

ESTUDIO DEL CAUDAL A TRAVÉS DE MODELOS LINEALES GENERALIZADOS DINÁMICOS

¹Bermúdez Dagoberto, ²D'Achiardi Elmen

Resumen. Este documento presenta una introducción a los modelos lineales dinámicos normales y generalizados que son una herramienta importante en el análisis de series de tiempo bajo un contexto bayesiano, donde el cambio dinámico de sus parámetros permite una adaptación a las circunstancias en que evolucionan los datos, facilitando su análisis e interpretación. Dicha metodología se implementará en el modelamiento del caudal y la precipitación para los datos recolectados en la estación Cabrera del río Sumapaz ubicado en la región central del territorio colombiano.

Palabras claves: Modelo lineal dinámico, inferencia Bayesiana, precipitación, caudal, Cadenas de Markov Monte Carlo.

Abstract. This document presents an introduction to the normal dynamic linear models and generalized that are an important tool in the analysis of time series under a Bayesian context, where the dynamic change of its parameters can be adapted to evolving circumstances in the data, making it easier their analysis and interpretation. This methodology is implemented for modeling the data of flow and precipitation collected at the station Cabrera Sumapaz River located in the central region of Colombia.

Key words. Bayesian inference, dynamic linear model, flow, Monte Carlo Markov Chains, rainfall.

Recibido: Noviembre, 2010

Aceptado: Febrero, 2011

1. INTRODUCCIÓN

Desde el trabajo pionero de Harrison & Stevens (1976), los Modelos Lineales Dinámicos (MLD's) se han convertido en una herramienta muy importante para el análisis de series temporales bajo el contexto Bayesiano, tal como indican los numerosos ejemplos y referencias listadas en (West & Harrison 1997). Varias son las razones que han contribuido al éxito de estos modelos. En primer lugar, el principio de modelamiento por componentes que guía su construcción, así como el dinamismo de sus parámetros que les permite adaptarse a las circunstancias en las que se desarrolla la evolución de la serie, facilitan el análisis y la interpretación de los resultados obtenidos.

Por otra parte, el procesamiento secuencial bayesiano de la información disponible en cada periodo de tiempo permite al analista incorporar, en tiempo real, informaciones externas a los propios datos, involucrando de forma directa en el proceso de modelamiento de la evolución de la serie lo cual redundará en beneficio de una toma de decisiones más adecuada. Finalmente, el avance espectacular de las técnicas de cálculo Bayesianas acontecido sobre todo a partir del desarrollo de los métodos Markov Chain Monte Carlo MCMC, han hecho posible la

construcción de modelos más flexibles y realistas ampliando así su marco de aplicación (ver, Capítulo 15 de (West & Harrison 1997)). Los modelos dinámicos Bayesianos son desarrollados para aplicaciones en problemas de regresión y en series de tiempo no normales y no lineales, proporcionando extensiones dinámicas de los modelos lineales normales generalizados. Una característica clave de este análisis es el uso de las distribuciones a priori y a posteriori conjugada para parámetros de la familia exponencial. La estructura del modelo depende de la evolución temporal de variables de estado, y la retroalimentación de la información observacional.

2. RESEÑA HISTÓRICA

El enfoque de modelamiento y predicción sintetiza conceptos, modelos y métodos cuyo desarrollo se ha visto influido por el trabajo en varias áreas. Mucho de lo involucrado en el modelamiento no puede ser descrito usando estructuras matemáticas formales, sobre todo en las etapas de formulación y elección del modelo; sin embargo, de acuerdo con los conceptos fundamentales del método científico, se identifica el enfoque bayesiano como marco para el aprendizaje de rutina y organización del conocimiento incierto dentro de los sistemas de pronóstico.

Durante los últimos sesenta años se ha incrementado rápidamente su apoyo al enfoque bayesiano como medio de aprendizaje científico y

¹ Bermúdez Rubio Dagoberto, Ing., Universidad Nacional de Colombia – Sede Bogotá. (e_mail: dbermudezr@unal.edu.co)

² D'Achiardi Barreto Elmen Damián, Ing., Universidad Nacional de Colombia – Sede Bogotá. (e_mail: eddachiardib@unal.edu.co)

toma de decisiones gracias a la reciente aceptación notable de los profesionales impulsados por la necesidad de adoptar el sentido común a los principios en que se fundamenta. Todos nos enfrentamos a decisiones que tienen que hacerse sin posibilidad de repetir la experiencia, así que el valor de la estadística clásica es muy limitado para los investigadores que se enfrentan a una toma de decisiones.

A principios de la década de los sesenta, se hizo evidente que en los modelos lineales dinámicos normales con varianzas conocidas y relaciones de recurrencia para la actualización secuencial de las distribuciones a posteriori son esencialmente equivalentes a las ecuaciones del Filtro de Kalman, es obvio que no como muchos suelen creer, que el pronóstico bayesiano se basa en el filtro de Kalman (ver (Harrison & Stevens 1976), y la discusión, y respuesta a la discusión por Davis et. al. 1985).

