Case Study I: Combining Ensembles

Markov Random Fields and Regional Climate Models

Stephan R. Sain

Geophysical Statistics Project

Institute for Mathematics Applied to Geosciences

National Center for Atmospheric Research

Boulder, CO

Supported by NSF ATM/DMS. Thanks also to Reinhard Furrer and Noel Cressie.

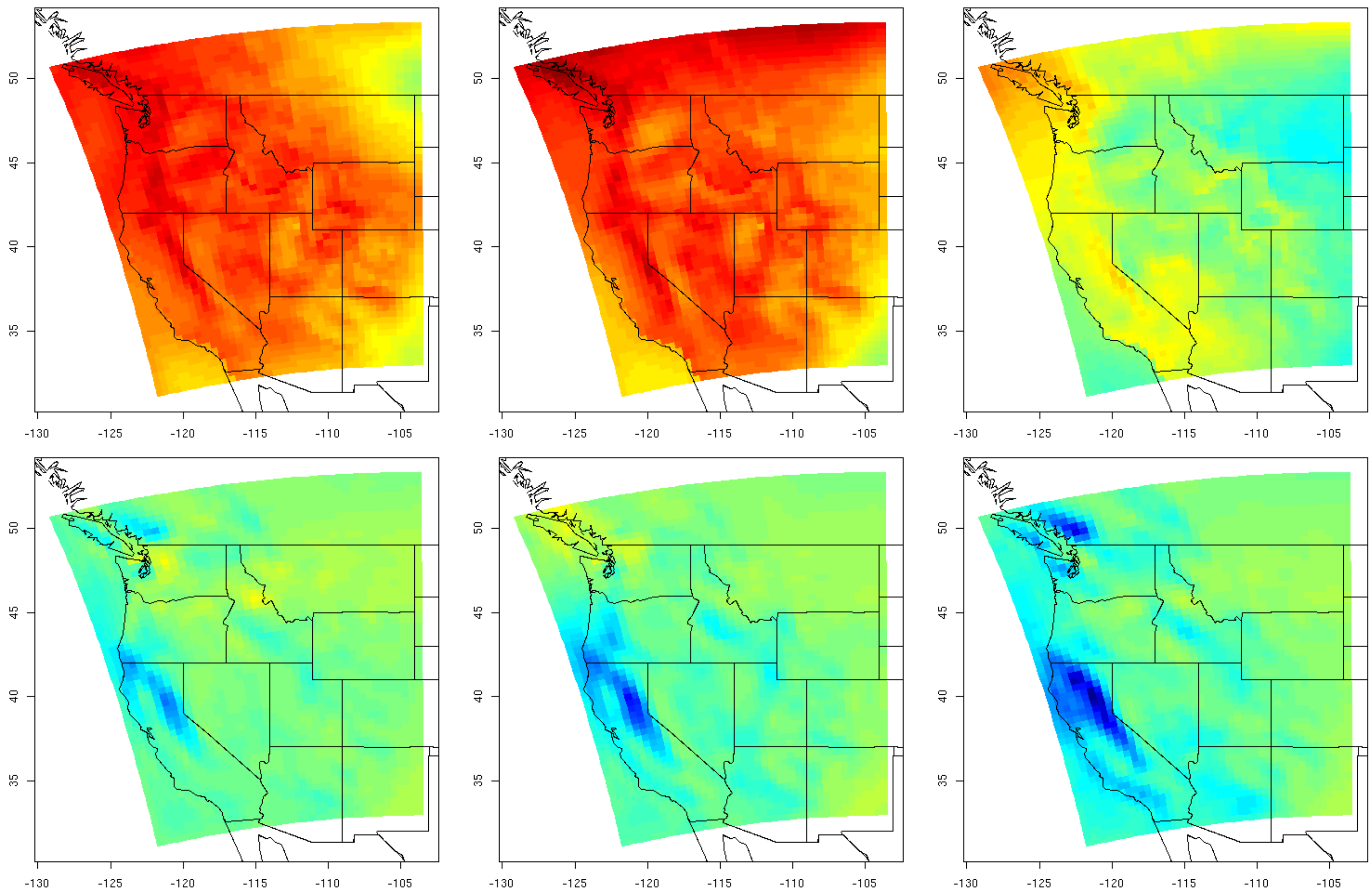
An RCM Example

- Driven by the NCAR/DOE Parallel Climate Model, the MM5 RCM was used to produce a control run and three future runs.
  - Control: 1995-2015
  - Future: 2040-2060
- Domain: western US, part of western Canada
- Climate scenario: "business as usual"
  - 1% yearly increase in greenhouse gasses
- Daily max/min temperature and precipitation

An RCM Example

- $n = 2464$ grid boxes on a regular lattice.
- Max/min temperature converted into midpoint (and range).
- 20-year winter (DJF) "averages" computed for each grid box.
- Two variables: Δ_{Tmid}, Δ_{Precip}
  - $\Delta = \text{Future}_i - \text{Control}, \; i = 1, 2, 3$

An RCM Example



A Hierarchical Model

- Data model:

$$\mathbf{y}_{rj} \sim \mathcal{N}\left(\mathbf{X}_1\boldsymbol{\alpha}_j + \mathbf{X}_2\boldsymbol{\beta}_{rj} + \mathbf{h}_{rj}, \boldsymbol{\Sigma}_j\right), \quad r = 1, \ldots, m, \quad j = 1, \ldots, p,$$

  - $\mathbf{X}$ includes intercept, longitude, latitude, and elevation.

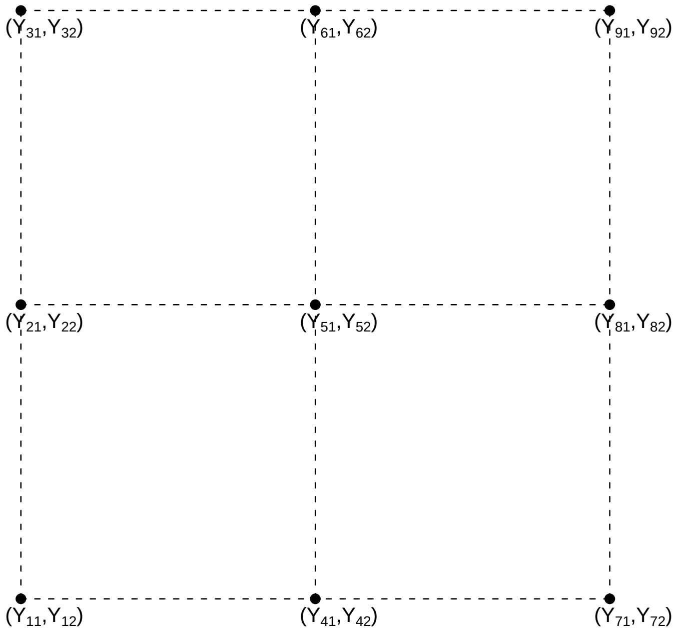
- Process model:

