

Seminario en la URJC

Móstoles, 3 de abril de 2006

“A Bayesian approach for the rainfall-runoff problem: The case of Rio Grande Basin”

Romy R. Ravines
romy@dme.ufrj.br
dme.ufrj.br/romy

Alexandra M. Schmidt
alex@im.ufrj.br
dme.ufrj.br/alex

Helio S. Migon
migon@im.ufrj.br
dme.ufrj.br/migon



Universidade Federal do Rio de Janeiro (UFRJ)
Instituto de Matemática (IM)
Departamento de Métodos Estatísticos (DME)

Motivación: la relación *rainfall-runoff*

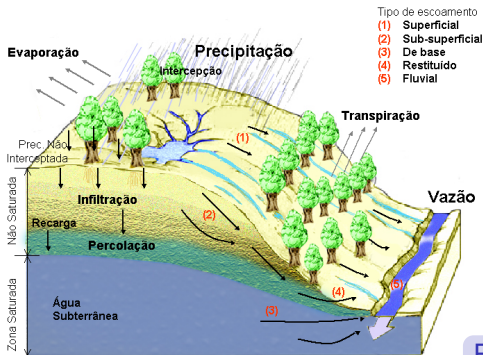


Figura: Procesos físicos en la generación del runoff

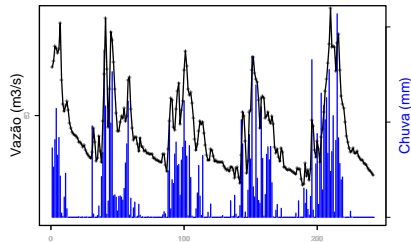
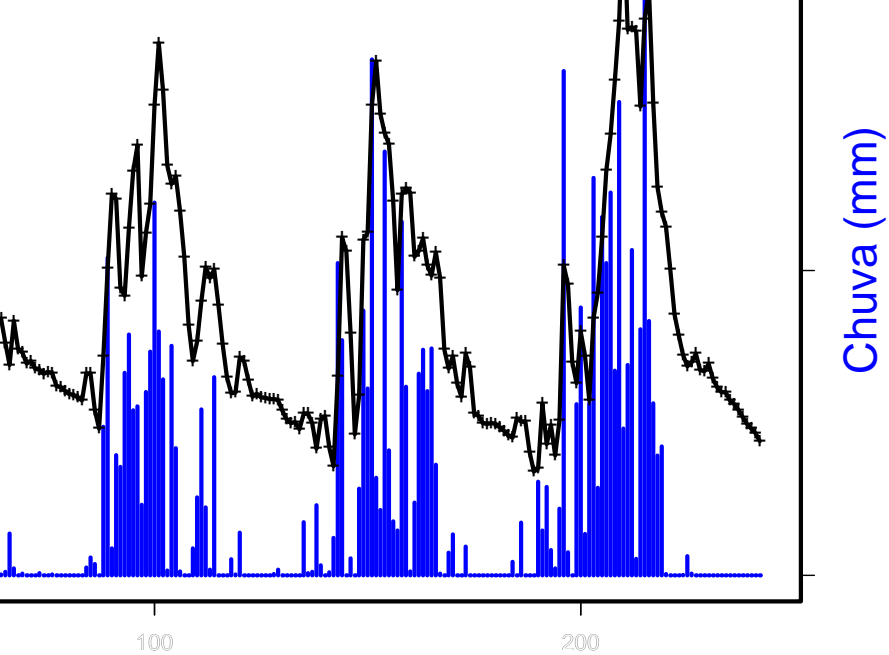


Figura: Datos reales

Palabras-Clave

- Modelar Conjuntamente
- Familia Exponencial
- Funciones de Transferencia



A Bayesian approach for the rainfall-runoff problem: The case of Rio Grande Basin

La cuenca del Rio Grande (BA): localización



La cuenca del Rio Grande (BA): datos

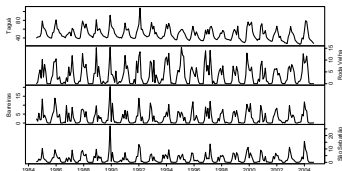
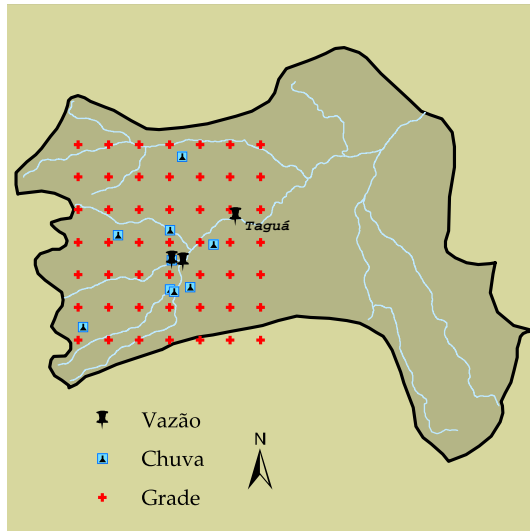