Hacia el final de la década de 1970, (Harrison & Stevens 1976) define la clase de modelos lineales dinámicos (DLM) y elaboran un método bayesiano para modelamiento dinámico y pronóstico. Este enfoque se basa en la práctica existente en el que modelos comunes pueden ser reformulados como casos especiales estáticos y estacionarios. El poder real, sin embargo, es que la nueva formulación proporciona muchas facilidades adicionales tantas como sean necesarios para trabajar con poca o ninguna información, introduciendo información subjetiva.

A finales de los setenta y principios de los ochenta se vio un gran desarrollo y aplicación de los métodos de modelamiento y pronóstico bayesiano. Notable entre estos, monitoreo e intervención (West 1986; West y Harrison, 1986; West, et. al. 1987, Harrison y Veerapen 1994; Pole, et. al. 1994), modelos de estructuras no normales y no lineales (Souza, 1981; Smith, 1979; Migon 1984; Migon y Harrison, 1985; West, Harrison y Migon 1985). Desde la evolución en la estadística computacional, a partir de 1990, las nuevas direcciones en la investigación, el desarrollo de modelos y su aplicación se han abierto, involucrando análisis mediante métodos de simulación (MCMC).

Dichos desarrollos computacionales dieron lugar a un uso más amplio de estas metodologías, facilitando la comunicación con los profesionales técnicamente menos orientados. Se cree que el modelamiento es un arte, y particularmente lo es el pronóstico Bayesiano.

3. MODELO LINEAL DINÁMICO

Los modelos dinámicos lineales son una amplia clase de modelos con parámetros variables en el

tiempo, útiles para el modelamiento de datos de series de tiempo y regresión. Fue introducido por (Harrison & Stevens 1976) y está muy bien documentado en el libro de (West & Harrison 1997) Los modelos dinámicos lineales son modelos paramétricos donde se describe la variación de los parámetros y la información de los datos disponibles probabilísticamente. Se caracterizan por un par de ecuaciones, denominada ecuación de observación y ecuación de evolución de parámetros. El DLM puede ser visto como una generalización de los modelos de regresión que permite cambios en los valores de los parámetros a través del tiempo. La ecuación observacional y la ecuación del sistema son respectivamente

$$\begin{aligned} y_t &= F_t' \theta_t + \epsilon_t, & \epsilon_t &\sim N(0, V_t) \\ \theta_t &= G_t \theta_{t-1} + w_t, & w_t &\sim N(0, W_t) \end{aligned} \quad (1)$$

donde y_t , es una serie de tiempo de observaciones, condicionalmente independientes dada la secuencia de parámetros θ_t , F_t es un vector de variables explicativas de tamaño $p \times 1$, θ_t un vector de parámetros $p \times 1$, G_t es una matriz $p \times p$ que describe la evolución de los parámetros y, por último, V_t y W_t son las varianzas de los errores asociados con la observación unidimensional y con el vector p-dimensional de los parámetros respectivamente. Esta clase incluye muchos de los modelos que se encuentran en la literatura estadística. Por ejemplo, si $G = I_p$, la matriz identidad de orden p y $\omega_t = 0 \forall t$, todos los modelos de regresión lineal pueden ser representados. Por otra parte, si F_t , V_t y W_t son constantes $\forall t$, entonces el modelo cubre los modelos de series de tiempo lineal, tales como los procesos ARIMA de Box y Jenkins (1976). En resumen, un modelo lineal dinámico está completamente especificado por el conjunto (F_t, G_t, V_t, W_t) .

Ejemplo 1 (Modelo polinomial de primer orden)

El modelo más simple de series temporales es el modelo polinomial de primer orden, el cual corresponde a una aproximación de primer orden de una serie de Taylor de una función suavizada del tiempo, llamada la tendencia de la serie. Este modelo está completamente definido por la cuádrupla $(1, 1, V_t, W_t)$. Reemplazando en el sistema (1) para

$$y_t = \theta_t + \epsilon_t, \epsilon_t \sim N(0, V_t), \theta_t = \theta_{t-1} + \omega_t, \omega_t \sim N(0, W_t)$$

donde θ_t es unidimensional y describe la tendencia del proceso. Aunque este modelo es muy simple, se puede aplicar en muchos sistemas de pronóstico a corto plazo, incluyendo un número de series de tiempo como en el control de inventarios o de

planificación de la producción. V_t y W_t también pueden evolucionar en el tiempo, ofreciendo un amplio margen para el modelamiento.