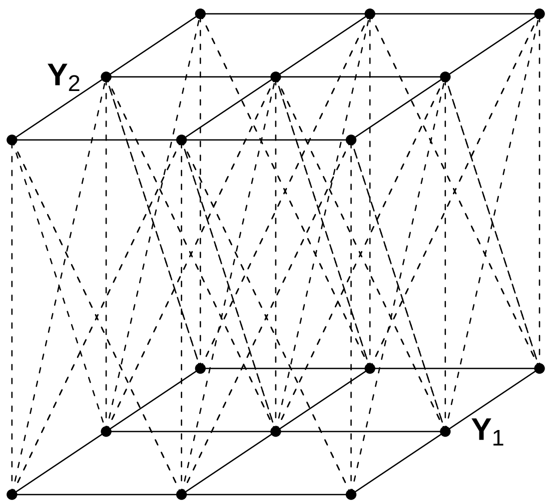
$$\begin{pmatrix} \boldsymbol{\beta}_{r1} \\ \vdots \\ \boldsymbol{\beta}_{rp} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_p \end{pmatrix}, \boldsymbol{\Sigma}_b\right),$$

$$\begin{pmatrix} \mathbf{h}_{r1} \\ \vdots \\ \mathbf{h}_{rp} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_p \end{pmatrix}, \mathbf{V}\left(\{\tau_j^2\}, \{\rho_{j\ell}\}, \{\phi_{j\ell}\}\right)\right).$$

- Priors: Non-informative.

A Different Approach





To improve interpretability and identifiability, rethink the lattice: rather than multivariate observations on a bivariate lattice, consider univariate observations on a stacked lattice.

A Different Approach

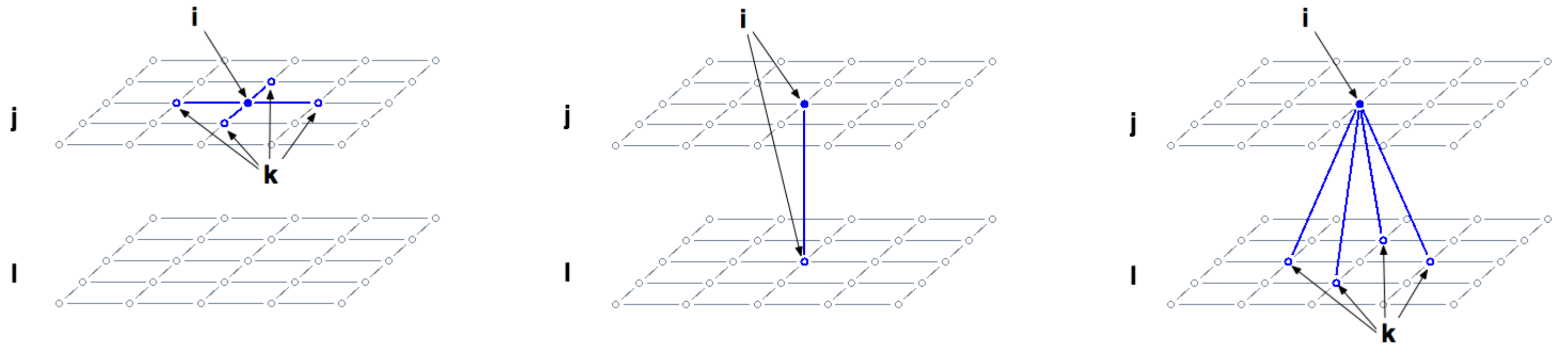
- Write the mean/variance of each conditional distribution as

$$
\begin{aligned}
\mathrm{E}[y_{ij}|y_{-\{ij\}}] = \mu_{ij} \;\; & + \;\; \sum_k b_{ijkj}(y_{kj} - \mu_{kj}) \qquad \text{(left)} \\
& + \;\; \sum_\ell b_{iji\ell}(y_{i\ell} - \mu_{i\ell}) \qquad \text{(middle)} \\
& + \;\; \sum_{k,\ell} b_{ijk\ell}(y_{k\ell} - \mu_{k\ell}) \qquad \text{(right)}
\end{aligned}
$$

and

$$
\mathrm{Var}[y_{ij}|y_{-\{ij\}}] = \tau_{ij}^2,
$$

A Different Approach



A Different Approach

- This formulation gives rise to a joint distribution of the form $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where

$$\boldsymbol{\Sigma} = \left[\mathbf{I}_n \otimes \boldsymbol{\tau}^{1/2}\right] \left[\mathbf{I}_n \otimes \mathbf{A} - \begin{bmatrix} \mathbf{0} & & \mathbf{B} I_{ij} \\ & \ddots & \\ \mathbf{B}' I_{ij} & & \mathbf{0} \end{bmatrix}\right]^{-1} \left[\mathbf{I}_n \otimes \boldsymbol{\tau}^{1/2}\right],$$

where $\boldsymbol{\tau}^{1/2} = [\tau_1, \ldots, \tau_p]'$ and

$$\mathbf{A} = \begin{bmatrix} 1 & & -\rho_{j\ell} \\ & \ddots & \\ -\rho_{j\ell} & & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} -\phi_{11} & & -\phi_{j\ell} \\ & \ddots & \\ -\phi_{\ell j} & & -\phi_{pp} \end{bmatrix}.$$

Bayesian Computation

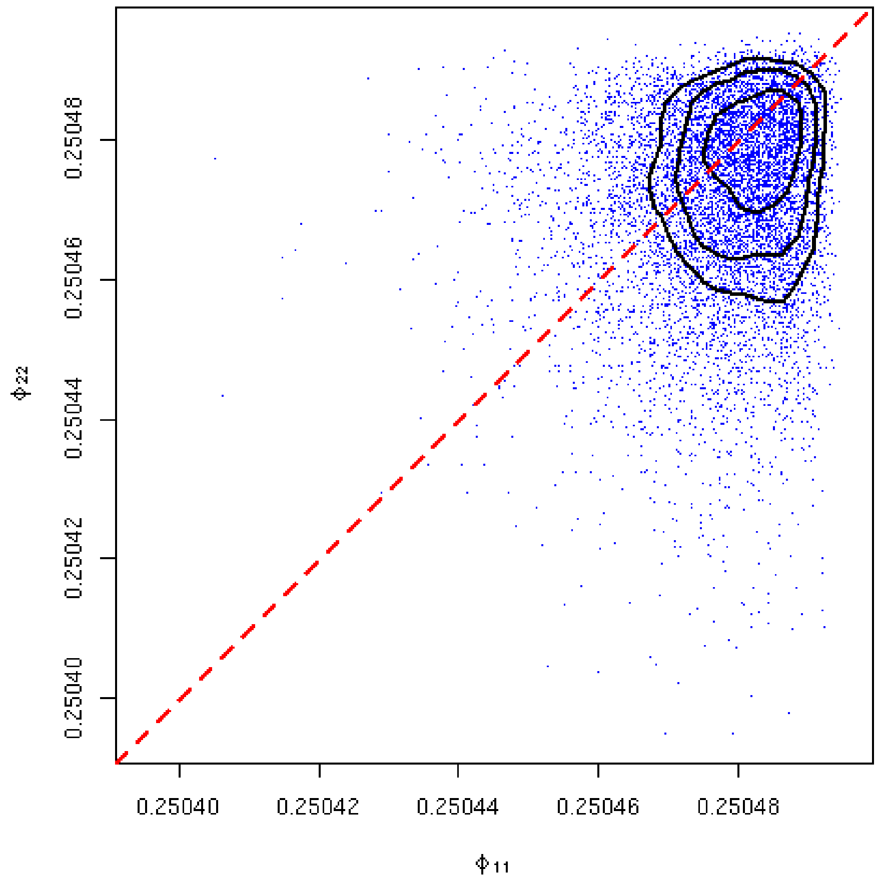
- Not much hope of computing posterior in closed form.

