

Introducción

El modelo climático regional (RCM) es una herramienta de reducción de escala que utiliza modelos dinámicos e híbridos, en conjunto con las variables de salida de modelos globales climáticos (GCMs), y el uso de fórmulas dinámicas complejas. Dado que esta técnica de reducción de escala dinámica es computacionalmente intensiva, es de gran valor realizar aproximaciones rápidas para llevar a cabo análisis de sensibilidad. Este estudio introduce un emulador de RCM que utiliza variables salida de GCMs como covariables y un marco estadístico Bayesiano para proveer una aproximación rápida a las variables de salida de un RCM. El emulador tiene coeficientes que varían espacialmente y previas informativas para realizar selección de variables por locación, de una manera eficiente. Para este fin, se modelan tanto los coeficientes como la variable respuesta como procesos espaciales. El método es aplicado para emular la precipitación del modelo regional CRCM utilizando variables de salida del modelo global CCSM con datos del Programa Regional de Evaluación del Cambio Climático de América del Norte (NARCCAP).

Datos

Se utilizan medidas espacio-temporales de precipitación, producidas por el modelo regional CRCM y 5 variables del modelo de circulación global CCSM.

Respuesta: Precipitación Mensual (PR) producida por el modelo regional CRCM de NARCCAP (locación s_k).

Covariables: 5 covariables del modelo de circulación global CCSM, por año (solo en verano) y para cada cuadrícula del CCSM: PSL, TREFHT, OMEGA (510.45), U (266.48), V (266.48), Qbot.

Table: Estadísticas Descriptivas para las covariables.

	log(PR)	PSL	TREFHT	OMEGA (510.45)	U (266.48)	V (266.48)	Qbot
Min.	-13.82	100.949	273	-0.07888	-4.51	-7.540	-14.796
1st Qu.	-12.20	101.399	287	-0.00963	10.04	0.557	-3.972
Median	-11.56	101.551	291	0.00732	15.12	3.762	-0.704
Mean	-11.62	101.674	292	0.01255	14.55	3.718	-1.664
3rd Qu.	-10.94	101.882	296	0.03098	19.44	6.660	0.448
Max.	-8.91	103.057	305	0.11237	32.57	16.438	7.481

Modelo Estadístico

Sea $Y_t(s)$ el output del modelo RCM para la locación s con la resolución de RCM para el año t . $Y_t(s)$ se modela como:

$$Y_t(s) = \mu_t(s) + \varepsilon_t(s), \quad (1)$$

donde $\varepsilon_t \sim GP(0, C(h, \Omega))$ y $C(h, \Omega)$ es la función de covariancia espacial que describe el proceso espacial suavizado utilizando la distancia h y un conjunto de parámetros Ω . Se asume independencia en el tiempo, sin embargo se podría incorporar en la estructura de ε . $\mu_t(s)$ está representado por:

$$\mu_t(s) = \sum_{k=0}^p X_{kt}(s) \beta_k(s), \quad \text{with} \quad (2)$$

$$\beta_k(s) = v_k(s) f_k(s) = v_k(s) \sum_{j=1}^J \alpha_{kj} F_j(s) \quad (3)$$

donde $X_{kt}(s)$ contiene el output de GCM para la covariable k en el año t para cada locación de GCM en la cuadrícula que contiene la locación s del modelo regional (RCM) correspondiente, y $X_{0t}(s) = 1$.

Sea $f_k(s)$ una representación de bases, y defina $\Phi_{ij} = \phi(s_i - s_j)$ dado la matriz Φ . Se utiliza la descomposición del valor espectral para escribir $\Phi = UDU^T$, donde D es una diagonal y U contiene las J eigen funciones respectivas, ordenadas según la magnitud de los eigen valores y evaluadas en cada locación s_i . Sea U_j el j avo eigenvector y Φ y $F_j(s_i)$ los elementos i ésimos de U_j . Se utilizan thin-plate splines basis functions para construir Φ , donde $\phi(|s_i - s_j|) = |s_i - s_j|^2 \log |s_i - s_j| / \eta$, y η es un parámetro de escala.

Estudio de Simulación

Table: Tiempo para correr 10000 MCMC iteraciones por modelo, $N = 100$, $T = 12$.

Model	Time (minutes)
(1b) Independent Y / pre-estimate w	21.6
(2b) Exponential covariance Y / pre-estimate w	70.2
(2c) Exponential covariance Y / fix covariance w	84.1
(Full Bayesian) Exponential covariance Y / complete model	324.0

Table: Resultados out-of-sample para los emuladores utilizando distintos valores de J : número de bases. Errores estándar menores a 0.05 para MSE, MAD and MLS.