Figura: Datos mensuales: agosto de 1984 a setiembre de 2004



Efecto de la precipitación: Función de Transferencia

Siguiendo lo establecido en Migon & Monteiro (1997), la relación *rainfall-runoff* puede ser representada por una función de transferencia. En particular, E_t puede seguir una de las siguientes alternativas:

Dos funciones de transferencia:

$$E_t = \rho_t E_{t-1} + \gamma_t X_t$$

$$E_t = \rho_t E_{t-1} + [1 - \exp(-\kappa_t X_t)] [\phi_t - \alpha_t - \rho_t E_{t-1}]$$

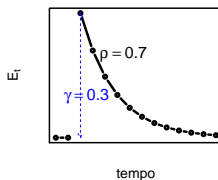
Casos particulares de γ_t :

Constante: $\gamma_t = \gamma$

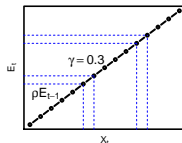
Temporal: $\gamma_t = \gamma_{t-1} + \delta_t$

Aleatorio: $\gamma_t = \gamma + \delta_t$

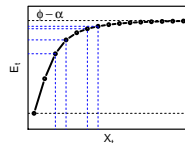
Jerárquico: $\gamma_t = \gamma + \delta_t$; $\gamma \sim N(a, b)$



(a) Dec. Exponencial



(b) Ret. Proporcionales



(c) Ret. Decrecientes

Figura: Hipótesis sobre E_t

Efecto de la precipitación: Función de Transferencia

Siguiendo lo establecido en Migon & Monteiro (1997), la relación *rainfall-runoff* puede ser representada por una función de transferencia. En particular, E_t puede seguir una de las siguientes alternativas:

Dos funciones de transferencia:

$$E_t = \rho_t E_{t-1} + \gamma_t X_t$$

$$E_t = \rho_t E_{t-1} + [1 - \exp(-\kappa_t X_t)] [\phi_t - \alpha_t - \rho_t E_{t-1}]$$

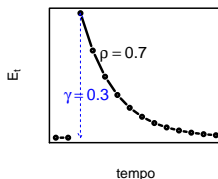
Casos particulares de γ_t :

Constante: $\gamma_t = \gamma$

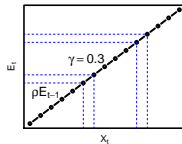
Temporal: $\gamma_t = \gamma_{t-1} + \delta_t$

Aleatorio: $\gamma_t = \gamma + \delta_t$

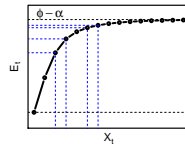
Jerárquico: $\gamma_t = \gamma + \delta_t; \quad \gamma \sim N(a, b)$



(a) Dec. Exponencial



(b) Ret. Proporcional



(c) Ret. Decrecientes

Figura: Hipótesis sobre E_t

Modelo propuesto: caso particular

Runoff

- ⇒ Distribución: Log-Normal o Gama
- ⇒ Serie no estacionaria
- ⇒ Efecto X_t : FT 1º orden.

Rainfall

- ⇒ Distribución: Normal Truncada (Sansó & Guenni, 2000)
- ⇒ Estacionalidad: 2 armónicos
- ⇒ $F_i = (1, \text{long}_i, 1, 0, 1, 0)'$
- ⇒ $G = \begin{pmatrix} G_1 & \mathbf{0} \\ \mathbf{0} & G_2 \end{pmatrix}$.

$$Y_t \sim p(\mu_t, \phi) \quad t = 1, \dots, T = 221 \quad (1a)$$

$$f(\mu_t) = \alpha_t + E_t \quad (1b)$$

$$\alpha_t = \alpha_{t-1} + w_{\alpha,t} \quad w_{\alpha,t} \sim N(0, \sigma_\alpha^2) \quad (1c)$$

$$E_t = \rho E_{t-1} + \gamma X_t + w_{E,t} \quad w_{E,t} \sim N(0, \sigma_E^2) \quad (1d)$$

$$X_t = |B|^{-1} \int_B X_t(s) ds \quad s = 1, \dots, S = 49 \quad (1e)$$

$$X_{it} = \begin{cases} w_{it}^\beta & \text{si } w_{it} > 0 \\ 0 & \text{si } w_{it} \leq 0 \end{cases} \quad i = 1, \dots, N = 9 \quad (1f)$$

$$w_t = z_t + v_t \quad v_t \sim N_N(\mathbf{0}, \tau^2 \mathbf{I}) \quad (1g)$$

$$z_t = F' \theta_t + \epsilon_t \quad \epsilon_t \sim N_N(\mathbf{0}, \sigma^2 V_t) \quad (1h)$$

$$\theta_t = G \theta_{t-1} + \epsilon_t \quad \epsilon_t \sim N_k(\mathbf{0}, W_t) \quad (1i)$$

En (1f)–(1i), $\sigma^2 > 0$, $V_t \in \mathbb{R}^{N \times N}$, $\theta \in \mathbb{R}^k$, $G \in \mathbb{R}^{k \times k}$, $F' \in \mathbb{R}^{N \times k}$. w_{it} es una variable latente Gaussiana, $\tau^2 \mathbf{I}$ es el efecto pepita, $\theta_t = (\theta_{t1}, \theta_{t2})'$, where θ_{t1} es el sub-vector de la tendencia espacial y θ_{t2} de la estacionalidad. V_t captura la correlación espacial, con $V_{ij} = \exp(-\lambda d_{ij})$. En (1b), $f(\cdot) = \log$ o identidad, dependiendo de $p(\mu_t, \phi)$.

