Spatial Models: Point Referenced Data

Point Referenced Models

- Primarily two classes of models
- Bayesian Kriging
  - 2D analog to AR(1)
  - Special case of Gaussian Process model
  - Includes parameter error
- Markov Random Field
  - 2D analog to state space

The basic spatial model

$$Z(s) = \underbrace{\mu(s|\beta)}_{trend} + \underbrace{w(s|\phi)}_{spatial\ error} + \underbrace{\epsilon(s)}_{residual\ error}$$

- $\mu(s)$ = process model (e.g. $X\beta$)
- $\varepsilon \sim N(0,\tau^2)$ = nugget
- $W \sim N(0,C_\phi)$
- $C_{ij} = Cov[z_i,z_j] = f(|s_i - s_j|)$
- Assumes isotropy, second-order stationarity

Spatial Covariance Matrix

- $C_{ij} = Cov[z_i, z_j] = f(|s_i - s_j|) = \sigma^2 \rho(d_{ij} \mid \phi)$
- D = pairwise distance matrix
- Example correlation functions
  - Exponential: $\rho(d|\phi) = \exp(-d_{ij}\phi)$
  - Gaussian: $\rho(d|\phi) = \exp(-d_{ij}^2\phi)$
  - Matern: $\rho(d|\phi) = \frac{\nu^{1/2}\phi d_{ij}}{2^{\nu}\Gamma(\nu)} K_{\nu}(2\phi\nu^{1/2} d_{ij})$
- Not all functions are valid correlation functions
  - Requirement: Positive Definite covariance matrix

Distance Matrices

- Matrix of all pairwise distances
- "Distance" need not be Euclidean
- Ultimately can generalize TS and Spatial model to any situation where correlation is based on distance
  - Phylogenetic distance
  - Graphs/networks

	Banff	Calgary	Columbia Icefield	Edmonton	Field, B.C.	Jasper	Lake Louise	Radium Hot Springs	Golden	Revelstoke
Calgary	128									
Columbia Icefield	188	316								
Edmonton	423	295	461							
Field, B.C.	85	213	157	508						
Jasper	291	419	100	361	260					
Lake Louise	58	186	130	481	27	233				
Radium Hot Springs	132	260	261	555	157	361	130			
Golden	134	262	207	557	49	307	76	105		
Revelstoke	282	410	355	705	197	455	224	253	148	
Vancouver	856	984	928	1279	771	798	794	818	713	565

Distances shown are in Kilometres.
To convert to miles multiply by 0.6

Spatial Likelihood

$$Z(s) = \mu(s|\beta) + w(s|\phi) + \epsilon(s)$$

$$\Downarrow$$

$$\vec{Z} \sim N(\mu, C_{\phi} + \tau^2 \boldsymbol{I})$$

```
gp_mle <- function(parm){
  mu <- parm[1]          ## mean
  sigma <- parm[2]    ## spatial variance
  phi <- parm[3]         ## spatial correlation
  tau <- parm[4]          ## nugget variance (optional)
  C <- sigma*exp(-psi*D)   ## exponential corr
  -sum(dmvnorm(Z,rep(mu,n),C + diag(tau,n),log=TRUE))
}
```

Bayesian Spatial Model

$$\vec{Z} \sim N(\mu, C_\phi + \tau^2 \boldsymbol{I})$$

$$\mu \sim N(M_{0,} V_\mu)$$

$$\sigma^2 \sim IG(s_{1,} s_2)$$

$$\tau^2 \sim IG(t_{1,} t_2)$$

$$\phi \sim Gamma(f_{1,} f_2)$$

Bayesian Approach

```
model{
  mu ~ dnorm(0,0.01)
  sigma ~ dgamma(0.01,0.01)
  tau ~ dgamma(0.01,0.01)
  phi   ~ dunif(0,100)
  SIGMA <- inverse(1/sigma*exp(-phi*D) + 1/tau)
  Z[] ~ dmnorm(mu,SIGMA)
}
```

