que la función de log-verosimilitud de θ basado en la información completa $U_t, t \in [0, t_n]$ y $Y = (Y_1, \dots, Y_n)$ está dada por

$$\begin{aligned} \log L(\theta; Y, U_t, t \in [0, t_n]) &= \sum_{i=1}^n \log \phi(Y_i; \int_{t_{i-1}}^{t_i} h^{-1}(U_s; \psi) ds, \tau^2) \\ &+ a(U_{t_n}; \psi) - a(U_0; \psi) - \frac{1}{2} \int_0^{t_n} (\mu(U_t; \psi)^2 + \mu'(U_t; \psi)) dt. \end{aligned}$$

Sea $\hat{\theta}$ cualquier valor del vector parametral θ . Entonces el algoritmo EM trabaja como sigue:

1. **Paso-E.** Generar M trayectorias del proceso de difusión $X, X^{(k)}, k = 1, \dots, M$, condicionadas en las observaciones Y_1, \dots, Y_n usando el valor del parámetro $\hat{\theta} = (\hat{\psi}, \hat{\tau}^2)$, y calcular

$$g(\theta) = \frac{1}{M - M_0} \sum_{k=M_0+1}^M \log L(\theta; Y_1, \dots, Y_n, h(X_t^{(k)}; \hat{\psi}), t \in [0, t_n]),$$

con un adecuado periodo de calentamiento M_0 y M suficientemente grande.

2. **Paso-M.** $\hat{\theta} = \operatorname{argmax} g(\theta)$.
3. Regresar a 1.

Para implementar este algoritmo, el principal problema es cómo generar trayectorias de X condicionadas en $Y = (Y_1, \dots, Y_n)$ donde la relación entre las Y_i 's y X está dada por la ecuación (12). Utilizando este algoritmo se obtiene una sucesión $X^{(k)}, k = 1, \dots, M$ para un valor de M suficientemente grande para asegurar que $g(\theta)$ aproxima de manera adecuada a la función de log-verosimilitud completa dada la información que se tiene.

Para generar muestras de

$$\{X_t; t \in [0, t_n]\} | (Y_1, \dots, Y_n)$$

para algún valor del parámetro θ , es decir, simular la difusión X condicionada en las observaciones $Y = (Y_1, \dots, Y_n)$ de integrales de X en subintervalos $[t_{j-1}, t_j], j = 1, \dots, n$ perturbadas por errores de medición, utilizamos el siguiente algoritmo propuesto y discutido a

detalle en [3]:

Algoritmo 1.

1. Generar una trayectoria inicial $\{X_t^{(0)} : t \in [0, t_n]\}$.
2. Sea $l = 1$.
3. Generar una trayectoria $\{X_t^{(l)} : t \in [0, t_n]\}$ condicionada con Y actualizando el subconjunto de trayectorias:
 - (a) Aleatoriamente se divide el intervalo de tiempo en K bloques

$$0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_K = t_n.$$

- (b) Obtenemos $X_0^{(l)}$ y simulamos la subtrayectoria condicional

$$\{X_t^{(l)} : t \in [\tau_{k-1}, \tau_k]\} | Y_{\{k\}}, X_{\tau_{k-1}}^{(l)}, X_{\tau_k}^{(l-1)} \quad (17)$$

para $k = 1, \dots, K - 1$. Finalmente, simulamos la trayectoria de

$$\{X_t^{(l)} : t \in [\tau_{K-1}, \tau_K]\} | Y_{\{K\}}, X_{\tau_{K-1}}^{(l)}.$$

4. $l=l+1$.
5. Ir a 3.