- Use computational methods (MCMC) to probe the posterior.
  - The Gibb's sampler.
  - Metropolis-Hastings – accept/reject method.

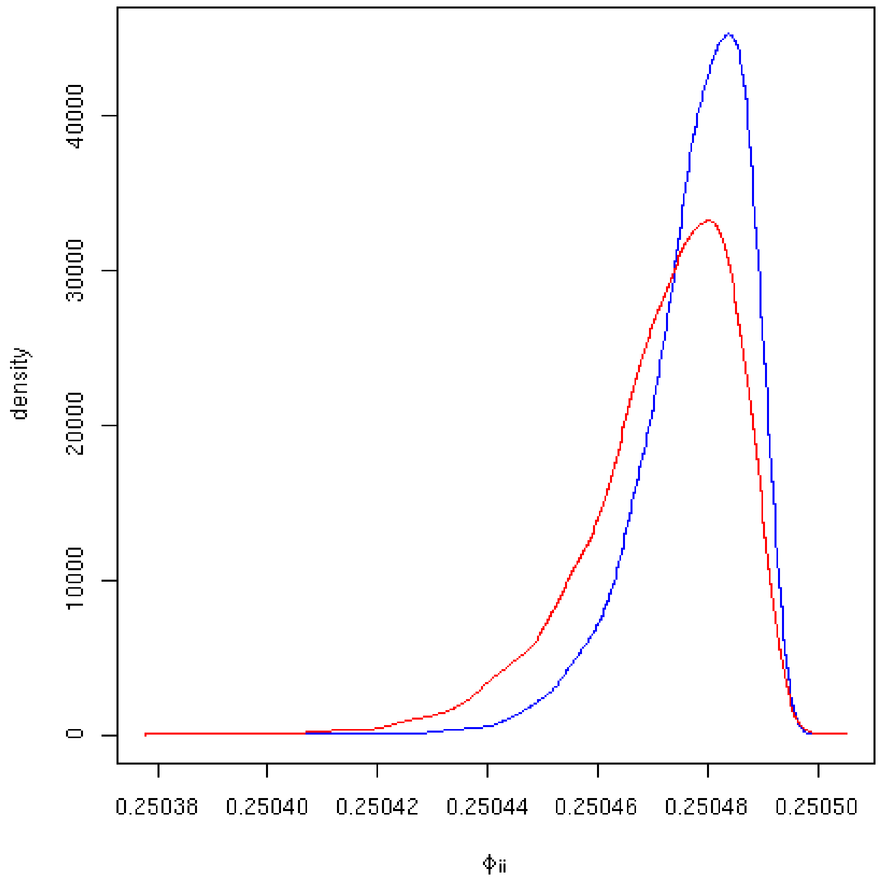
MCMC

- 3 regimes of 10K iterations each
  - Regime 1: all single variable M-H with periodic updates of the proposal distribution.
  - Regime 2: ρ, ϕ_{12}, and ϕ_{21} blocked M-H with periodic updates of the proposal distribution.
  - Regime 3: ρ, ϕ_{12}, and ϕ_{21} blocked M-H with no further updates.
- 10 chains run with random starts from the prior.
  - Choose ρ and $\boldsymbol{\phi}$ to yield positive-definite $\mathbf{H}$.
- Sparse matrix methods (`spam`) crucial to computation.

ϕ_{11}, ϕ_{22}

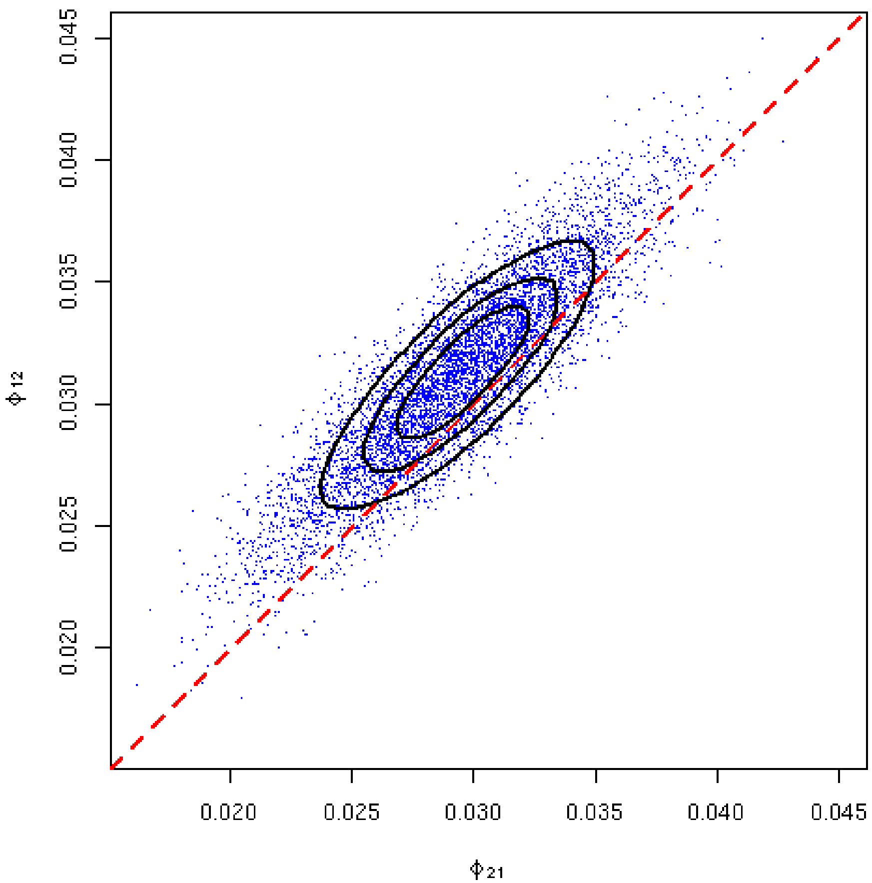


0.25040
0.25042
0.25044
0.25046
0.25048
ϕ11
0.25040
0.25042
0.25044
0.25046
0.25048
ϕ22