Inferencia

- Aprovechando la factorización de la verosimilitud en $p(Y_t|X_t)p(X_t)$, la inferencia puede ser realizada en dos etapas: Primero se ajusta un modelo para X_{it} , (1e)–(1i), por ejemplo, con diferentes especificaciones de la tendencia polinomial, después se ajustan los casos particulares del modelo para Y_t , (1a)–(1d).
- Métodos MCMC: Metrópolis-Hastings, Slice Sampling.
- **“An Efficient Sampling Scheme for Generalized Dynamic Models”**
- Comparación de modelos: DIC, EPD, ECM, EAM.
- Algoritmos escritos en

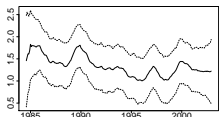
[▶ CUBS](#)0x: <http://www.doornik.com/>

Modelo Espacio-Temporal de la Precipitación: resultados (1)

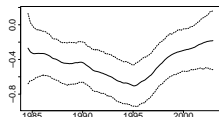
El modelo seleccionado para la precipitación tiene un intercepto y un efecto lineal de la longitud en la tendencia espacial. El ajuste de otros modelos mostró que la latitud no tiene un efecto significativo en esta región.

Cuadro: Estadísticas a posteriori dos parámetros estáticos

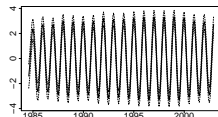
	media	ds	2.5%	25%	50%	75%	97.5%	\hat{R}	n.eff
β	1.339	0.014	1.311	1.329	1.339	1.348	1.367	1.001	5000
λ	0.065	0.013	0.044	0.056	0.063	0.073	0.094	1.001	5000
ρ^2	0.632	0.035	0.566	0.607	0.630	0.655	0.702	1.001	5000
σ^2	1.073	0.046	0.988	1.041	1.071	1.103	1.166	1.001	5000



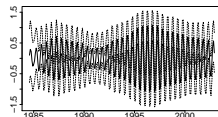
(a) Intercepto



(b) Longitude



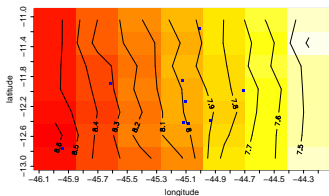
(c) 1º armónico



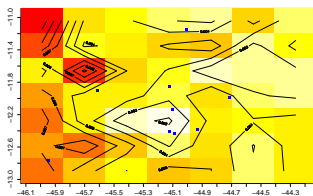
(d) 2º armónico

Figura: Trayectoria estimada para os parámetros dinámicos

Modelo Espacio-Temporal de la Precipitación: resultados (2)

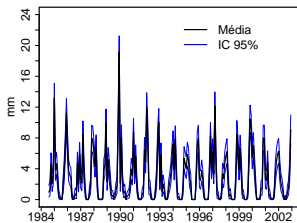


(a) Diciembre, 2000

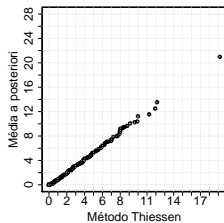


(b) Junio, 2002

Figura: Media a posteriori de la precipitación en dos meses. Los puntos marcan la localización de las estaciones.



(a) Intervalo de 95 % de credibilidad



(b) Comparación con Thiessen

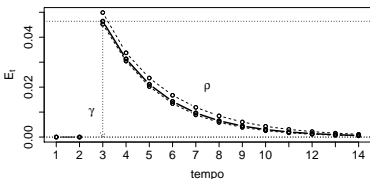
Figura: Precipitación en la cuenca del Rio Grande

Modelo de Función de Transferencia para *Runoff*: resultados

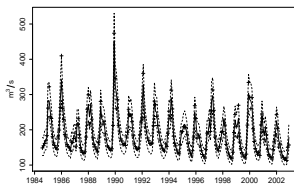
El modelo seleccionado para la escorrentía considera una respuesta Log-Normal, un nivel estático y una función de transferencia estocástica.