Ejemplo 2 (Regresión lineal simple dinámica)

Supongamos, que los pares de valores (x_t, y_t) son observados a través del tiempo y que se desea modelar la relación existente entre x_t y y_t . Suponiendo que el modelo lineal es una buena aproximación de la relación entre estos valores, un modelo de regresión lineal simple puede ajustarse. Puesto que la relación lineal es sólo una aproximación local para la función de dependencia real que involucra a x y y , un modelo con parámetros variables es más apropiado. Por ejemplo, la omisión de algunas variables, la no linealidad de la relación funcional de x y y o algunos cambios estructurales que ocurren en el proceso de investigación, pueden ser responsables de la inestabilidad de los parámetros. Estas situaciones se puede modelar como $y_t = F_t'\theta_t + \epsilon_t, \theta_t = \theta_{t-1} + \omega_t$ donde $F_t = (1, x)'$ y $\omega_t \sim N(0, W_t)$. Nótese que, en este caso, $G_t = I_2$.

Es de notar que la escogencia de F_t y G_t depende del modelo y de la naturaleza intrínseca de los datos que se están analizando. Para completar la especificación del modelo, las varianzas V_t, W_t deben establecerse. Esto último describe la velocidad de la evolución de los parámetros. En las aplicaciones, V_t es a menudo más grande que los elementos de W_t . Para facilitar la estimación del parámetro para su análisis, W_t es escalado por V_t y la varianza condicional de ω_t se vuelve $V_t W_t$. Por lo tanto, la matriz W_t se puede interpretar como una matriz de pesos relativos con respecto a la varianza observacional.

La matriz de varianzas de la evolución de los parámetros se puede asignar subjetivamente por el usuario del método, para poder hacer esto, el concepto de factor de descuento será de utilidad (véase Ameen y Harisson, 1985). Por lo tanto, las ecuaciones presentadas antes se pueden reescribir como:

$$\begin{aligned} (y_t | \theta_t) &\sim N(F_t'\theta_t, V_t) \\ (\theta_t | \theta_{t-1}) &\sim N(G_t\theta_{t-1}, V_t W_t) \end{aligned} \quad (2)$$

4. INFERENCIA EN DLM

La inferencia en DLM sigue los pasos usuales de la inferencia Bayesiana. Se exploran los aspectos secuenciales de la inferencia bayesiana combinando dos operaciones principales: la *evolución* para crear la *priori* y la *actualización* para incorporar observaciones nuevas en el tiempo t . Sea

$D_t = D_{t-1} \cap (y_t)$ denotan la información hasta el tiempo t , incluyendo los valores de x_t y $G_t, \forall t$, que se suponen conocidos, con D_0 representando la información a priori. Entonces para cada tiempo t la distribución a priori, la predictiva y la a posteriori están respectivamente dadas por:

$$\begin{aligned} p(\theta_t | D_{t-1}) &= \int p(\theta_t | \theta_{t-1}) p(\theta_{t-1} | D_{t-1}) d\theta_{t-1} \\ p(y_t | D_{t-1}) &= \int p(y_t | \theta_t) p(\theta_t | D_{t-1}) d\theta_t, \\ p(\theta_t | D_t) &\propto p(\theta_t | D_{t-1}) p(y_t | D_{t-1}) \end{aligned} \quad (3)$$

Donde la ultima ecuación del sistema (3) es obtenida vía teorema de Bayes. La constante de integración en la especificación anterior es “algunas veces” obtenida fácilmente. Este es justamente el caso cuando $(F, G, V, W)_t$ son todas conocidas y es asumida la normalidad. El algoritmo resultante en este caso es muy restringido y se conoce como filtro de Kalman. Usualmente las matrices anteriores dependen de algunos parámetros desconocidos denotados genéricamente por ψ .

Función de pronóstico y diseño DLM

En series de tiempo es muy útil encontrar la respuesta media de la función de pronóstico

$$E\left(E(y_{t+h} | \theta_{t+h}) | D_t\right) = F'E(\theta_{t+h} | D_t) = F'G^h m_t$$

donde h es el horizonte de predicción, $m_t = E(\theta_t | D_t)$. La estructura de esta función depende principalmente de los valores propios de la matriz G . Se puede ver en (West & Harrison 1997) cómo aplicar estos resultados teóricamente.

Ecuaciones de evolución y actualización

Las ecuaciones descritas antes, permiten la descripción conjunta de (y_t, θ_t) dadas las observaciones pasadas

D_{t-1} via $p(y_t, \theta_t | D_{t-1}) = p(y_t | \theta_t) p(\theta_t | D_{t-1})$ es to nos lleva a la distribución predictiva después de integrar sobre θ_t .

Una de las principales características de los modelos lineales dinámicos, es que en cada instante del tiempo, toda la información disponible es usada para describir la distribución a posteriori del vector de estados. El siguiente teorema muestra cómo evolucionar desde la distribución a posteriori en el tiempo $t-1$ hasta la distribución a posteriori en el tiempo t .