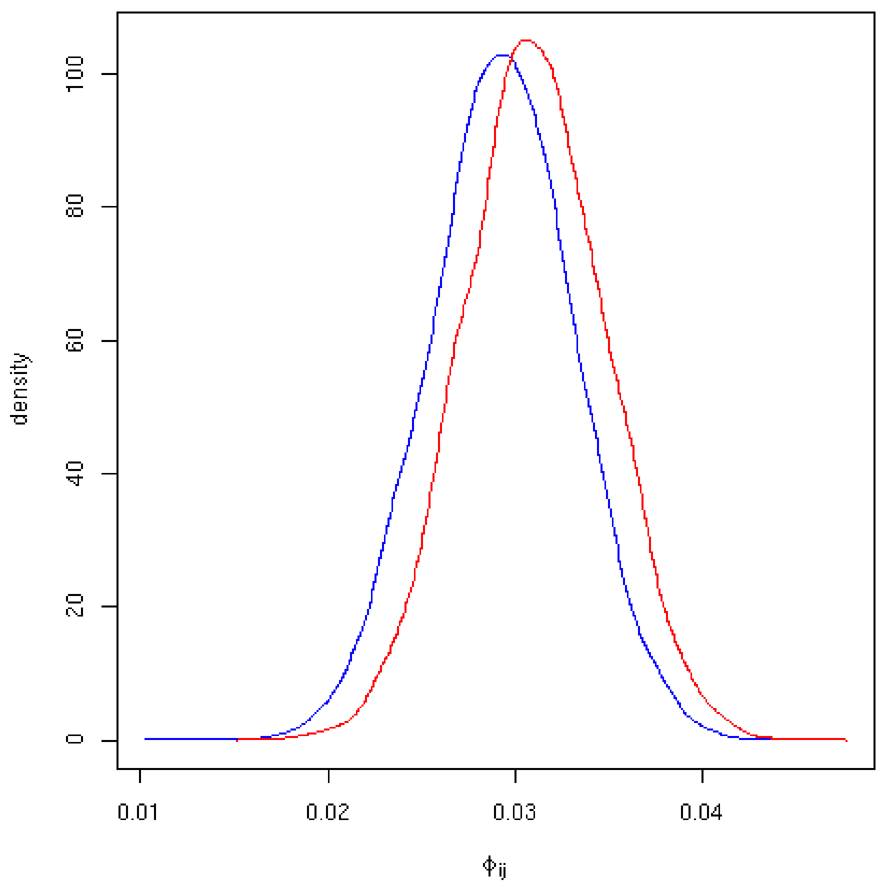


0.25038
0.25040
0.25042
0.25044
0.25046
0.25048
0.25050
ϕii
0
10000
20000
30000
40000
density