Bayesian Kriging:
Prediction from the Bayesian Spatial Model

- First fit spatial model
- Second compute D_{pred} = n+m pairwise distance matrix of current (n) AND prediction (m) points
- Third, for each stored MCMC iteration
  - Calculate C_{pred}

$$E[Z_{pred}] = \mu_{pred} + C_{s',s} C_{s,s}^{-1} (Z - \mu)$$

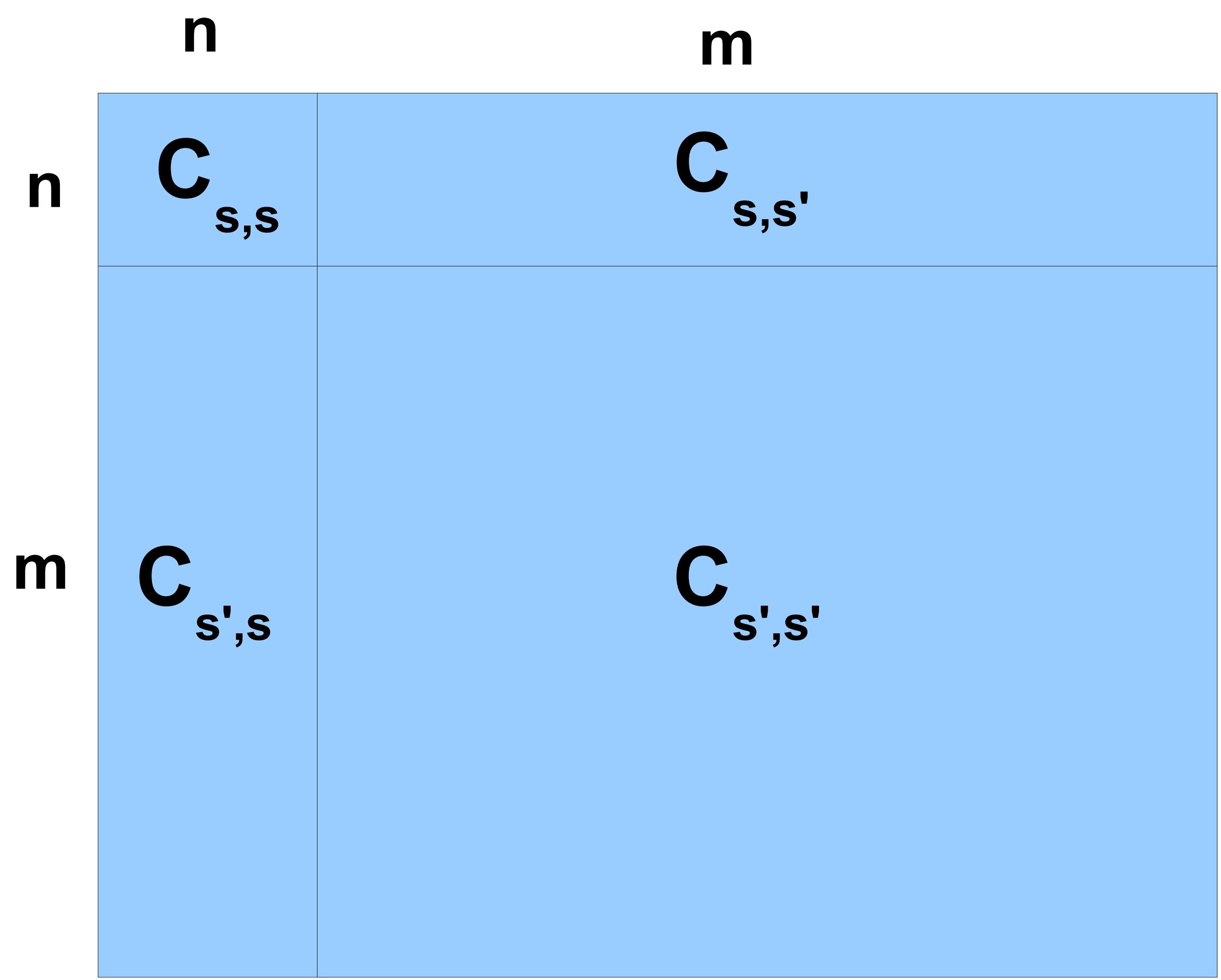
Store for CI

$$Var[Z_{pred}] = C_{s',s'} - C_{s',s} C_{s,s}^{-1} C_{s,s'} + \tau^2 \boldsymbol{I}$$