Teorema: Considere un modelo lineal dinámico normal con $V_t = V, \forall t$. Se denota la distribución a posteriori en $t-1$ por $(\theta_{t-1} | D_{t-1}, V) \sim N(m_{t-1}, VC_{t-1})$ y la distribución

marginal a posteriori de $\phi = V^{-1}$, como

$\phi|D_{t-1} \sim G(n_{t-1}/2, n_{t-1}s_{t-1}/2)$. Entonces,

1. Condicionalmente sobre V se tiene:

a) Evolución. La distribución a priori en el tiempo t será $\theta_t|V, D_{t-1} \sim N(a_t, VR_t)$, con $a_t = G_t m_{t-1}$ y $R_t = G_t C_{t-1} G_t' + W_t$.

b) La distribución predictiva un paso adelante será $y_t|V, D_{t+1} \sim N(f_t, VQ_t)$, con $f_t = F_t' a_t$ y $Q_t = F_t' R_t G_t + 1$

c) Actualización. La distribución a posteriori en el tiempo t será $\theta_t|V, D_t \sim N(m_t, VC_t)$, con $m_t = a_t + A_t e_t$ y $C_t = R_t - A_t A_t' Q_t$, donde $A_t = R_t F_t' / Q_t$ y $e_t = y_t - f_t$

2. La precisión ϕ es actualizada por la relación $\phi|D_t \sim G(n_t/2, n_t s_t/2)$, con $n_t = n_{t-1} + 1$ y $n_t s_t = n_{t-1} s_{t-1} + e_t^2 / Q_t$

3. Incondicionalmente sobre V se tiene:

a) $\theta_t|D_{t-1} \sim t_{n_{t-1}}(a_t, s_{t-1} R_t)$

b) $y_t|D_{t-1} \sim t_{n_{t-1}}(f_t, Q_t^*)$, con $Q_t^* = s_{t-1} Q_t$

c) $\theta_t|D_t \sim t_{n_{t-1}}(m_t, s_t C_t)$

5. MODELO LINEAL GENERALIZADO DINÁMICO

Una extensión de DLM's que permite observaciones de la familia exponencial fue presentado por (West et al. 1985) basándose en los modelos lineales generalizados de Nelder & Wedderburn (1972). La primera ecuación del sistema (1) se sustituye por

$$p(y_t|\eta_t) \propto \exp\left(\frac{y_t \eta_t - b(\eta_t)}{\phi_t}\right) \quad (4)$$

y además, una función de enlace adecuada se introduce, relacionando la media

$\mu_t = E(y_t|\eta_t) = b'(\eta_t)$ para los regresores F_t a través de $g(\mu_t) = \lambda_t = F_t' \theta_t$. Una distribución a priori conjugada para (4) está dada por

$p(\eta_t|D_{t-1}) \propto \exp\left(\frac{\gamma_t \eta_t - b(\eta_t)}{s_t}\right)$. Las integrales en

(3) no se pueden obtener en forma cerrada, por lo que la inferencia debe hacerse de forma aproximada. Un procedimiento que permite un

análisis secuencial de MLDG fue implementado en (West et al. 1985) utilizando una estimación lineal Bayesiana. Por ejemplo, si $y_t \sim \text{Gamma}(\varphi, \lambda_t)$

entonces $\eta_t = -\lambda_t / \varphi$, $b(\eta_t) = -\log(-\eta_t)$, con $\phi = 1/\varphi$. Para garantizar una respuesta media positiva μ_t se puede adoptar, en lugar del enlace canónico μ^{-1} el logaritmo de la función enlace: $\log(\varphi/\lambda_t) = \eta_t$.

La ecuación de evolución en (1) es sólo parcialmente especificada. Esto significa que las distribuciones de $\theta_{t-1}|D_{t-1}$ y ω_t son especificadas solamente por los momentos de primer y segundo orden, es decir $\theta_{t-1}|D_{t-1} \sim [m_{t-1}, C_{t-1}]$ y $\omega_t \sim [0, W_t]$.

Entonces, la distribución a priori de los parámetros de estado también es parcialmente especificada como $\theta_t|D_t \sim [a_t, R_t]$ tal como se indicó anteriormente. Entonces, la distribución a priori de $\lambda_t = g(\mu_t)$ es $\lambda_t|D_{t-1} \sim [f_t, q_t]$, donde $f_t = F_t' a_t$ y $q_t = F_t' R_t F_t$. Los parámetros (r_t, s_t) en la distribución a priori para η_t se pueden relacionar para f_t y q_t a través del $E(g(b'(\eta_t))|D_{t-1}) = f_t$ y $\text{var}(g(b'(\eta_t))|D_{t-1}) = q_t$.

Entonces la distribución a posteriori para η_t tiene la misma forma de esta (conjugada) distribución a priori con parámetros $(r_t/s_t + y_t/\phi_t, 1/s_t + 1/\phi_t)$.