ϕ_{12}, ϕ_{21}

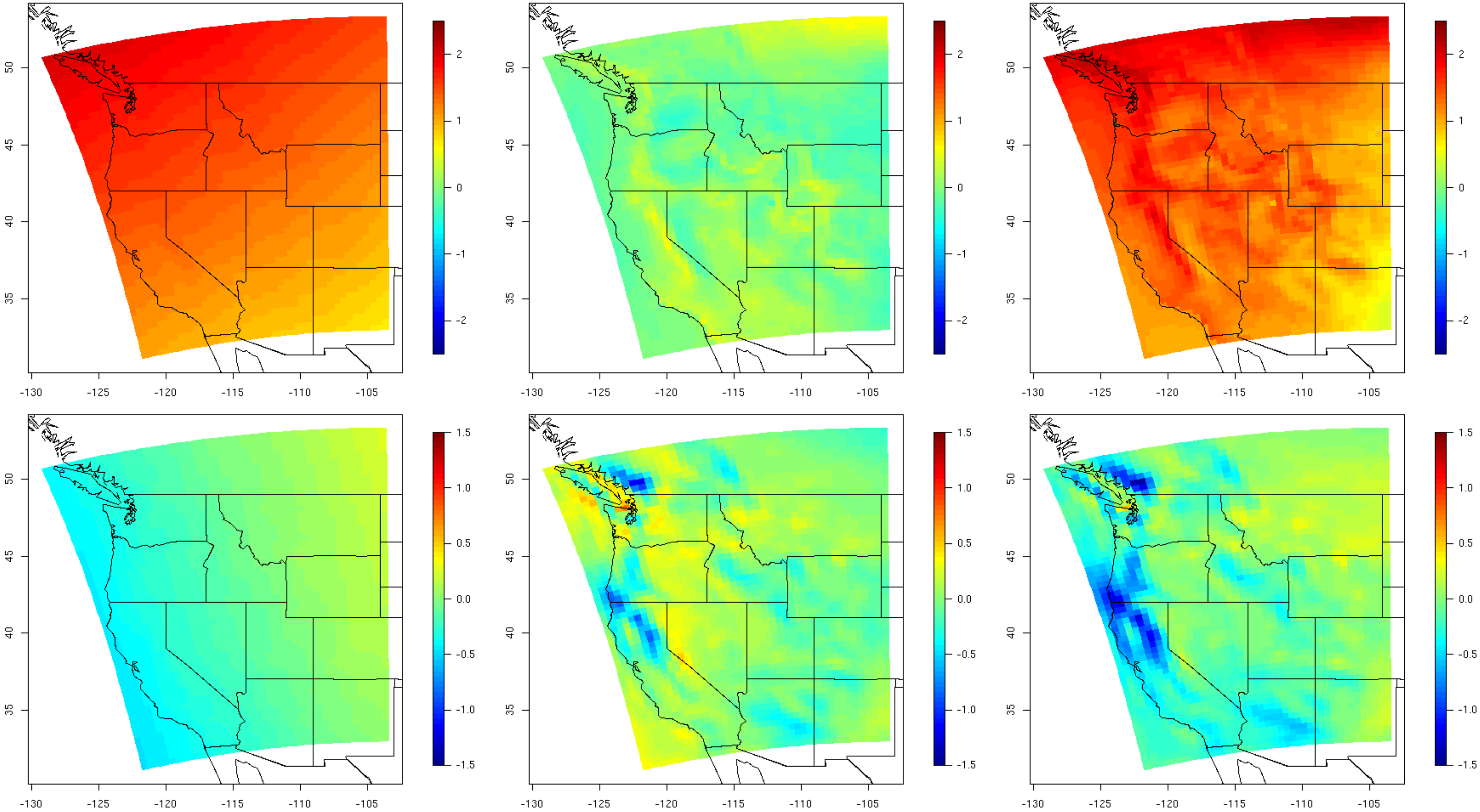


0.045
0.040
0.035
0.030
0.025
0.020
ϕ12
0.020
0.025
0.030
0.035
0.040
0.045
ϕ21

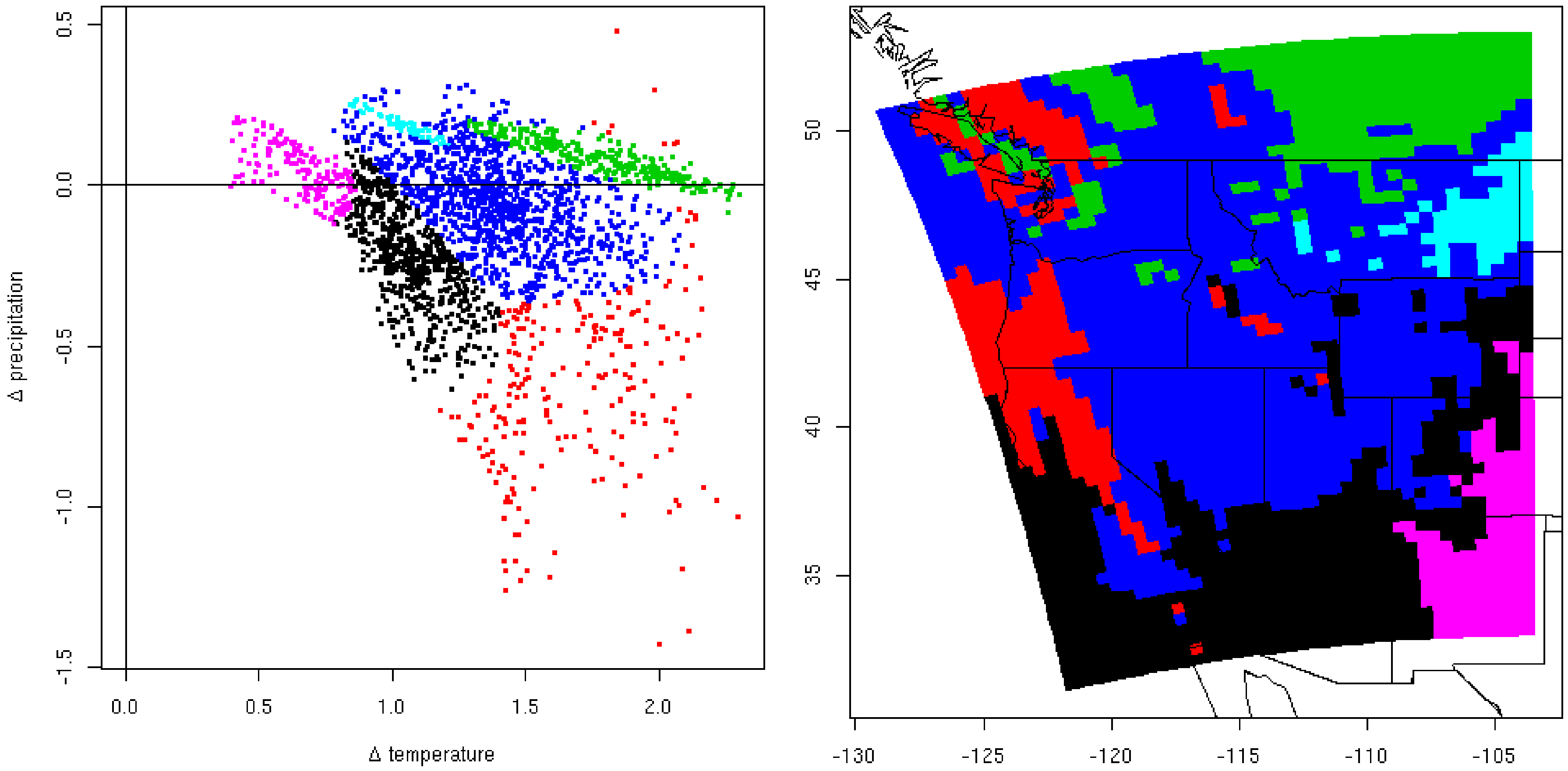


100
80
60
40
20
0
density
0.01
0.02
0.03
0.04
ϕij

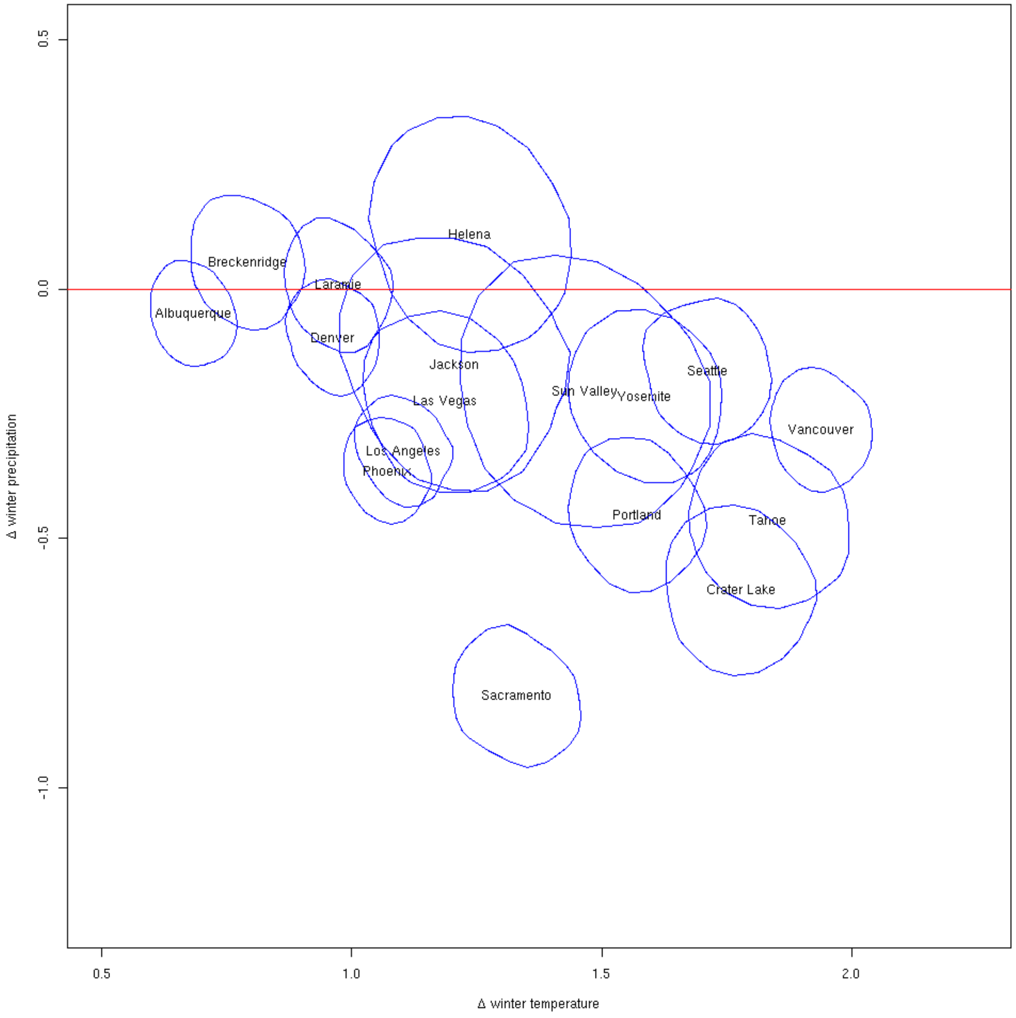
Posterior Means



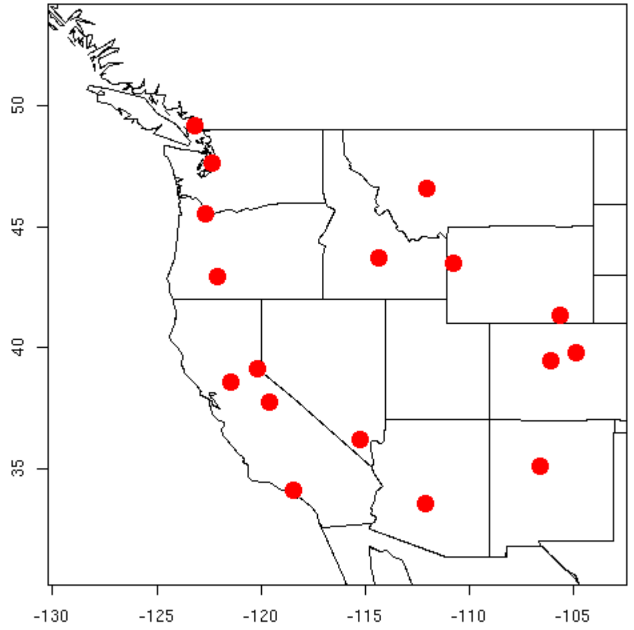
Clustering


0.5
0.0
-0.5
-1.0
-1.5