  - For PI directly draw $z_{pred} \sim N(E[z_{pred}], var[z_{pred}])$

C_{pred} submatrices

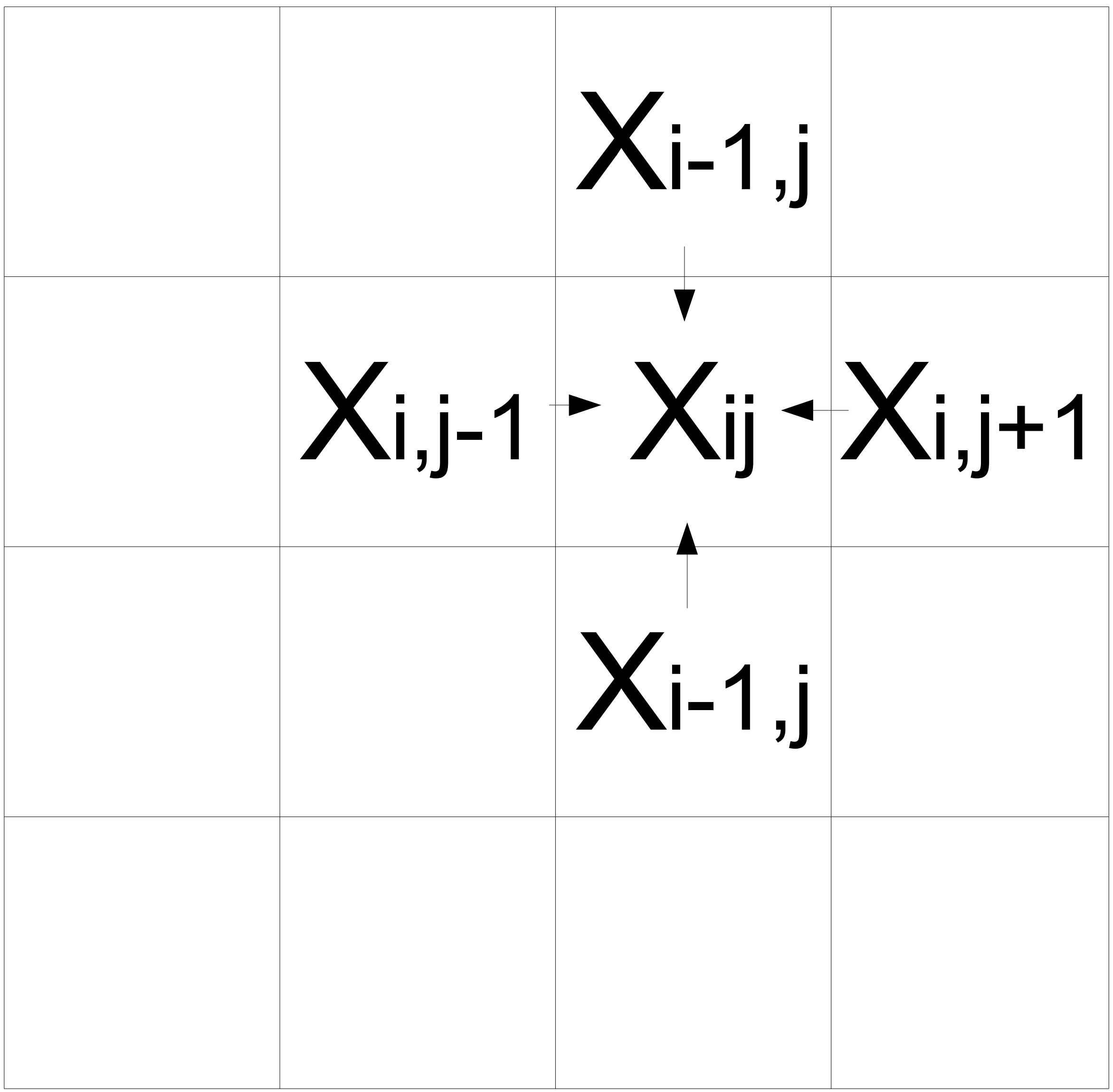
$$E[Z_{pred}] = \mu_{pred} + C_{s',s} C_{s,s}^{-1} (Z - \mu)$$