La distribución a posteriori del predictor lineal es $\lambda_t|D_t \sim [f_t^*, q_t^*]$ donde, de nuevo

$$f_t^* = E(g(b'(\eta_t))|D_{t-1}) \quad y$$

$$q_t^* = \text{var}(g(b'(\eta_t))|D_{t-1}).$$

Más aun, para completar el análisis, la distribución a posteriori de los parámetros de estado se pueden obtener. La estimación lineal Bayesiana es usada para aproximar los primeros y segundos momentos de esta distribución llevándonos a:

$$\hat{E}(\theta_t|\eta_t, D_{t-1}) = a_t + R_t F_t (\eta_t - f_t) / q_t \quad y$$

$$\hat{\text{var}}(\theta_t|\eta_t, D_{t-1}) = R_t - R_t F_t F_t' R_t / q_t.$$

Los momentos de $\theta_t|D_t$ son calculados usando la ley de esperanza iterada dando $\theta_t|D_t \sim [m_t, C_t]$,

donde
$$m_t = a_t + R_t F_t (f_t' - f_t) / q_t \quad y$$

$$C_t = R_t - R_t F_t F_t' R_t (1 - q_t^* / q_t) / q_t .$$

6. MARKOV CHAIN MONTE CARLO

Los modelos dinámicos presentados anteriormente permiten inferir sólo cuando los F_t 's, G_t 's y W_t 's son completamente conocidos y una forma conjugada es impuesta sobre $V_t = V, \forall t$.. En general, estas cantidades u otras cantidades usadas en su definición son desconocidas y la inferencia acerca de ellas se basa en su distribución a posteriori. Esta distribución por lo general no se puede hallar analíticamente. Acá, el problema de hacer una inferencia sobre todos los parámetros del modelo en los DLM's es considerado. Se inicia considerando DLM's normales y luego se generalizan las ideas a DLM's no normales. Lo ideal en los problemas no resueltos analíticamente en inferencia Bayesiana se pueden resolver aproximadamente muestreando de la distribución a posteriori de referencia. Las distribuciones involucradas aquí son demasiado complicadas para extraer muestras directamente de ellas. La herramienta para resolver este problema es MCMC, una poderosa colección de técnicas de muestreo que ha revolucionado la inferencia Bayesiana para modelos complejos en las últimas décadas. Las limitaciones de espacio sólo permiten unas líneas descriptivas aquí sobre MCMC. Una descripción más extendida de MCMC se puede encontrar en (Gamerman 1997).

MCMC significa cadenas de Markov Monte Carlo y se ocupa de hacer muestreo de una distribución complicada (en este caso, la distribución a posteriori conjunta de todos los parámetros del modelo) donde el muestreo directo no está disponible. Es un procedimiento en dos pasos. En el primer paso, una cadena de Markov con un kernel de transición es construido tal que la distribución límite está dada por la distribución a posteriori conjunta. En el segundo paso, una trayectoria es la muestra de esta cadena. Para una repetición suficientemente grande, los valores son prácticamente muestreados de la distribución a posteriori. Los kernel de transición más utilizados se basan en el muestreador de Gibbs y Metropolis-Hastings. En el primer grupo, el kernel está compuesto por productos de las condicionales a posteriori de parámetros divididos en bloques. En el segundo grupo, el kernel se mantiene compuesto por productos de la condicionales propuestas, no necesariamente deriva de la a posteriori. Un término de corrección es impuesto para bloques no extraídos de la distribución a posteriori para asegurar la convergencia de la cadena para la a posteriori. Este

sistema también puede aplicarse si todos los parámetros se agrupan en un solo bloque. En este caso, la distribución propuesta sustituye a la distribución a posteriori conjunta. En teoría, cualquier partición de los parámetros en bloques asegura convergencia. Sin embargo, en la práctica, esta elección juega un papel importante proporcionando una solución computacionalmente factible.

6.1. DLM NORMAL

Ahora se considerará la solución a la inferencia en el DLM normal por muestreo. Sistemas alternativos de muestreo son presentados y comparados en términos teóricos y empíricos. La diferencia básica entre ambos radica en la especificación de los bloques de parámetros. Los esquemas son presentados en el orden de la dimensión cada vez mayor de los bloques, o en otras palabras, la disminución de número de bloques.

Los DLM normales considerados aquí son dados por el sistema (1) con F_t y G_t conocidas y las varianzas constantes del error observacional y del sistema, es decir, $V_t = V$ y $W_t = W$, para todo t . Esta restricción prevé un modelo más parsimonioso. La extensión para el caso general de varianzas diferentes no es difícil, pero puede dificultar significativamente la inferencia a menos que la información a priori esté disponible para las diferentes varianzas de los parámetros. Extensiones a los casos donde F_t y G_t sean cantidades desconocidas no son difíciles de aplicar a la metodología descrita a continuación.