	MSE(Y)	MAD(Y)	$Y \in CI$	MLS
Model 1b, $J = 5$	1.36	0.93	0.92	0.8634
Model 1b, $J = 10$	1.94	1.10	0.95	0.9988
Model 1b, $J = 20$	2.05	1.21	0.96	1.0034
Model 2b, $J = 5$	1.06	0.81	0.89	0.8116
Model 2b, $J = 10$	1.84	1.08	0.92	0.8542
Model 2b, $J = 20$	1.95	1.15	0.91	0.8481
Gaussian Model	2.29	1.29	1.00	2.9967

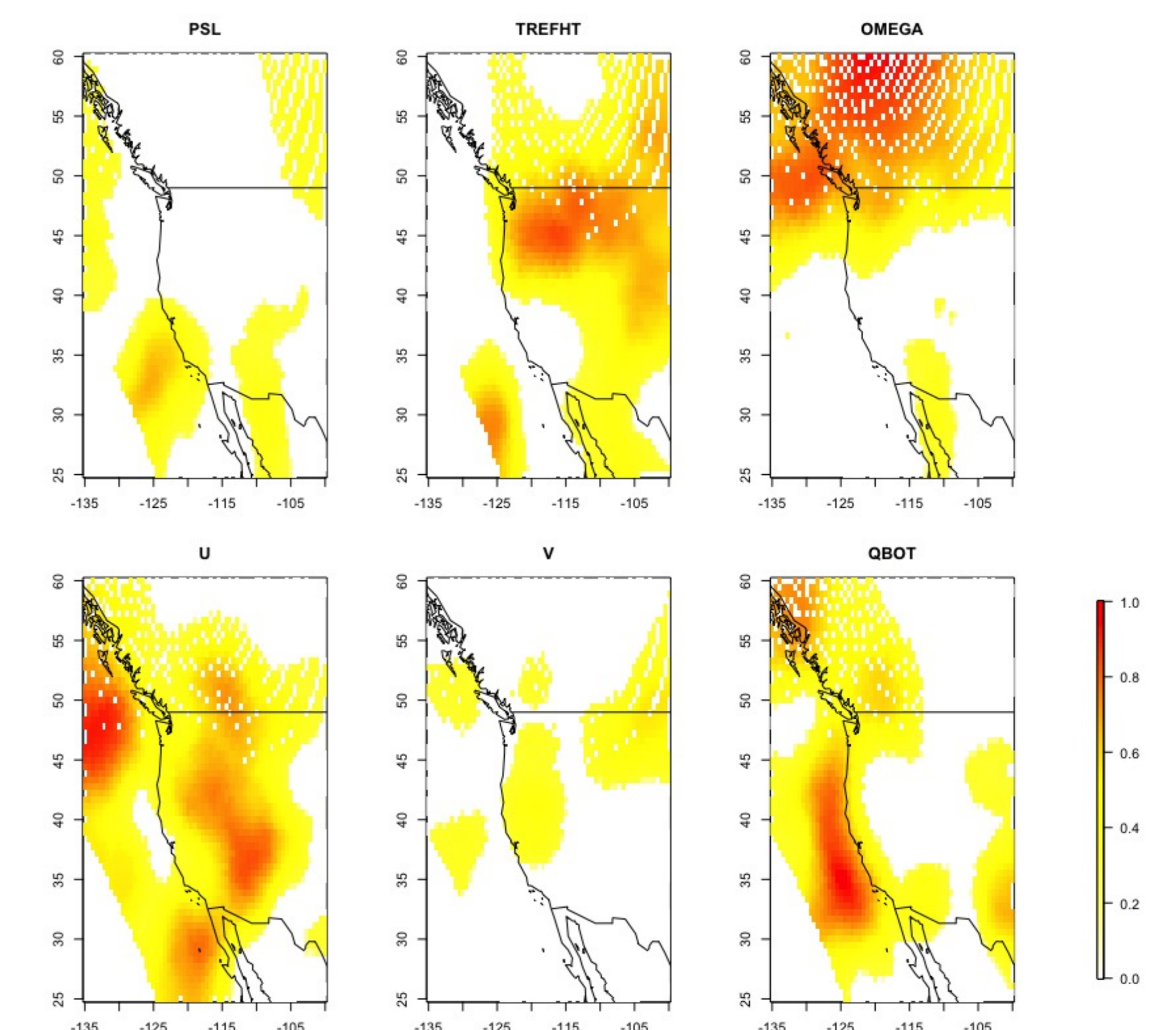


Figure: Pesos $v_k(s)$ para cada covariable en el espacio. Areas en blanco representan valores que no son significativamente distintos de cero: areas en las que la covariable respectiva no representa un driver significativo para la precipitación durante el verano.

Discusión

El modelo propuesto prueba ser una opción flexible para describir estructuras espaciales complejas. El uso de previas basadas en los datos facilita la selección de locaciones en el espacio, manteniendo el ajuste del modelo computacionalmente posible. Como trabajo en proceso, el modelo se está traduciendo a INLA, para probar nuevas formas de estimación.

Referencias: [1] Raftery et al (2003). *JASA*, 464:931–938. [2] Hoeting et al (1999). *Statistical Science*, 3:193–195. [3] Mearns et al (2009). *Eos, Transactions American Geophysical Union* 36:311:311.

Reconocimientos: STAMOS Fellowship, MICIT CR. Esta investigación forma parte del Proyecto B7166 de la Vicerrectoría de Investigación de la Universidad de Costa Rica: "Propuesta, desarrollo y prueba de un emulador estadístico para la reducción de escala de modelos climáticos de circulación global".