El método es una modificación del método presentado en [10], donde se usa el algoritmo para la simulación de puentes de difusión propuesto en [5]. En [3] se puede encontrar un algoritmo para subdividir los intervalos aleatoriamente. El punto principal de este algoritmo es cómo generar muestras de variables del tipo (17), que son puentes de difusión no lineales. Para resolver este problema, se utiliza el método para la simulación de puentes de difusión propuesto en [5]. Muestras de trayectorias de puentes condicionados de la forma (17) pueden ser generados por medio del algoritmo Metropolis-Hastings. Denotando a un puente de difusión por (t, a, s, b) en el intervalo $[t, s]$ con $X_t = a$ y $X_s = b$, después de un adecuado periodo de calentamiento, el siguiente algoritmo, propuesto en [3], genera muestras de un puente de difusión $(\tau_{k-1}, a, \tau_k, b)$ condicionado a $Y_{\{k\}}$, donde $Y_{\{k\}}$ denota las observaciones Y_i 's en el intervalo $(\tau_{k-1}, \tau_k]$:

Algoritmo 2

1. Simular un puente $(\tau_{k-1}, a, \tau_k, b)$, $X^{(0)}$, y hacer $l = 1$.
2. Proponer una nueva trayectoria simulando un puente $(\tau_{k-1}, a, \tau_k, b)$, $X^{(l)}$.

3. Aceptar el puente de difusión con probabilidad

$$\min \left(1, \prod_{i=1}^{n_k} \frac{\phi(Y_{j+i}; \int_{t_{j+i-1}}^{t_{j+i}} X_s^{(l)} ds, \tau^2)}{\phi(Y_{j+i}; \int_{t_{j+i-1}}^{t_{j+i}} X_s^{(l-1)} ds, \tau^2)} \right).$$

En otro caso hacemos $X^{(l)} = X^{(l-1)}$.

4. Tomamos $l = l + 1$ y regresamos a 2.

4.3 Un estudio de simulación

En esta sección aplicaremos el método descrito al proceso Ornstein-Uhlenbeck, el cual es solución a la ecuación

$$dX_t = -\alpha X_t dt + \sigma dW_t, \quad (18)$$

donde $\alpha, \sigma > 0$ son constantes y $X_0 = x$. (18). Suponemos que α y σ son parámetros desconocidos para ser estimados por medio del método descrito en la sección anterior.

4.3.1. Verosimilitud y algoritmo EM

Para el proceso Ornstein-Uhlenbeck la transformación (13) está dada por

$$h(x; \sigma) = \frac{x}{\sigma},$$

de este modo $U_t = X_t/\sigma$, es solución de la ecuación

$$dU_t = -\alpha U_t dt + dW_t.$$

Tenemos que $\mu(u; \alpha, \sigma) = -\alpha u$ y $a(u; \alpha, \sigma) = -\alpha u^2/2$ y la función de log-verosimilitud es

$$\begin{aligned} \log L(\theta; Y, U_t, t \in [0, t_n]) &= \sum_{i=1}^n \log \phi(Y_i; \sigma \int_{t_{i-1}}^{t_i} U_s ds, \tau^2) \\ &\quad + \frac{\alpha}{2} (U_0^2 - U_{t_n}^2 + t_n) - \frac{\alpha^2}{2} \int_0^{t_n} U_t^2 dt, \end{aligned}$$

donde $\theta = (\alpha, \sigma, \tau^2)$ y el algoritmo EM trabaja como sigue:

Paso E La función objetivo $g(\theta)$ para el proceso Ornstein-Uhlenbeck

está dada por

$$\begin{aligned}
g(\theta) &= \frac{1}{M - M_0} \sum_{k=M_0+1}^M \sum_{i=1}^n -\frac{(Y_i - \sigma \int_{t_{i-1}}^{t_i} U_t^{(m)} dt)^2}{2\tau^2} - \frac{n}{2} \log(2\pi\tau^2) \\
&+ \frac{\alpha}{2} t_n + \frac{\alpha}{2(M - M_0)} \sum_{k=M_0+1}^M ((U_0^{(m)})^2 - (U_{t_n}^{(m)})^2) \\
&- \frac{\alpha^2}{2(M - M_0)} \sum_{k=M_0+1}^M \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (U_t^{(m)})^2 dt,
\end{aligned}$$

donde $U_t^{(m)}$ es la m -ésima trayectoria del proceso U condicionado a las observaciones Y simulado por medio de los algoritmos 1 y 2 con los parámetros obtenidos en un paso anterior.