Como se mostró antes de realizar inferencia exacta en DLM normal cuando la matriz de varianzas de evolución es conocida. También se mostró algunas alternativas simples a la inferencia con respecto a un W desconocido, es decir, la comparación de algunos valores de W y su especificación mediante el uso de factores de descuento, basados en el escalamiento sobre V .

Esquemas de muestreo componente a componente

Una muestra de la distribución a posteriori de $(\theta_1, \dots, \theta_T, V, W)$ es extraída después de observar la serie hasta el tiempo n . La posteriori conjunta es

$$\pi(\theta, V, W | D_T) \propto \prod_{i=1}^T p(y_i | \theta_i, V) \prod_{i=2}^T p(\theta_i | \theta_{i-1}, W) p(\theta_1) p(V, W)$$

donde, $\theta = (\theta_1, \dots, \theta_T)$. La especificación del modelo se puede tomar como base para el bloqueo de parámetros. Por lo tanto, una elección inicial de bloqueo es $\theta_1, \dots, \theta_T, V$ y W . La condicional

completa de los θ_t 's, denotado por $\pi(\theta_t | \theta_{-t} V, W, D_T)$, son distribuciones normales que se obtienen fácilmente. Si la a priori conjunta para (V, W) es el producto de una Gamma inversa para V y una Wishart inversa para W , entonces la condicional también será así. Así, el muestreo de las condicionales es fácilmente disponible.

Estas distribuciones condicionales completan un ciclo del sistema de muestreador de Gibbs. Todos ellos son fáciles de muestrear de MCMC basado en una inferencia bayesiana.

La distribución a priori de θ es altamente dependiente del valor de W . Para valores pequeños de W , grandes correlaciones son inducidas y conservadas a través de la distribución a posteriori. Desafortunadamente, este suele ser el caso observado en la práctica con los parámetros del sistema que experimentan pequeñas perturbaciones a través del tiempo. Esto significa que el muestreo de los componentes θ por separado puede implicar un kernel que va a ser muy lento y sólo puede alcanzar el equilibrio después de un número innecesariamente grande de iteraciones.

Algoritmo I MCMC para DLM

1. Inicialización: Un conjunto de valores iniciales $(\theta^{(0)}, V^{(0)}, W^{(0)})$ y el contador de iteraciones $j = 1$;

2. Muestreando $\theta : \theta^{(j)}$ es muestreado componente a componente como sigue

a. Hacer $t = 1$;

b. Muestrear $\theta_t^{(j)}$ de $\pi(\theta_t | \theta_1^{(j)}, \dots, \theta_{t-1}^{(j)}, \theta_{t+1}^{(j-1)}, \dots, \theta_T^{(j-1)}, V^{(j-1)}, W^{(j-1)}, D_T)$;

c. Hacer $t \rightarrow t + 1$ y retorne a 2b, si $t < T$;

3. Muestreando V y W : $V^{(j)}$ y $W^{(j)}$, son muestreadas sucesivamente de las respectivas funciones condicionales $\pi(V | \theta^{(j)}, W^{(j-1)}, D_T)$ y

$\pi(W | \theta^{(j)}, V^{(j)}, D_T)$;

4. Actualización: Sea $j \rightarrow j + 1$ y retornar a 2 hasta obtener convergencia.

Modelo lineal generalizado dinámico

Las extensiones anteriores implican una solución MCMC que se basa en pasos extras en los algoritmos de muestreo para DLM's normal. El problema principal del método anterior fue la carga computacional de muestrear a θ eficientemente,

pero no siempre se pueden obtener directamente las funciones condicionales.

En el caso de MLDG, la función condicional de θ no tiene siempre una forma conocida y no puede ser muestreada directamente. Ni siquiera la función condicional de θ_t es conocida. Por lo tanto, el muestreo de ellos se descarta y nuevas propuestas deben ser empleadas. Las propuestas adecuadas para θ con términos de corrección eficaces son extremadamente difíciles de encontrar debido a la dimensión de θ . Propuestas adecuadas para θ_t son más fáciles de encontrar y de construir. Sin embargo, conducen a una lenta convergencia de la cadena por las mismas razones presentadas para los modelos normales.

Shephard y Pitt (1997) y Knorr-Held (1999) sugieren el uso de bloques aleatorios con algunos de los componentes de θ . Estas propuestas se basan en aproximaciones normales a la verosimilitud y caminatas aleatorias móviles, respectivamente. Gamerman (1998) sugirió el uso de bloques de tiempos específicos pero utilizando una reparametrización en términos de los errores w 's para evitar la fuerte correlación entre θ_t 's. Lo anterior puede ser reescrito como $w_t = \theta_t - G_t \theta_{t-1}, t = 2, \dots, T$, en términos de los parámetros del sistema θ_t , con $w_1 = \theta_1$. Los parámetros del sistema θ_t son fácilmente despejados por la relación inversa

$$\theta_t = \sum_{i=1}^t \left(\prod_{k=1}^{t-i} G_{t-k+1} \right) w_i \quad (5)$$

para $t = 2, \dots, n$ y $\theta_1 = w_1$. Nótese que

$\pi(w | W, D_T) = \pi(\theta(w) | W, D_T)$ para

$w = (w_1, \dots, w_T)$ puesto que el Jacobiano de la transformación es uno. El algoritmo de muestreo está dado por:

ALGORITMO II MCMC PARA MLDG BASADO EN REPARAMETRIZACIÓN DE ESTADOS

1. Inicialización: Un conjunto de valores iniciales $(\theta^{(0)}, W^{(0)})$ y el contador de iteraciones $j = 1$;

2. Muestreando $\theta : \theta^{(j)}$ es muestreado componente a componente como sigue

a) Hacer $t = 1$;

b) Muestrear w_t^* de $q_t(w_t | w_t^{(j-1)}, w_{-t}^{(j-1)}, W_t^{(j-1)}, D_T)$, con

$$w_{-t}^{(j-1)} = \left(w_1^{(j)}, \dots, w_{t-1}^{(j)}, w_{t+1}^{(j-1)}, \dots, w_T^{(j-1)} \right);$$

c) Sea $w_t^{(j)} = w_t^*$ con probabilidad α_t y

$w_t^{(j)} = w_t^{(j-1)}$ con probabilidad $1 - \alpha_t$ donde

$\alpha_t = \min \{1, A\}$ donde A es

$$\frac{\pi\left(w_t^* \mid w_{-t}^{(j)}, W^{(j-1)}, D_t\right) q_t\left(w_t^{(j-1)} \mid w_t^*, w_{-t}^{(j-1)}, W^{(j-1)}, D_t\right)}{\pi\left(w_t^{j-1} \mid w_{-t}^{(j)}, W^{(j-1)}, D_t\right) q_t\left(w_t^* \mid w_t^{(j-1)}, w_{-t}^{(j-1)}, W^{(j-1)}, D_t\right)}$$

d) Reconstruir $\theta_t^{(j)} = G_t \theta_{t-1}^{(j)} + w_t^{(j)}$;

e) Haga $t \rightarrow t + 1$ y regrese a 2b, si $t < T$;

3. Muestreando $W : W^{(j)}$ es muestreado desde la función condicional $\pi\left(W \mid \theta^{(j)}, D_T\right)$;

4. Actualización haga $j \rightarrow j + 1$ y regrese a 2 hasta encontrar convergencia.

La densidad condicional de w_t requerido en la expresión de α_t es obtenida a partir de la densidad

conjunta $\pi(w \mid W, D_T)$ tras la eliminación de

términos que no involucren a w_t . Tal y como está expresado no es necesario conocer la constante. Las densidades propuesta q_t utilizado por (Gamerman

1998) para MLDG. las variables están dadas por la función (normal) condicional de w_t en un DLM con el mismo sistema sistema de ecuaciones y modificando la ecuación observacional

$\tilde{y}_t \sim N(F_t' \theta_t, V_t)$, con

$$\tilde{y}_t = g'(\mu_t) + g'(\mu_t)(y_t - \mu_t) \quad y$$

$$\tilde{V}_t = \left(g'(\mu_t) \right)^2 \text{var}(y_t \mid \theta_t).$$

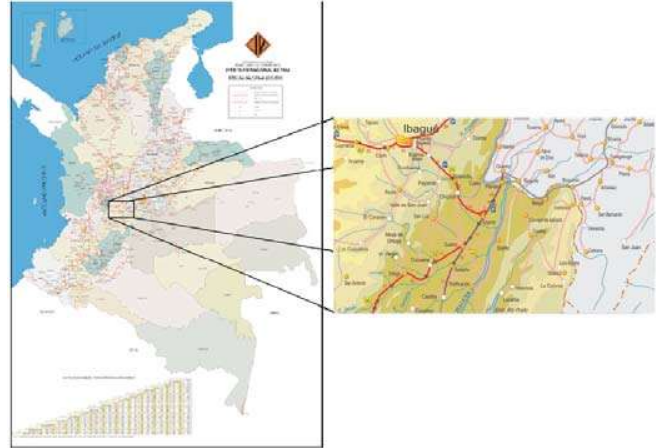
Estas variables son una extensión de las variables de trabajo utilizados en la estimación de máxima verosimilitud de los modelos lineales generalizados. Ferreira & Gamerman (2000) detallan e ilustran el uso de este esquema de muestreo.

Las iteraciones son más costosas aquí por la reconstrucción de las ecuaciones anteriores, pero presentan un ahorro en el costo computacional.

Las simulaciones se van a llevar a cabo usando WinBUGS, que es un software estadístico para análisis bayesiano usando los métodos MCMC. Está basado en el proyecto BUGS (Bayesian Inference Using Gibbs Sampling), iniciado en 1989. Es desarrollado por un grupo de investigadores del Unidad de Bioestadística del Medical

Research Council en Cambridge y el Imperial College School of Medicine de Londres.