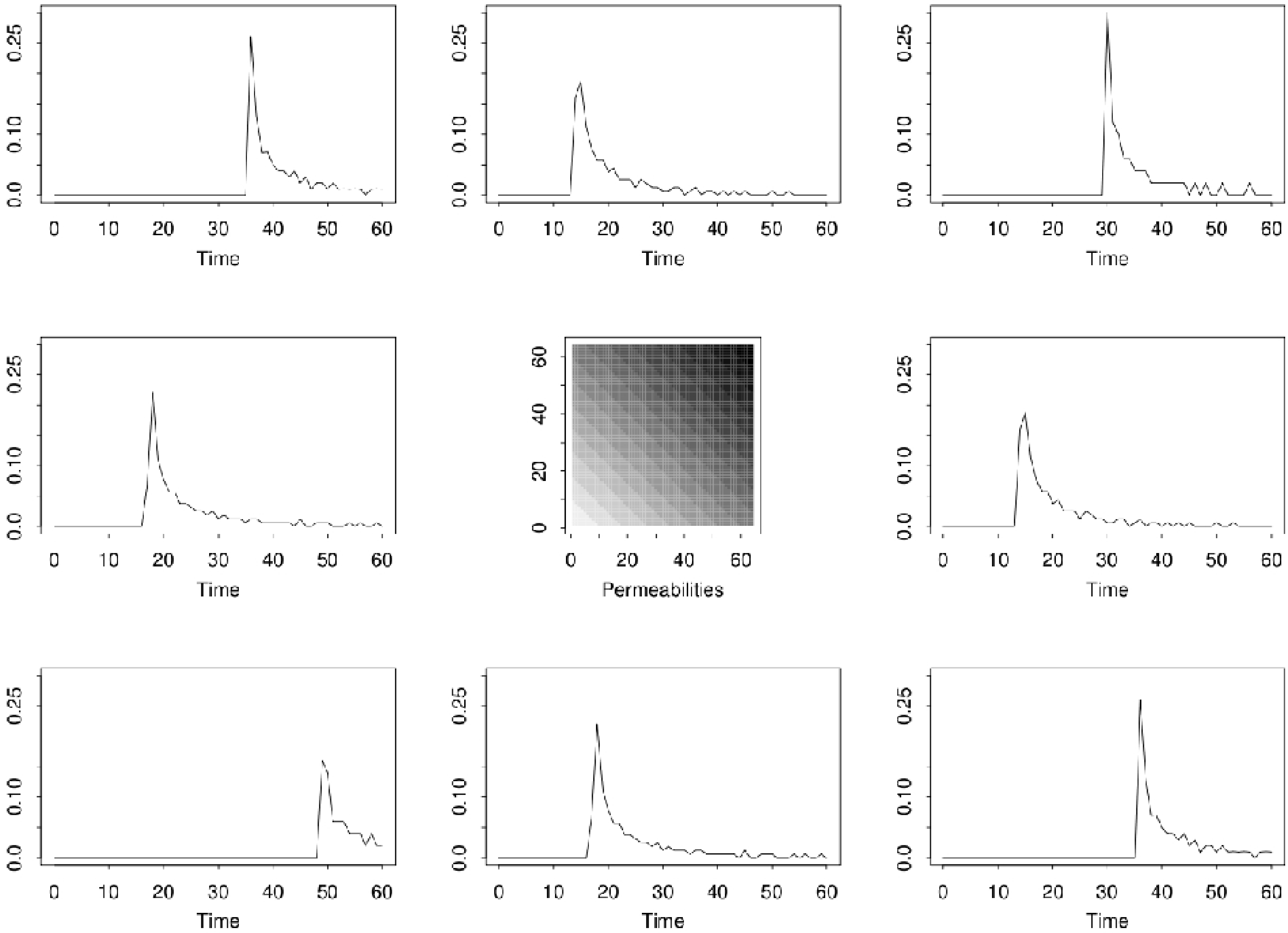
Words of Warning

- Singularities
  - There can be no off-diagonal zeros in D
  - $\min(d_{ij}) \ll \max(d_{ij})$ can cause numerical singularity
- Matrix inversion
  - Kriging requires large & computationally expensive matrix multiplications and inversions
  - Have to do for each MCMC iteration!
- Memory
  - Have to store a MAP for every MCMC iteration