Cuadro: Estadísticas a posteriori dos parámetros estáticos

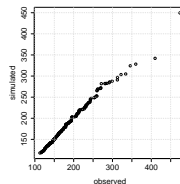
	media	ds	2.5 %	25 %	50 %	75 %	97.5 %	\hat{R}	n.eff
α	4.771	0.029	4.708	4.754	4.772	4.791	4.823	1.001	2500
ρ	0.674	0.021	0.631	0.660	0.675	0.688	0.712	1.001	2500
γ	0.047	0.002	0.043	0.045	0.047	0.048	0.050	1.001	2500
σ_{γ}^2	0.005	0.001	0.003	0.004	0.005	0.005	0.007	1.001	2500
σ_E^2	0.004	0.001	0.002	0.003	0.003	0.004	0.005	1.001	2500



(a) Función Respuesta-Impulso



(b) Media e 95 % IC para runoff



(c) Q-Q plot

Figura: Escorrentía: valores ajustados, con (1a)-(1d). Media a posteriori (línea sólida) y límites del IC de 95 % (líneas con puntos). + son los datos observados.

Previsión

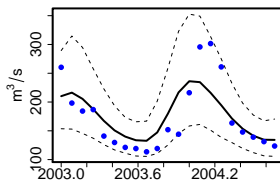
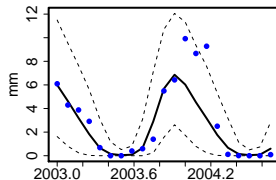
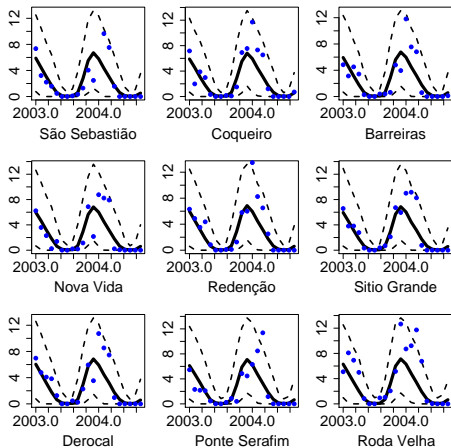
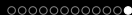


Figura: Previsión: media (línea sólida) e intervalo de 95% de credibilidad (líneas punteadas). Los puntos en azul: en (a) y (c) corresponden a los datos observados, en (b) corresponden a los valores obtenidos con el método Thiessen



La contribución es:

- **Modelo conjunto:** combinar un modelo de función de transferencia para *runoff*, condicionado en la precipitación de la cuenca, con un modelo espacio-temporal simultáneo para la precipitación medida en varias localizaciones dentro del área de drenaje de la estación de *runoff* en consideración.

► Extensiones

Consideraciones Finales:

- Los modelos utilizados tienen una **interpretación física** clara y constituyen una representación parsimoniosa de todos los procesos envueltos en la relación *rainfall-runoff*: **elicitación de prioris**.
- Una extensión: utilización de **modelos jerárquicos** para trabajar con varias cuencas simultáneamente.
- Uso de los resultados de esta forma de modelar en el **manejo de embalses**: Teoría de la Decisión.

La contribución es:

- **Modelo conjunto:** combinar un modelo de función de transferencia para *runoff*, condicionado en la precipitación de la cuenca, con un modelo espacio-temporal simultaneo para la precipitación medida en varias localizaciones dentro del área de drenaje de la estación de *runoff* en consideración.

► Extensiones

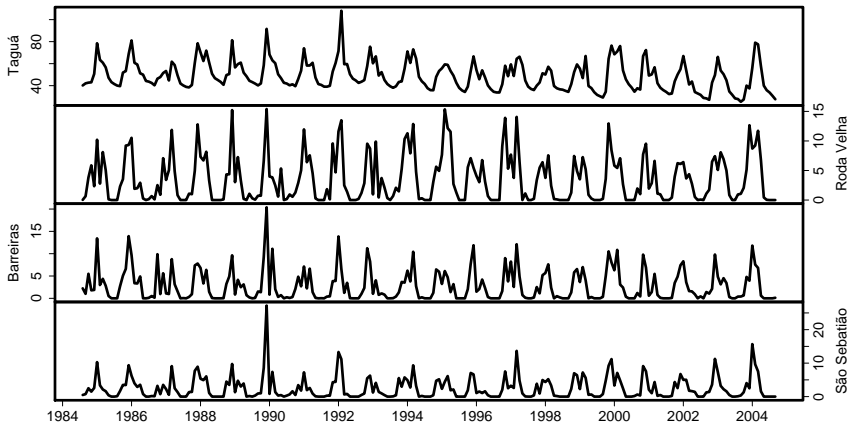
Consideraciones Finales:

- Los modelos utilizados tienen una **interpretación física** clara y constituyen una representación parsimoniosa de todos os procesos envueltos en la relación *rainfall-runoff*: **elicitación de prioris**.
- Una extensión: utilización de **modelos jerárquicos** para trabajar con varias cuencas simultáneamente.
- Uso de los resultados de esta forma de modelar en el **manejo de embalses**: Teoría de la Decisión.