Paso M Tenemos que la función $g(\theta)$ se maximiza en:

$$\hat{\alpha} = \frac{t_n(M - M_0) + \sum_{k=M_0+1}^M ((U_0^{(m)})^2 - (U_{t_n}^{(m)})^2)}{2 \sum_{k=M_0+1}^M \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (U_t^{(m)})^2 dt} \quad (19)$$

y

$$\hat{\sigma} = \frac{\sum_{k=M_0+1}^M M \sum_{i=1}^n Y_i \int_{t_{i-1}}^{t_i} U_t^{(m)} dt}{\sum_{k=M_0+1}^M \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (U_t^{(m)})^2 dt}. \quad (20)$$

Se simularon 1,000 conjuntos de datos estimando los parámetros con el método propuesto. Cada conjunto de datos fue obtenido de la simulación de trayectorias de longitud 1,500 con distribución inicial $X_0 \sim N(0, \sigma^2/(2\alpha))$, y entonces calculando los datos Y_i , $i = 1, \dots, 1500$ de la ecuación (11) con $t_i = i$, $i = 0, \dots, n$. Los parámetros usados fueron $\alpha = 0,1$, $\sigma = 0,5$ y $\tau^2 = 1,25$. El algoritmo EM fue corrido con $M = 10,000$ y $M_0 = 1,000$ y dividiendo aleatoriamente los intervalos modelados por una variable aleatoria Poisson(λ) (véase [16]). Las estimaciones obtenidas se presenta en el cuadro 1

λ	α	σ	τ^2
10	0.106	0.523	1.229
20	0.101	0.507	1.235
30	0.084	0.458	1.252

Cuadro 1. Promedio de los parámetros obtenidos de 1000 base de datos simulados con valores parametrales $\alpha = 0,1$, $\sigma = 0,5$ y $\tau^2 = 1,25$.

5. Análisis de paleotemperaturas

La información que se analiza en este trabajo son muestras de conos de hielo de Groenlandia (véase [21]). Los datos $\delta^{18}O$ provienen de 60,000 años pasados y cada dato es el promedio de un intervalo de 20 años. De este modo tenemos una base de datos de 3,000 elementos. Se puede ver una gráfica de los datos en la figura 1.

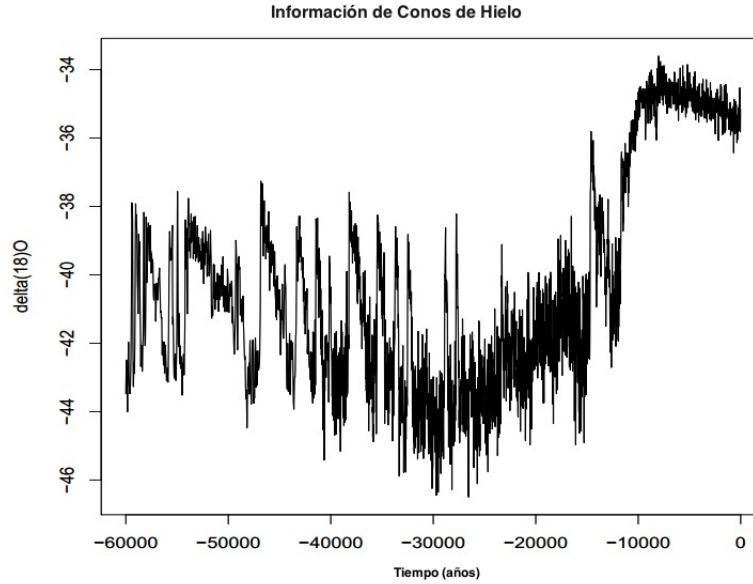


Figura 1. Información de los conos de hielo.

Como puede observarse en la serie de datos, obviamente no se trata de un proceso estacionario. Después de un análisis preliminar los datos se dividieron en tres subintervalos, $[-60000, -30000]$, $[-30000, -10000]$ y $[-10000, 0]$. Para cada uno de estos subintervalos, las observaciones correspondientes fueron modeladas como observaciones de la forma (11) donde X es un proceso Ornstein-Uhlenbeck y las integrales son consideradas en intervalos de 20 años de longitud.

