A Bayesian Approach for the Rainfall-Runoff Problem in Multiple Basins

Romy R. Ravines — romy@dme.ufrj.br

Alexandra M. Schmidt — alex@im.ufrj.br

Helio S. Migon — migon@im.ufrj.br

Instituto de Matemática. Universidade Federal do Rio de Janeiro. Brazil

Introduction: Rainfall-Runoff relationship

- **Rainfall-Runoff model**

Let Y_t be the basin **runoff** and X_t the basin **rainfall** at time t:

$$
\begin{aligned}
Y_t &\sim p(Y_t|\mu_t, \phi_t), \quad t = 1, 2, \ldots \\
g(\mu_t) &= f_1(\alpha_t, E_t) \\
E_t &= f_2(E_{t-1}, \ldots, E_0, X_t)
\end{aligned}
$$

- $p(Y_t|\mu_t, \phi_t)$ defined in $\mathbb{R}^+$: Gamma, log-normal, etc;
- α_t denotes a trend and E_t denotes the total effect of rainfall at time t;
- $g(\cdot)$, $f_1(\cdot)$ and $f_2(\cdot)$ are known non-linear functions describing the dynamics of the hydrological process.

- **Rainfall Effect on Runoff: A transfer function**

Following Migon and Monteiro (1997):

$$
\begin{aligned}
E_t &= \rho_t E_{t-1} + \gamma_t X_t \\
E_t &= \rho_t E_{t-1} + [1 - \exp(-\kappa_t X_t)][\vartheta_t - (\alpha_t + \rho_t E_{t-1})].
\end{aligned}
$$

- **Parameters interpretation**: α is the basic level or **stream flow**, ρ is the **recharge factor** or permanence rate of the rainfall effect and γ is the **velocity of response** to precipitation, related to soil saturation.
- **Some particular cases**:

$$
\begin{aligned}
&\gamma_t = \gamma \\
&\gamma_t = \gamma_{t-1} + \delta_t \\
&\gamma_t = \gamma + \delta_t; \quad \gamma \sim N(a, b)
\end{aligned}
$$

- **Basin Rainfall**

Let $\{X_t(s), s \in B \subset \mathbb{R}^2, t = 1, 2, \cdots\}$ be a stochastic process at discrete time t and over spatial domain B. In particular, $X_t(s)$ here represents the level of rainfall at time t and location s. The basin rainfall at time t is

$$
X_t = |B|^{-1} \int_B X_t(s) ds,
$$

where $|B|$ is the basin area. Following Sansó and Guenni (2000), $X_t(s)$ can be:

$$
\begin{aligned}
X_t(s_i) &= \begin{cases} w_t(s_i)^\beta & \text{if } w_t(s_i) > 0, \quad s_i \in B, \quad i = 1, \ldots, S, \\ 0 & \text{if } w_t(s_i) \leq 0, \end{cases} \\
\boldsymbol{w}_t(\boldsymbol{s}) &= \boldsymbol{\theta}_t \boldsymbol{f}(\boldsymbol{s}) + Z_t(\boldsymbol{s}) + \epsilon_t(\boldsymbol{s}), \quad \boldsymbol{s} = (s_1, \ldots, s_S)' \\
Z_t(\boldsymbol{s}) &\sim GP(\boldsymbol{0}, \sigma^2 \varrho(||s_1, s_2||, \lambda)),
\end{aligned}
$$

where $w_t(s_i)$ is a latent variable; $\boldsymbol{\theta}_t \boldsymbol{f}(\boldsymbol{s})$ is a polynomial trend; $Z_t(\boldsymbol{s})$ is a process with variance σ^2 and spatial correlation function $\varrho(||s - s'||, \lambda)$ (depending on λ); and $\epsilon_t(\boldsymbol{s})$ is a random error.

Proposed approach: Rainfall-Runoff in multiple basins

- **Multivariate Rainfall-Runoff model**

Let Y_t^m be the runoff at time t from basin m, $t = 1, \ldots, T; m = 1, \ldots, M$. The joint distribution of $\boldsymbol{Y}_t = (Y_t^1, \ldots, Y_t^M)'$ can be expressed, for example, as:

$$
p(\boldsymbol{Y}_t|\boldsymbol{\theta}) = p(Y_t^M|Y_t^{M-1}, \boldsymbol{\theta}) p(Y_t^{M-1}|Y_t^{M-2}, \boldsymbol{\theta}) \cdots p(Y_t^2|Y_t^1, \boldsymbol{\theta}) p(Y_t^1, \boldsymbol{\theta}),
$$

where $p(Y_t^m)$ is the (conditional) distribution of Y_t^m.