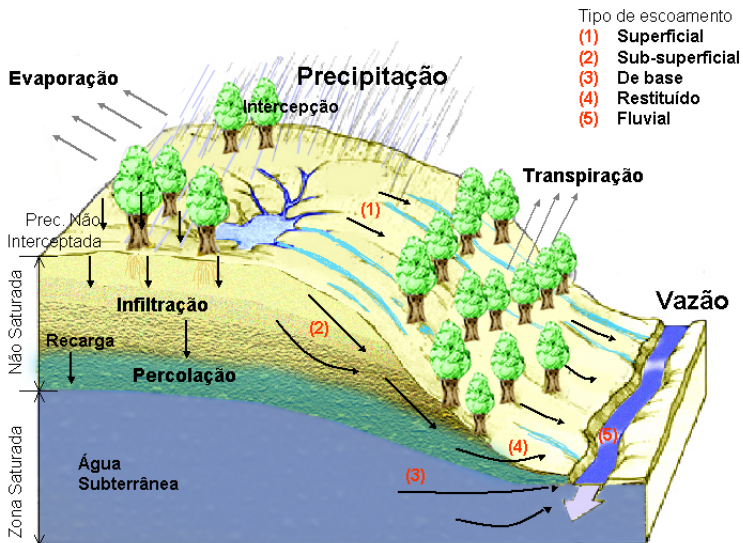
Algunas Referencias

- Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, *85*(1), 215–227.
- Geweke, J., & Tanizaki, H. (2001). Bayesian estimation of state space models using Metropolis-Hastings algorithm within Gibbs sampling. *Computacional Statistics & Data Analysis*, *37*, 151-170.
- Kitagawa, G. (1987). Non-gaussian state-space modeling of non-stationary time series. *Journal of the American Statistical Association*, *82*(400), 1032–1041.
- Migon, H., & Monteiro, A. B. (1997). Rain-fall modelling: An application of Bayesian forecasting. *Stochastic Hydrology and Hydraulics*, *11*, 115–127.
- Sansó, B., & Guenni, L. (2000). A non-stationary multi-site model for rainfall. *Journal of the American Statistical Association*, *95*, 1089–1100.
- West, M., Harrison, J., & Migon, H. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, *80*(389), 73–83.

Cuenca del Rio Grande (BA): series temporales



Procesos físicos envueltos en la generación de la escorrentía



Extensión 1: runoff en varios puestos

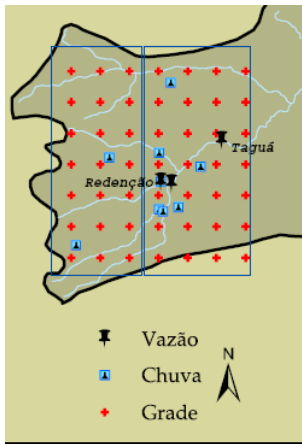


Figura: Estaciones y división del grid de previsión

El modelo propuesto puede ser extendido para modelar las series de runoff de dos estaciones de la misma cuenca:

$$Y_{Ta,t} \sim p(\mu_{Ta,t}, \phi_{Ta}) \quad t = 1, \dots, T = 221$$

$$\mu_{Ta,t} = \alpha_{Ta} + \beta \log(Y_{Re,t}) + E_{Ta,t}$$

$$E_{Ta,t} = \rho_{Ta} E_{Ta,t-1} + \gamma_{Ta} X_{Ta,t} + w_{Ta,t} \quad w_{Ta,t} \sim N(0, \sigma_{E_{Ta}}^2)$$

$$Y_{Re,t} \sim p(\mu_{Re,t}, \phi_{Re}) \quad t = 1, \dots, T = 221$$

$$\mu_{Re,t} = \alpha_{Re} + E_{Re,t}$$

$$E_{Re,t} = \rho_{Re} E_{Re,t-1} + \gamma_{Re} X_{Re,t} + w_{Re,t} \quad w_{Re,t} \sim N(0, \sigma_{E_{Re}}^2)$$

$$X_{Ta,t} = |B_{Ta}|^{-1} \int_{B_{Ta}} X_t(s) ds \quad s = 1, \dots, S_1 = 21$$

$$X_{Re,t} = |B_{Re}|^{-1} \int_{B_{Re}} X_t(s) ds \quad s = 1, \dots, S_2 = 28$$

$$X_t(s) = \beta_t f(s) + Z_t(s) + \epsilon_t(s), \quad s = 1, \dots, S = 9$$

$$Z(s) \sim GP(\mathbf{0}, \sigma^2 \varrho(\|s_1, s_2\|, \lambda))$$

donde $Y_{j,t}$ y $X_{j,t}$ denotan la escorrentía y la precipitación en el tiempo t en la localización $j = Ta, Re$. $Ta = \text{Taguá}$, $Re = \text{Redenção}$. B_j es el área de drenaje relativa a la estación j , $B_{Ta} \cup B_{Re} = B$, es el área de toda a cuenca.