Un resumen de los resultados de la corrida del algoritmo EM está en el cuadro 2. El algoritmo EM fue corrido con $M = 10000$, $M_0 = 1000$, $\lambda = 20$ para cada intervalo. El parámetro α^{-1} puede interpretarse como un correlación, y $\sigma/\sqrt{2\alpha}$ es la desviación estándar de la distribución estacionaria.

Por otro lado, utilizando la ecuación (2) y obteniendo un ajuste por mínimos cuadrados para la temperatura, la correspondiente ecuación es

$$T = 20,45 + 1493\delta^{18}O + 1,052 \times (10^{-10})\delta^{18}O^2, \quad (21)$$

Parameter	Block [-10000,0]	[-30000,-10000]	[-60000,-30000]
α^{-1}	205.3	669.8	321.1
σ	0.0303	0.1167	0.1395
$\sigma/\sqrt{2\alpha}$	0.307	2.136	1.767
τ^2	0.1062638	0.6966518	0.225974

Cuadro 2. Valores estimados obtenidos por el algoritmo EM para las muestras de conos de hielo usando el proceso Ornstein Uhlenbeck para cada uno de los tres intervalos.

para esta estimación se utilizó solo la información del subintervalo $[-10000, 0]$ que corresponde a la época actual. Con los resultados obtenidos en la estimación del proceso de difusión integrado podemos hacer una proyección anual del proceso de difusión y así obtener una estimación de la temperatura promedio anual para el periodo correspondiente utilizando la ecuación (21). En la figura 2 presentamos los resultados correspondientes.

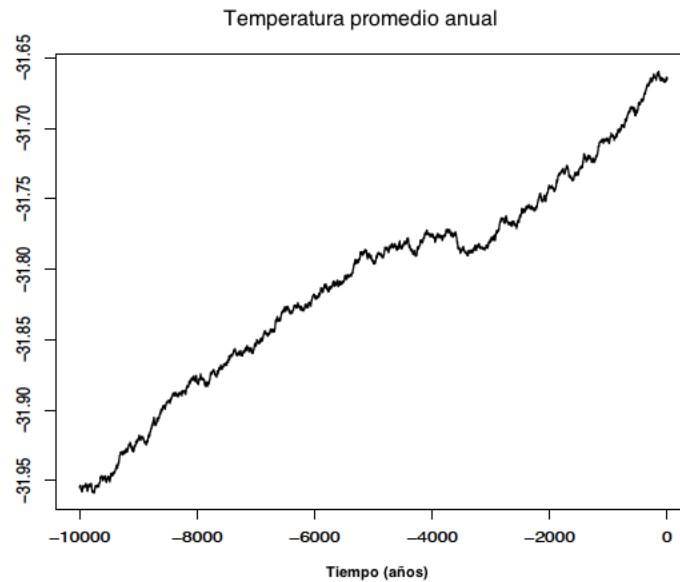


Figura 2. Temperatura promedio anual de los últimos 10,000 años en grados Celsius

6. Conclusiones

En este trabajo se presentó una forma de hacer estimación de la temperatura de la Tierra con base en la variación de la razón isotópica

($^{18}\text{O}/^{16}\text{O}$) en conos de hielo en el tiempo. Para resolver el problema de la falta de registros de las paleotemperaturas se utilizó su relación con la razón isotópica ($^{18}\text{O}/^{16}\text{O}$) en conos de hielo de Groenlandia, de los cuales sí se tiene información suficiente. La variación de la paleotemperatura puede ser modelada por una ecuación diferencial estocástica, y es natural modelar los datos de los conos de hielo como un proceso de difusión integrado, de este modo el problema se redujo a la estimación de los parámetros de dicha ecuación diferencial estocástica.