- **Rainfall Effect**

Let X_t^m be the rainfall at time t from basin m. Let $m = A, B, B \subset A$, then $\boldsymbol{Y}_t = (Y_t^A, Y_t^B)'$ and $\boldsymbol{X}_t = (X_t^A, X_t^B)'$ are the time series of both basins. Let $X_t^{A|B} = X_t^A - X_t^B$ be the rainfall at an area that belongs to basin A but not to basin B, then

$$
p(\boldsymbol{Y}_t|\boldsymbol{X}_t, \boldsymbol{\theta}) = p(Y_t^A|Y_t^B, X_t^{A|B}, \boldsymbol{\Theta}) p(Y_t^B|X_t^B, \boldsymbol{\Theta})
$$

- **Basins Rainfall**

Following Gelfand et al. (2001) one can use Monte Carlo integration, such that:

$$
X_t^m = \frac{1}{|m|} \int_m X_t(s) ds \approx \frac{1}{N_m} \sum_{i=1}^{N_m} \hat{X}_t(s_i) \quad i = 1, \ldots, N_m,
$$

where N_m is the number of points of a grid constructed **inside the limits** of the basin m and $\hat{X}_t(i)$ is the interpolated value for the s_i location of that grid.

- **Inference Procedure**

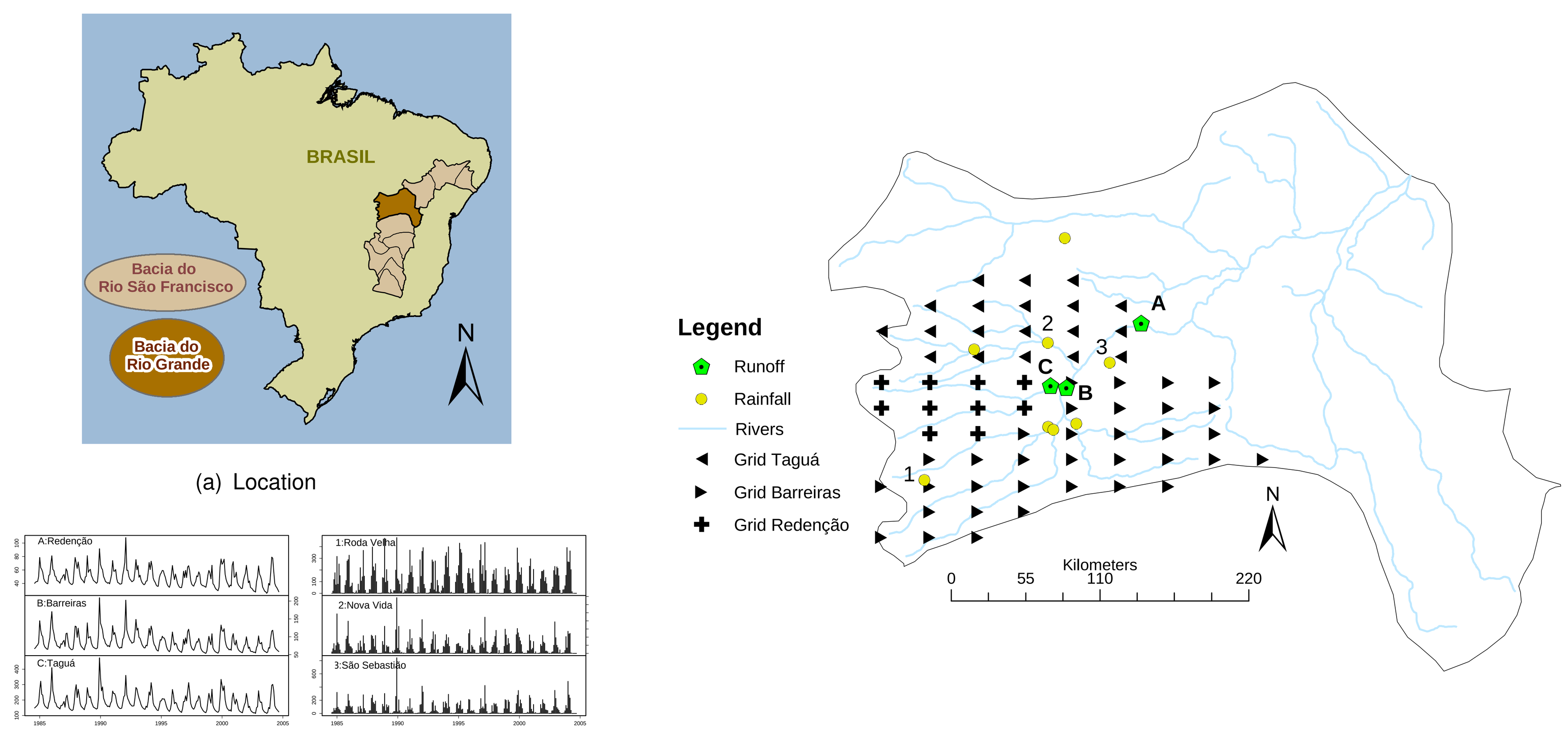
The joint distribution of $\boldsymbol{Y} = (Y_1, \ldots, Y_T)'$, $\boldsymbol{X} = (X_1, \ldots, X_T)'$ and $\boldsymbol{X}_t(\boldsymbol{s}) = (X_t(s_1), \ldots, X_t(s_S))'$ is