Extensión 2: modelo jerárquico

Una extensión natural del modelo de función de transferencia usado aquí es un **modelo jerárquico** como:

$$Y_{t,k} \sim p(\mu_{t,k}, \sigma_Y^2) \quad t = 1, \dots, T; \quad (2a)$$

$$k = 1, \dots, K,$$

$$\mu_{t,k} = \alpha_k + E_{t,k} \quad (2b)$$

$$E_{t,k} = \rho E_{t-1,k} + \gamma_k X_{t,k} \quad (2c)$$

$$\alpha_k = \alpha + v_{k1}, \quad v_{k1} \sim N(0, \sigma_\alpha^2) \quad (2d)$$

$$\gamma_k = \gamma + v_{k2}, \quad v_{k2} \sim N(0, \sigma_\gamma^2) \quad (2e)$$

donde $Y_{t,k}$ es la escorrentia $X_{t,k}$ es la precipitación en el tiempo t en la cuenca k . Este modelo asume que el nivel básico e el efecto de la precipitación de las K cuencas varía en torno de un nivel común, denotados por α y γ respectivamente. Naturalmente, la memoria del efecto de la precipitación también puede tener una estructura jerárquica. La *exchangeability* de las cuencas es una hipótesis de este modelo.

Como ejercicio, ajustamos el modelo (2) a los datos de Rio Grande.

Cuadro: Posterior mean and sd

	media	dp
α	4.1201	0.8614
σ_α^2	2.6817	12.9231
α_1	3.5084	0.0150
α_2	4.0589	0.0152
α_3	4.8131	0.0160
γ	0.0438	0.0399
σ_γ^2	0.0052	0.0266
γ_1	0.0421	0.0014
γ_2	0.0420	0.0014
γ_3	0.0462	0.0014
ρ	0.6158	0.0117
σ_Y^2	0.0131	0.0007

Otros resultados: escorrentía en varios puestos

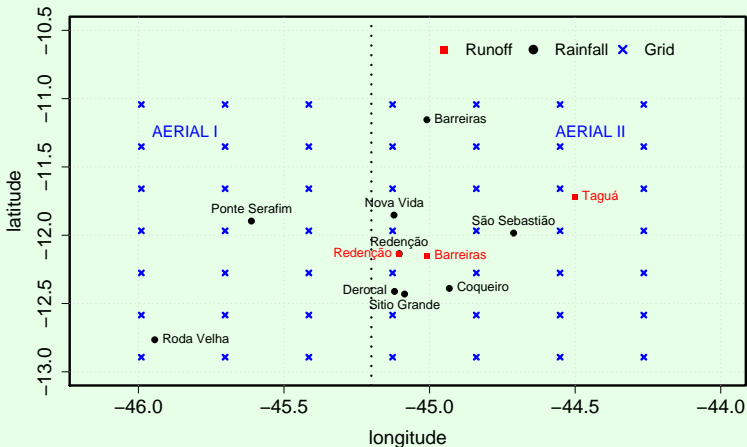


Figura: Localización de las estaciones y división del grid de previsión.

An Efficient Sampling Scheme for Generalized Dynamic Models

Modelos dinámicos generalizados

- Modelos dinámicos generalizados:

$$y_t | \mu_t \sim p(\mu_t, \boldsymbol{\phi}), \quad t = 1, \dots, T. \quad (3a)$$

$$g(\mu_t) = \mathbf{F}_t(\boldsymbol{\psi}_1)' \boldsymbol{\theta}_t \quad (3b)$$

$$\boldsymbol{\theta}_t = \mathbf{G}_t(\boldsymbol{\psi}_2) \boldsymbol{\theta}_{t-1} + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim N(\mathbf{0}, \mathbf{W}_t) \quad (3c)$$

donde $p(\mu_t, \boldsymbol{\phi})$ es una distribución de la familia exponencial, μ_t es el valor esperado de y_t , $\boldsymbol{\phi}$ representa otros parámetros de $p(\cdot)$, y $\boldsymbol{\psi}_1$ y $\boldsymbol{\psi}_2$ denotan parámetros envueltos en la definición de \mathbf{F}_t y \mathbf{G}_t , respectivamente.

- Los vectores $\boldsymbol{\theta}_t$ son conocidos como parámetros de estado y están relacionados a través del tiempo via (3c), la ecuación del sistema.
- Muestrear $\boldsymbol{\theta}_t$ puede ser complicado.
- Alternativa de muestreo: paso Metropolis-Hastings.
- Propuestas: Gamerman (1998), Geweke & Tanizaki (2001), etc.

Modelos dinámicos lineales generalizados

- West, Harrison, & Migon (1985) presentaron los **modelos dinámicos lineales generalizados**: la distribución de la respuesta pertenece a la familia exponencial con parámetro natural η_t . η_t tiene una priori conjugada, $CP(r_t, s_t)$. La ecuación del sistema es similar a (3c), mas la distribución de sus errores está parcialmente especificada.

$$y_t | \eta_t, \phi \sim \exp[\phi \{y_t \eta_t - a(\eta_t)\}] b(y_t, \phi), \quad t = 1, \dots, T. \quad (4a)$$

$$\eta_t | D_{t-1} \sim CP(r_t, s_t) \quad (4b)$$

$$g(\eta_t) = F_t' \theta_t \quad (4c)$$

$$\theta_t = G_t \theta_{t-1} + w_t, \quad w_t \sim [0, W_t] \quad (4d)$$

$$\theta_0 | D_0 \sim [m_0, C_0]$$

donde D_t denota la información hasta el tiempo t , m_t e C_t denotan, respectivamente, el primer y segundo momento del vector de estados θ_t , dado D_t .