Para resolver este problema se consideró a los datos de los conos de hielo como observaciones a tiempo discreto de la integral de un proceso de difusión contaminadas por error de medición y utilizando el algoritmo presentado en [3] para hacer estimación de los parámetros de la ecuación diferencial estocástica correspondiente se presentó una estimación de la razón isotópica ($^{18}\text{O}/^{16}\text{O}$). Dada la estimación de la razón isotópica ($^{18}\text{O}/^{16}\text{O}$) se utilizó la ecuación (2) para hacer una estimación de la temperatura. Se presentó una estimación de la temperatura promedio anual de los últimos 10,000 años. Utilizando las estimaciones de las temperaturas presentadas en este trabajo se puede hacer estimaciones de las temperaturas futuras.

Bibliografía

- [1] T. G. Andersen, T. Bollerslev, F. X. Diebold y H. Ebens, «The distribution of realized stock return volatility», *J. Finan. Econ.*, vol. 61, 2001a, 43–76.
- [2] T. G. Andersen, T. Bollerslev, F. Diebold y P. Labys, «The distribution of exchange rate volatility», *J. Amer. Statist. Assoc.*, vol. 96, 2001b, 42–55.
- [3] F. Baltazar-Larios y M. Sørensen, *Maximum likelihood estimation for integrated diffusion processes*, Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen, Springer, 2010.
- [4] O. E. Barndorff-Nielsen y N. Shephard, «Econometric analysis of realized volatility and its use in estimating stochastic volatility models», *J. R. Statist. Soc B*, vol. 64, 2002, 253–280.
- [5] M. Bladt y M. Sørensen, «Simple simulation of diffusion bridges with application to likelihood inference for diffusions», 2009, Working paper.
- [6] T. Bollerslev y H. Zhou, «Estimating stochastic volatility diffusion using conditional moments of integrated volatility», *Journal of Econometrics*, vol. 109, 2002, 33–65.
- [7] W. Broecker, «Massive icebergs as triggers for global climate change», *Nature*, vol. 372, 1994, 421–424.
- [8] S. Brooks, «Markov chain monte carlo method and its application», *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 47, 1998, 69–100.
- [9] R. Brown, «A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies», *Philosophical Magazine N. S.*, vol. 4, 1828, 161–173.
- [10] S. Chib, M. Pitt y N. Shephard, «Likelihood based inference for diffusion driven state space models», 2006.

- [11] F. Comte, V. Genon-Catalot y Y. Rozenholc, «Nonparametric adaptive estimation for integrated diffusions», *Stochastic processes and their applications*, vol. 119, 2009, 811–834.
- [12] P. Ditlevsen, *Turbulence and climate dynamics*, J and R Frydeberg A/S, Copenhagen, 2004.
- [13] P. D. Ditlevsen, S. Ditlevsen y K. K. Andersen, «The fast climate fluctuations during the stadial and interstadial climate states», *Annals of Glaciology*, vol. 35, 2002, 457–462.
- [14] S. Ditlevsen y M. Sørensen, «Inference for observations of integrated diffusion processes», *Scandinavian Journal of Statistics*, vol. 31, 2004, 417–429.
- [15] A. Einstein, *Investigations on the theory of the brownian movement*, Dover, 1956.
- [16] W. Feller, *An introduction to probability theory and its applications*, 3.^a ed., Wiley Series in Probability and Mathematical Statistics. Vol. 1, 1968.
- [17] J. L. Forman y M. Sørensen, «The pearson diffusions: A class of statistically tractable diffusion processes», *Scand. J. Statist.*, vol. 35, 2008, 438–465.
- [18] A. Gloter, «Parameter estimation for a discrete sampling of an integrated ornstein-uhlenbeck process», *Statistics*, vol. 35, 2000, 225–243.
- [19] ———, «Parameter estimation for a discretely observed integrated diffusion process», *Scandinavian Journal of Statistics*, vol. 33, 2006, 83–104.
- [20] B. Øksendal, *Stochastic differential equations*, Springer, New York, 1998.
- [21] A. Svensson, K. Andersen, M. Bigler, H. Clausen, D. Dahl-Jensen, S. M. Davies, S. J. Johnsen, R. Muscheler, F. Parrenin, S. O. Rasmussen, R. Röthlisberger, I. Seierstad, J. P. Steffesen y B. M. Vinther, «A 60 000 year greenland stratigraphic ice core chronology», *Climate of the Past*, vol. 4, 2008, 47–57.