$$
p(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{X}(s)|\boldsymbol{\Theta}) = \prod_{t=1}^{T} p(Y_t|X_t, \boldsymbol{X}_t(\boldsymbol{s}), \boldsymbol{\Theta}_Y) p(X_t|\boldsymbol{X}_t(\boldsymbol{s}), \boldsymbol{\Theta}_X) \prod_{i=1}^{S} p(\boldsymbol{X}_t(\boldsymbol{s}_i)|\boldsymbol{\Theta}_X),
$$

where $\boldsymbol{\Theta} = (\boldsymbol{\Theta}_Y, \boldsymbol{\Theta}_X)$, $\boldsymbol{\Theta}_Y$ are parameters in the runoff model and $\boldsymbol{\Theta}_X$ are parameters in the rainfall model.

MCMC techniques: Conjugate updating backward sampling, **CUBS** (Ravines, Migon, & Schmidt, 2007), for Gamma transfer functions.

Application: Rio Grande Basin

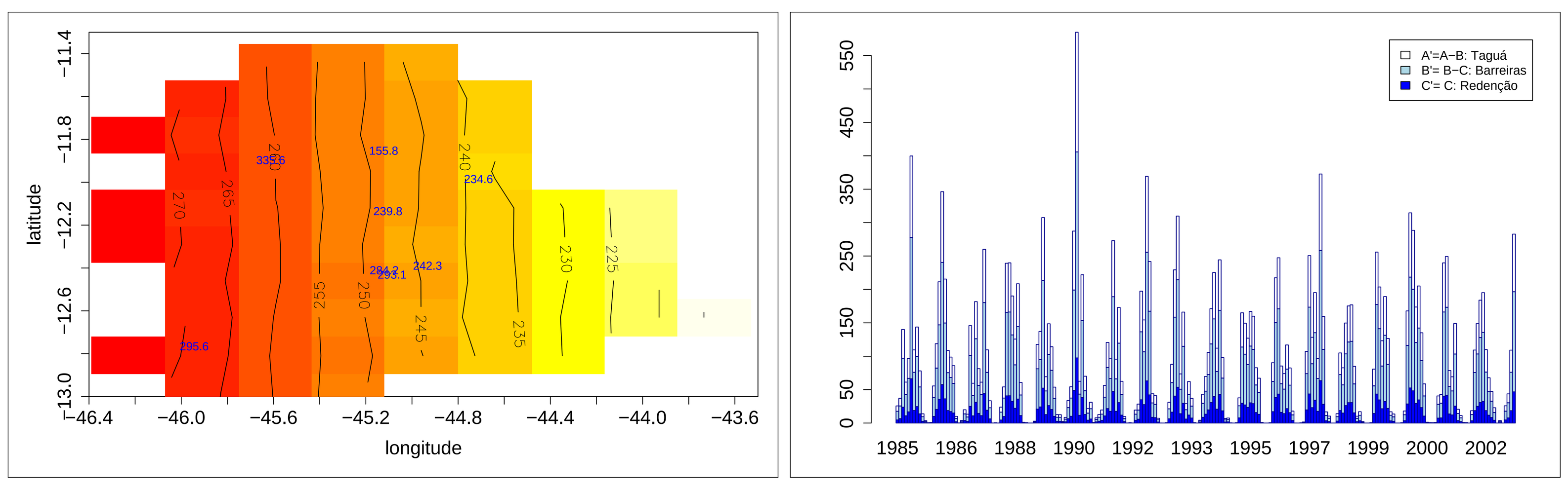


(a) Location (b) Runoff (c) Rainfall (d) Sub-basins and Stations

Figure 1: **Rio Grande Basin**: Location, sub-basins, monitoring sites, interpolation grids and some time series.