- En (4), **F_t y G_t son conocidas**, y, dado η_t , y_t y θ_t son condicionalmente independientes, mientras que en (3) esta estructura es diferente.
- West et al. (1985) propusieron un sistema recursivo que explora la conjugación del modelo para aproximar las distribuciones de θ_t secuencialmente. Este sistema es conocido como **Conjugate Updating**.

Análisis Secuencial

- Sea $D_t = \{y_1, \dots, y_t\}$ - la información total obtenida hasta t .
- En los **modelos lineales normales**:

$$\dots (\theta_{t-1}|D_{t-1}) \xrightarrow{\text{Evolución}} (\theta_t|D_{t-1}) \xrightarrow{\text{Actualización}} (\theta_t|D_t) \dots$$

- En los **modelos lineales dinámicos generalizados**:

$$\begin{array}{ccccc} \dots (\theta_{t-1}|D_{t-1}) & \xrightarrow{\text{Evolución}} & (\theta_t|D_{t-1}) & & (\theta_t|D_t) \dots \\ & & \Downarrow & & \Uparrow \\ & & (\eta_t|D_{t-1}) & \xrightarrow{\text{Actualización}} & (\eta_t|D_t) \end{array}$$

Análisis Secuencial

- Sea $D_t = \{y_1, \dots, y_t\}$ - la información total obtenida hasta t .
- En los **modelos lineales normales**:

$$\dots (\boldsymbol{\theta}_{t-1} | D_{t-1}) \xrightarrow{\text{Evolución}} (\boldsymbol{\theta}_t | D_{t-1}) \xrightarrow{\text{Actualización}} (\boldsymbol{\theta}_t | D_t) \dots$$

- En los **modelos lineales dinámicos generalizados**:

$$\begin{array}{ccccc} \dots (\boldsymbol{\theta}_{t-1} | D_{t-1}) & \xrightarrow{\text{Evolución}} & (\boldsymbol{\theta}_t | D_{t-1}) & & (\boldsymbol{\theta}_t | D_t) \dots \\ & & \Downarrow & & \Uparrow \\ & & (\boldsymbol{\eta}_t | D_{t-1}) & \xrightarrow{\text{Actualización}} & (\boldsymbol{\eta}_t | D_t) \end{array}$$

Aproximación utilizada por CUBS

- Sea D_t la información hasta el tiempo t . Sea $\Phi = (\psi, \phi)$.
- La **distribución condicional conjunta completa** de $\theta = (\theta_1, \dots, \theta_T)$ es:

$$\begin{aligned} p(\theta|Y, \Phi) &\propto p(\theta_T|D_T, \Phi) \prod_{t=1}^{T-1} \underbrace{p(\theta_t|\theta_{t+1}, D_t, \Phi)}_{\text{Densidad retrospectiva}} \\ &\propto p(\theta_T|D_T, \Phi) \prod_{t=1}^{T-1} p(\theta_{t+1} | \theta_t, D_t, \Phi) \underbrace{p(\theta_t | D_t, \Phi)}_{\text{A posteriori}} \end{aligned}$$

- Los momentos de las **distribuciones a posteriori** son aproximados por:

$$\begin{aligned} p(\theta_t | D_t, \Phi) &\propto p(\theta_t | D_{t-1}, \Phi) p(Y_t | \theta_t, \Phi) \\ &= \int p(\theta_t | \eta_t, D_{t-1}, \Phi) \underbrace{p(\eta_t | D_{t-1}, \Phi) p(Y_t | \eta_t, \Phi)}_{\text{Análise Conjugada}} d\eta_t \\ &\propto \int \underbrace{p(\theta_t | \eta_t, D_{t-1}, \Phi)}_{\text{Linear Bayes}} p(\eta_t | D_t, \Phi) d\eta_t = [\mathbf{m}_t, \mathbf{C}_t] \end{aligned}$$

$$\mathbf{m}_t = E[\theta_t | D_t, \Phi] = E[\hat{E}\{\theta_t | \eta_t, D_{t-1}\} | D_t, \Phi]$$

$$\mathbf{C}_t = V[\theta_t | D_t, \Phi] = V[\hat{E}\{\theta_t | \eta_t, D_{t-1}\} | D_t, \Phi] + E[\hat{V}\{\theta_t | \eta_t, D_{t-1}\} | D_t, \Phi]$$