Δ precipitation
0.0
0.5
1.0
1.5
2.0
Δ temperature
50
45
40
35
-130
-125
-120
-115
-110
-105

Impacts

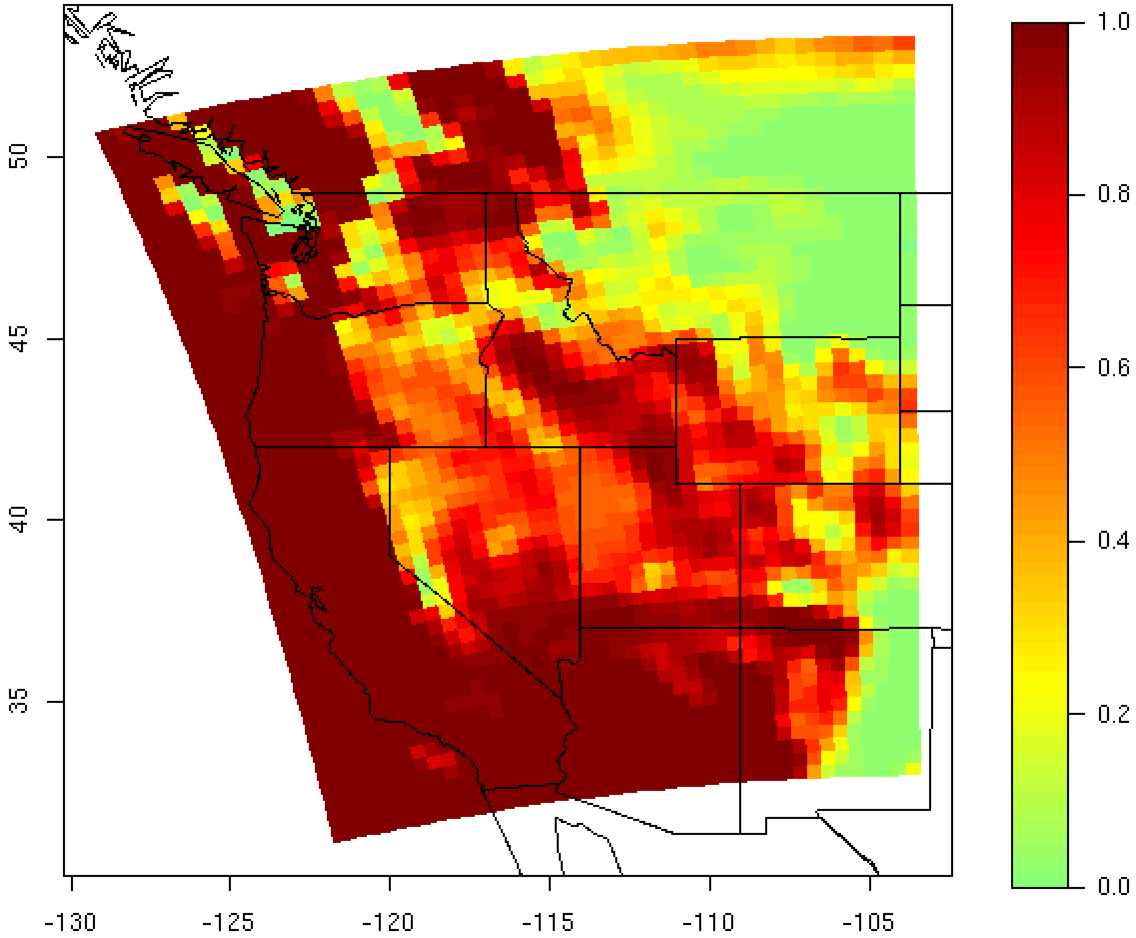


Helena
Breckenridge
Laramie
Albuquerque
Denver
Jackson
Las Vegas
Sun Valley
Yosemite
Seattle
Vancouver
Los Angeles
Phoenix
Portland
Tahoe
Crater Lake
Sacramento
0.5
0.0
-0.5
-1.0
0.5
1.0
1.5
2.0
Δ winter precipitation
Δ winter temperature