Data: monthly recorded series from January 1984 to September 2004, at three runoff stations and nine rainfall stations irregularly located in an area of drainage of $37522.48\,km^2$.

- **Results: Spatial Interpolation**



(a) December 2000 (b) Basins rainfall

Figure 2: Posterior mean of rainfall in a specific month and the aereal rainfall for all the sub-basins

- **Results: Multivariate transfer functions**

The model:

$$
\begin{aligned}
Y_t^A|Y_t^B &\sim \text{Lognormal}(\mu_t^{A|B}, \sigma^2_{A|B}) \quad t = 1, \ldots, T \\
Y_t^B|Y_t^C &\sim \text{Lognormal}(\mu_t^{B|C}, \sigma^2_{B|C}) \\
Y_t^C &\sim \text{Lognormal}(\mu_t^C, \sigma^2_C) \\
\mu_t^{A|B} &= \alpha^{A|B} + \eta^{A|B} Y_t^B + E_t^{A|B} \\
\mu_t^{B|C} &= \alpha^{B|C} + \eta^{B|C} Y_t^C + E_t^{B|C} \\
\mu_t^C &= \alpha^C + E_t^C \\
E_t^m &= \rho^m E_{t-1}^m + \gamma^m X_t^m |m| + w_t^m \\
w_t^m &\sim \text{Normal}(0, W_m) \qquad m = A|B, B|C, C;
\end{aligned}
$$

Table 1: Comparison with univariate models

	MSE	MAE
Multivariate model		
$A\|B$	120.684	7.025
$B\|C$	9.060	1.820
C	2.568	1.009
Univariate independent models		
A	162.991	7.756
B	8.047	1.652
C	2.568	1.009

Table 2: Some posterior summaries

	$A\|B$ = **Taguá | Barreiras**						$B\|C$ = **Barreiras | Redenção**						C = **Redenção**				
	mean	25%	50%	75%	$\hat{R}$		mean	25%	50%	75%	$\hat{R}$		mean	25%	50%	75%	$\hat{R}$
$\alpha^{A\|B}$	1.151	1.022	1.142	1.276	1.00	$\alpha^{B\|C}$	0.587	0.507	0.581	0.654	1.01	α^C	3.543	3.525	3.544	3.561	1.01
$\eta^{A\|B}$	0.915	0.887	0.916	0.945	1.01	$\eta^{B\|C}$	0.989	0.971	0.990	1.009	1.01						
$\rho^{A\|B}$	0.510	0.387	0.533	0.658	1.00	$\rho^{B\|C}$	0.743	0.692	0.752	0.806	1.00	ρ^C	0.594	0.575	0.594	0.613	1.03
$\gamma^{A\|B}$	0.025	0.014	0.024	0.035	1.00	$\gamma^{B\|C}$	0.005	0.003	0.005	0.008	1.00	γ^C	1.645	1.597	1.645	1.696	1.00
$W_{A\|B}$	0.003	0.002	0.003	0.004	1.01	$W_{B\|C}$	0.003	0.002	0.002	0.003	1.00	W_C	0.005	0.005	0.005	0.006	1.01
$\sigma^2_{A\|B}$	0.007	0.006	0.007	0.009	1.04	$\sigma^2_{B\|C}$	0.004	0.003	0.004	0.005	1.07	σ^2_C	0.002	0.002	0.002	0.003	1.10

Final Remarks

- Main contribution: **proposed approach.** Modeling **simultaneously** rainfall and runoff taking into account the different spatial units in which they are measured.
- Features of the proposed models: its parameters have **physical interpretations** and assumptions of normality or stationarity of the time series are not needed.
- Results show that our approach is a **promising tool** for the runoff-rainfall analysis.
- An extension: **hierarchical models** to handle data from several basins simultaneously.

References

Gelfand, A. E., Zhu, L., & Carlin, B. (2001). On the change of support problem for spatio-temporal data. *Biostatistics*, *2*(1), 31–45.

Migon, H., & Monteiro, A. B. (1997). Rain-fall modelling: An application of Bayesian forecasting. *Stochastic Hydrology and Hydraulics*, *11*, 115–127.

Ravines, R., Migon, H., & Schmidt, A. (2007). *An efficient sampling scheme for generalized dynamic models* (Tech. Rep.). Instituto de Matemática, Universidade Federal do Rio de Janeiro.

Sansó, B., & Guenni, L. (2000). A non-stationary multi-site model for rainfall. *Journal of the American Statistical Association*, *95*, 1089–1100.