CUBS para un modelo dinámico Gama

1 Haga $t = 1$

2 Calcule m_t y C_t :

1 Obtenga los momentos a priori de θ_t y $g(\eta_t)$, a partir do modelo:

$$\begin{array}{ll} \theta_t | D_{t-1} \sim [a_t, R_t] & \text{onde: } a_t = G_t m_{t-1}, \quad R_t = G_t C_{t-1} G_t' + W_t \\ g(\eta_t) | D_{t-1} \sim [f_t, q_t] & \text{onde: } f_t = F_t' a_t, \quad q_t = F_t' R_t F_t \end{array}$$

2 Determine a partir del análisis conjugado, los momentos a priori de $g(\eta_t)$:

$$\begin{array}{llll} E[g(\eta_t) | D_{t-1}] & = & \log r_t - \gamma(s_t + 1) & \approx \log r_t - \log(s_t + 1) = f_t \\ \text{Var}[g(\eta_t) | D_{t-1}] & = & \gamma'(s_t + 1) & \approx 1/(s_t + 1) = q_t \end{array}$$

3 Calcule los momentos a posteriori de $g(\eta_t)$:

$$\begin{array}{llll} E[g(\eta_t) | D_t] & = & \log(r_t + \phi y_t) - \gamma(s_t + \phi + 1) & = f_t^* \\ \text{Var}[g(\eta_t) | D_t] & = & \gamma'(s_t + \phi + 1) & = q_t^* \end{array}$$

4 Obtenga los momentos de la distribución a posteriori de θ_t , no modelo:

$$\theta_t | D_t \sim [m_t, C_t], \quad \text{donde: } m_t = a_t + R_t F_t (f_t^* - f_t) \frac{1}{q_t} \quad C_t = R_t - R_t F_t F_t' R_t \left(1 - \frac{q_t^*}{q_t}\right) \frac{1}{q_t}$$

3 Haga $t = t + 1$ y regrese para 2 si $t < T$;

4 Muestree θ_T de $N(m_T, C_T)$;

5 Haga $t = R - 1$, muestree θ_t de $p(\theta_t | \theta_{t+1}, D_t, \theta) = N(m_t^s, C_t^s)$;

6 Haga $t = t - 1$ y regrese para 5 si $t > 1$;

MCMC + CUBS

- ➊ **Inicio:** dar valores iniciales $\theta^{(0)}, \psi^{(0)}$ e iniciar las iteraciones, $i = 1$;
- ➋ **Muestrear $\theta^{(i)}$ usando CUBS:**
 - ➊ Calcular los momentos de $p(\theta_t | D_t, \psi^{(i-1)})$, $m^{(i)}$ e $C^{(i)}$, con el *Conjugate Updating*;
 - ➋ Muestrear θ^* con el *Backward Sampling*.
 - ➊ Muestrear θ_T^* de $\text{Normal}(m_T^{(i)}, C_T^{(i)})$
 - ➋ Muestrear $\theta_t^*, t = T - 1, \dots, 1$, de $p(\theta_t | \theta_{t+1}^*, \psi^{(i-1)})$
 - ➌ Haga $\theta^{(i)} = \theta^*$ con probabilidad p_t y $\theta^{(i)} = \theta^{(i-1)}$ con probabilidad $1 - p_t$, donde $p_t = \min(1, A)$ y A es la razón de aceptación del Metropolis-Hastings:

$$A = \min \left\{ 1, \frac{\omega(\theta^*)}{\omega(\theta)} \right\}, \quad \omega(\theta^*) = \frac{\pi(\theta^*)}{q(\theta^*)},$$

- ➌ **Muestrear $\psi^{(i)}$** usando, en general, un paso de Metropolis-Hastings ;
- ➍ **Muestrear $\phi^{(i)}$** usando, en general, un paso de Metropolis-Hastings;
- ➎ **Actualización:** haga $i = i + 1$ y regrese para 2 hasta la convergencia.

Precipitación en Tokyo

Referencias Kitagawa (1987) e Gamerman (1998).

Objetivo Estimar a probabilidad de ocurrencia de lluvia para cada día calendario.

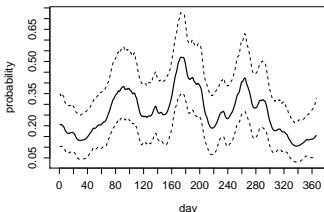
Datos Número de días con precipitación superior a 1mm en Tokyo para cada día entre 1983-1984.

Modelo

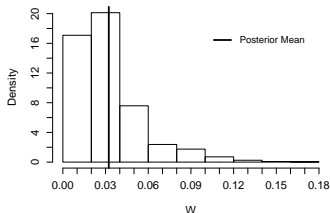
$$Y_t \sim \text{Binomial}(n, \pi_t), \quad t = 1, \dots, T$$

$$\text{logit}(\pi_t) = \alpha_t$$

$$\alpha_t = \alpha_{t-1} + w_t \quad w_t \sim N(0, W)$$



(a) Probabilidad de Lluvia



(b) Muestra a posteriori de W

Figura: Resultados obtenidos con MCMC+CUBS. **(a)** Probabilidad de lluvia: media a posteriori (línea sólida) e intervalo de 95% de credibilidad (línea punteada) **(b)** Distribución a posteriori empírica de W .