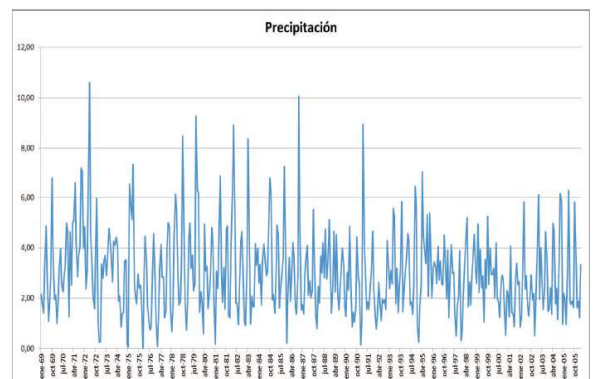
FIGURA 1
Estudio del caudal a través de modelos lineales generalizados dinámicos
Ubicación de la estación Cabrera



7. APLICACIÓN

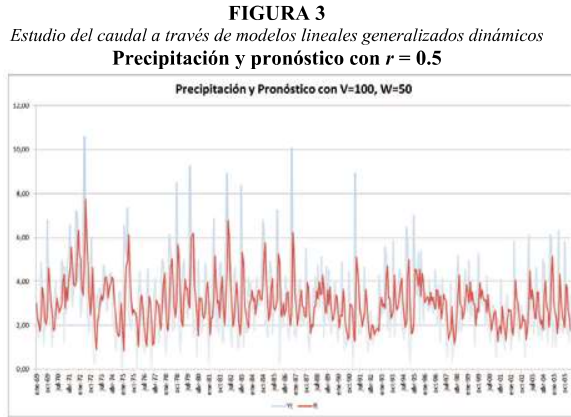
Los datos que se van a utilizar en el estudio corresponden a una serie mensual de precipitación (ver Figura 2) desde Enero de 1969 hasta Marzo de 2006 de la estación pluviométrica 2119009-Cabrera del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) ubicada a los 3°59' de latitud Norte, 74°29' de longitud Oeste con una elevación de 1900 msnm en el municipio de Cabrera Cundinamarca para la corriente del río Sumapaz. (Ver Figura 1.)

FIGURA 2
Estudio del caudal a través de modelos lineales generalizados dinámicos
Precipitación estación Cabrera



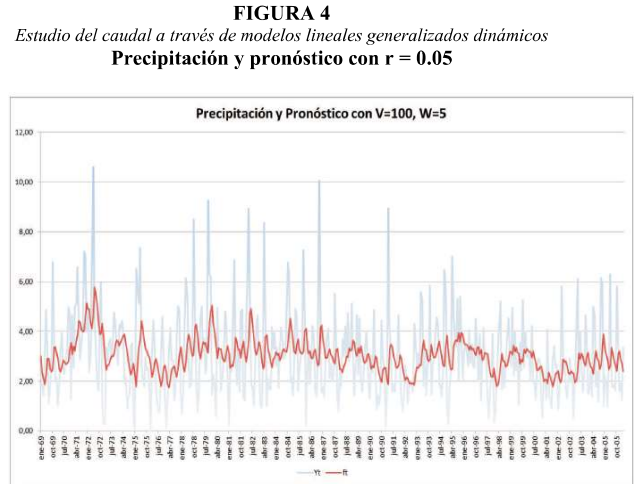
Al implementar en los datos un modelo DLM constante (1, 1, 100, 50), $r = 0.5$ es una relación señal-ruido fuerte con información inicial

$(\mu | D_0) \sim N(3, 400)$, la Figura 3 muestra la serie de tiempo de las observaciones y el pronóstico a un paso.



Por otro lado, al implementar un nuevo modelo DLM constante $(1, 1, 100, 5)$, $r = 0.05$ es una relación señal-ruido débil con información inicial

$(\mu | D_0) \sim N(3, 400)$, la Figura 4 muestra la serie de tiempo de las observaciones y el pronóstico a un paso.



REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS

- [1]. **GAMERMAN, D.** (1997), *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference.*, Chapman & Hall, London.
- [2]. **GAMERMAN, D.** (1998), 'Markov chains Monte Carlo for dynamic generalized linear models', *Biometrika* 85, 215–227.
- [3]. **HARRISON, P. J. & STEVENS, C. V.** (1976), 'Bayesian Forecasting.', *Journal of the Royal Statistical Society* (38), 205–247.
- [4]. **WEST, M. & HARRISON, J.** (1997), *Bayesian Forecastin and Dinamic Models*, Springer-Verlag, New York.
- [5]. **WEST, P., MIGON, H. & HARRISON, J.** (1985), 'Dynamic Generalized Linear Models and Bayesian Forecasting', *Journal of the American Statistical Association* 80(389), 73–83.