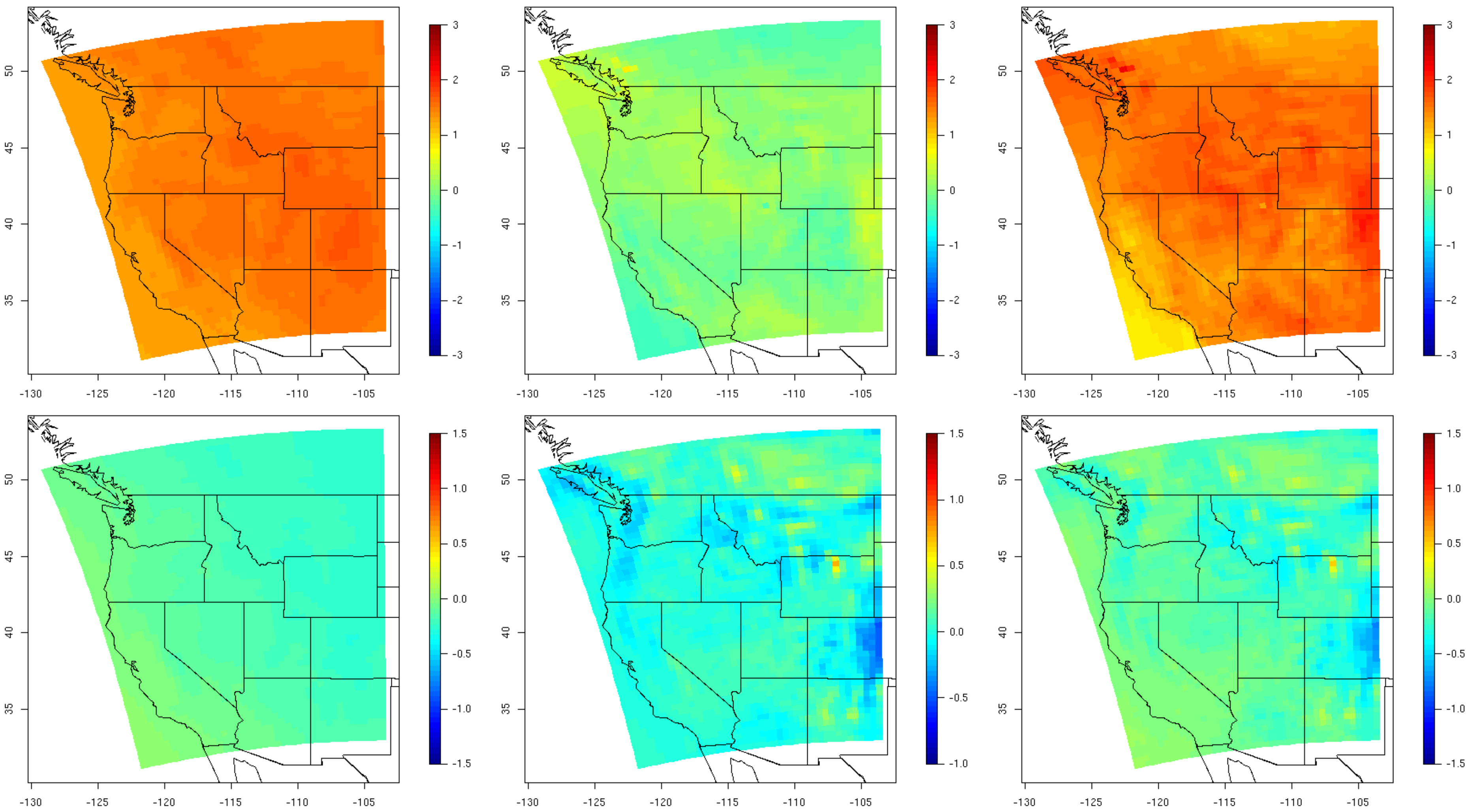


50
45
40
35
-130
-125
-120
-115
-110
-105

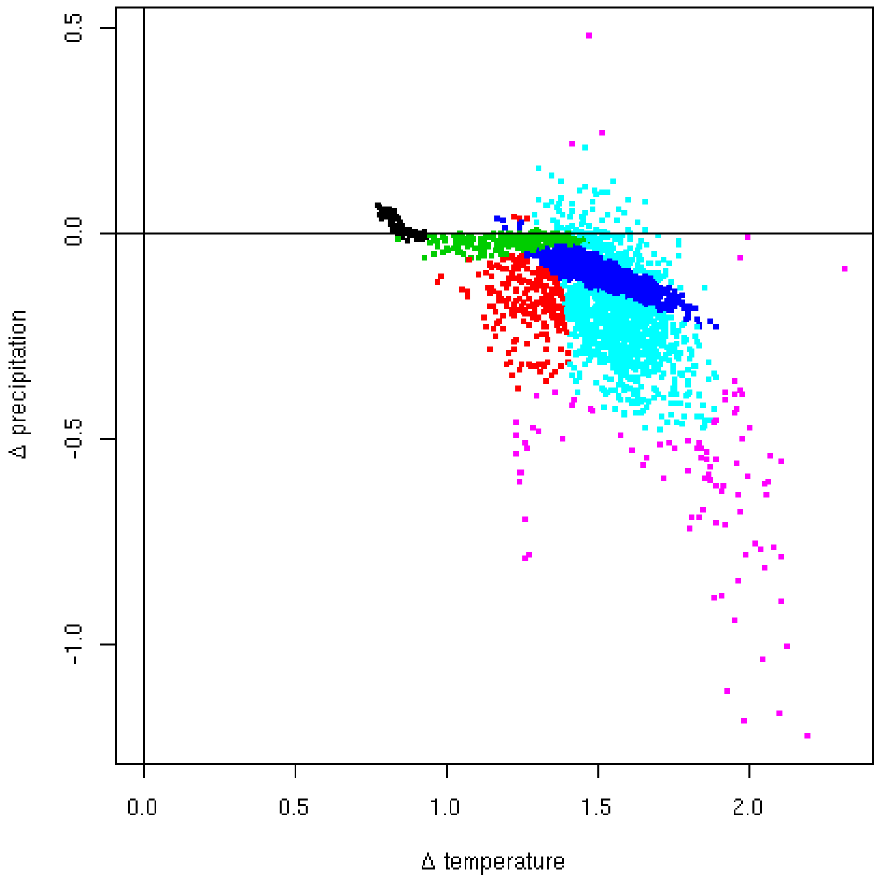
Impacts

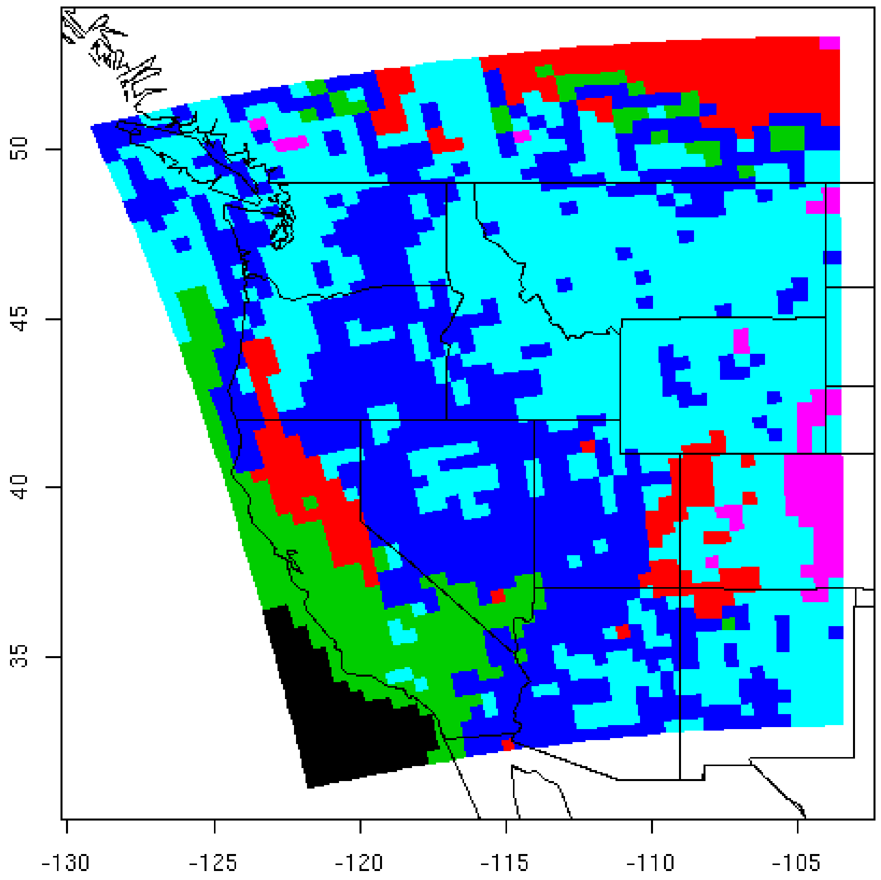


Posterior Means

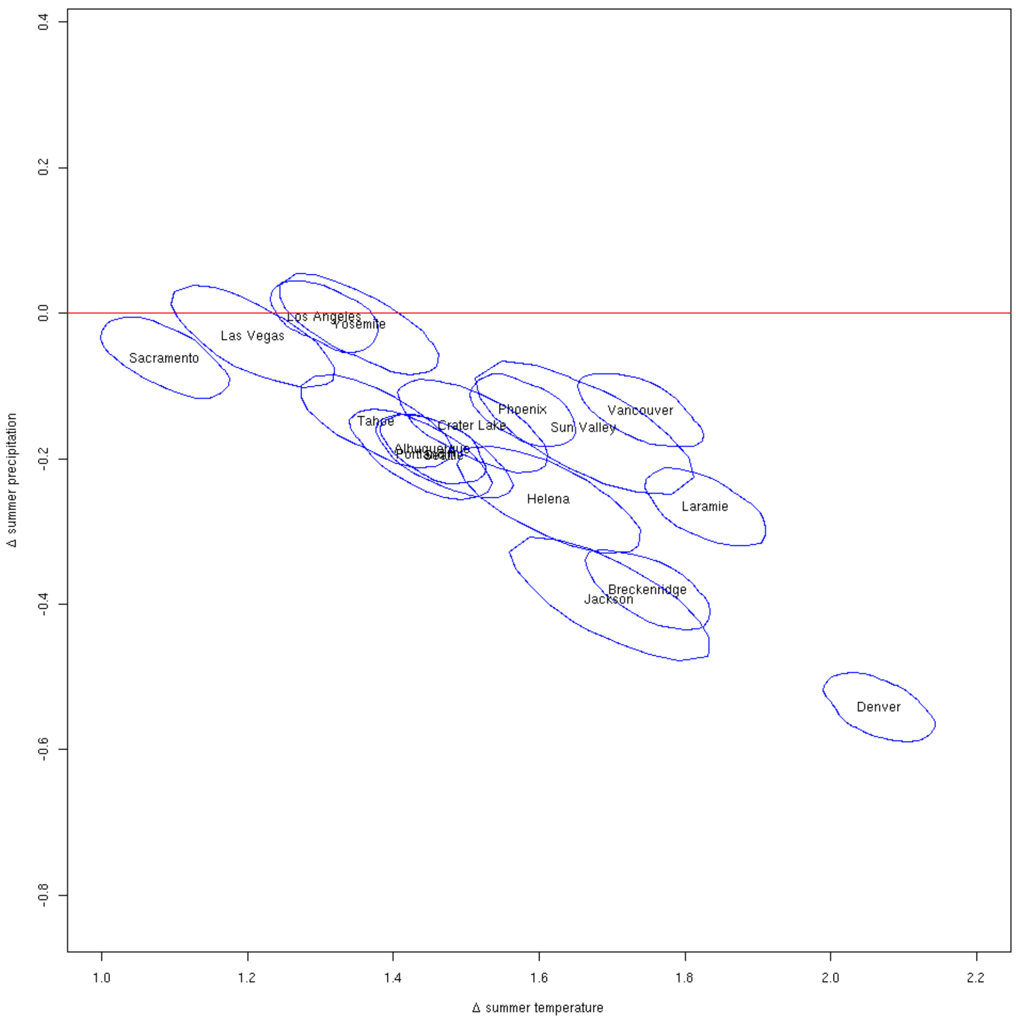


Clustering

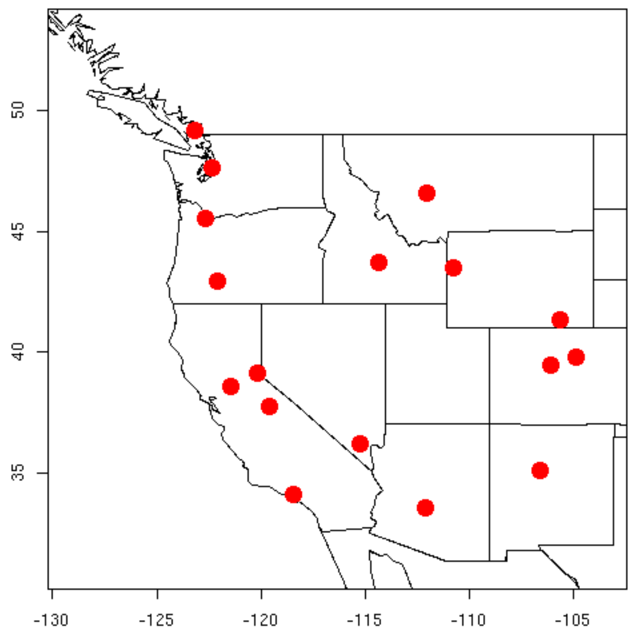

0.5
0.0
-0.5
-1.0
Δ precipitation
0.0
0.5
1.0
1.5
2.0
Δ temperature


50
45
40
35
-130
-125
-120
-115
-110
-105

Impacts



Δ summer precipitation
Δ summer temperature
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Sacramento
Las Vegas
Los Angeles
Yosemite
Tahoe
Crater Lake
Albuquerque
Phoenix
Sun Valley
Vancouver
Helena
Laramie
Jackson
Breckenridge
Denver



50
45
40
35
-130
-125
-120
-115
-110
-105

Impacts