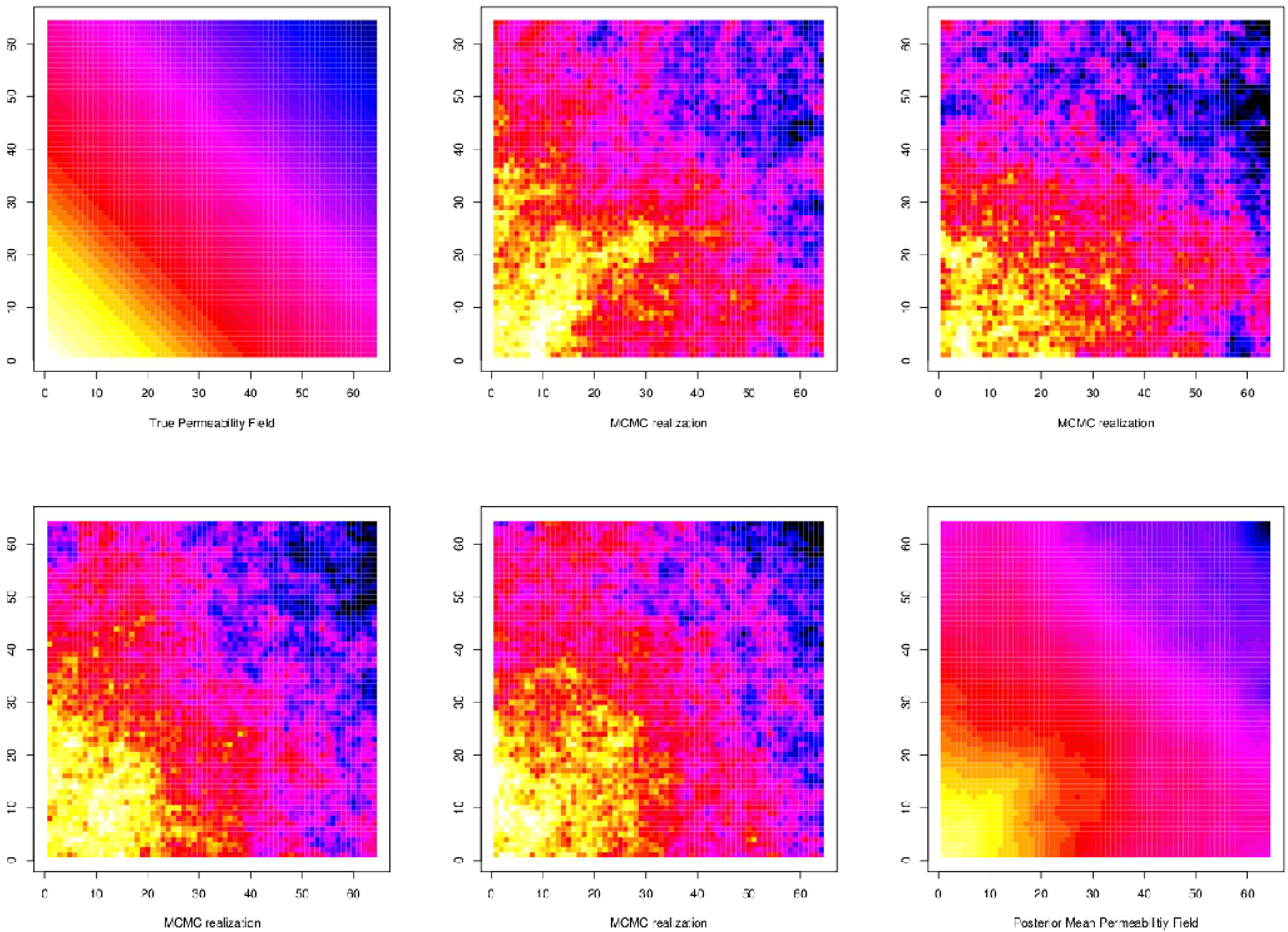
Markov Random Fields

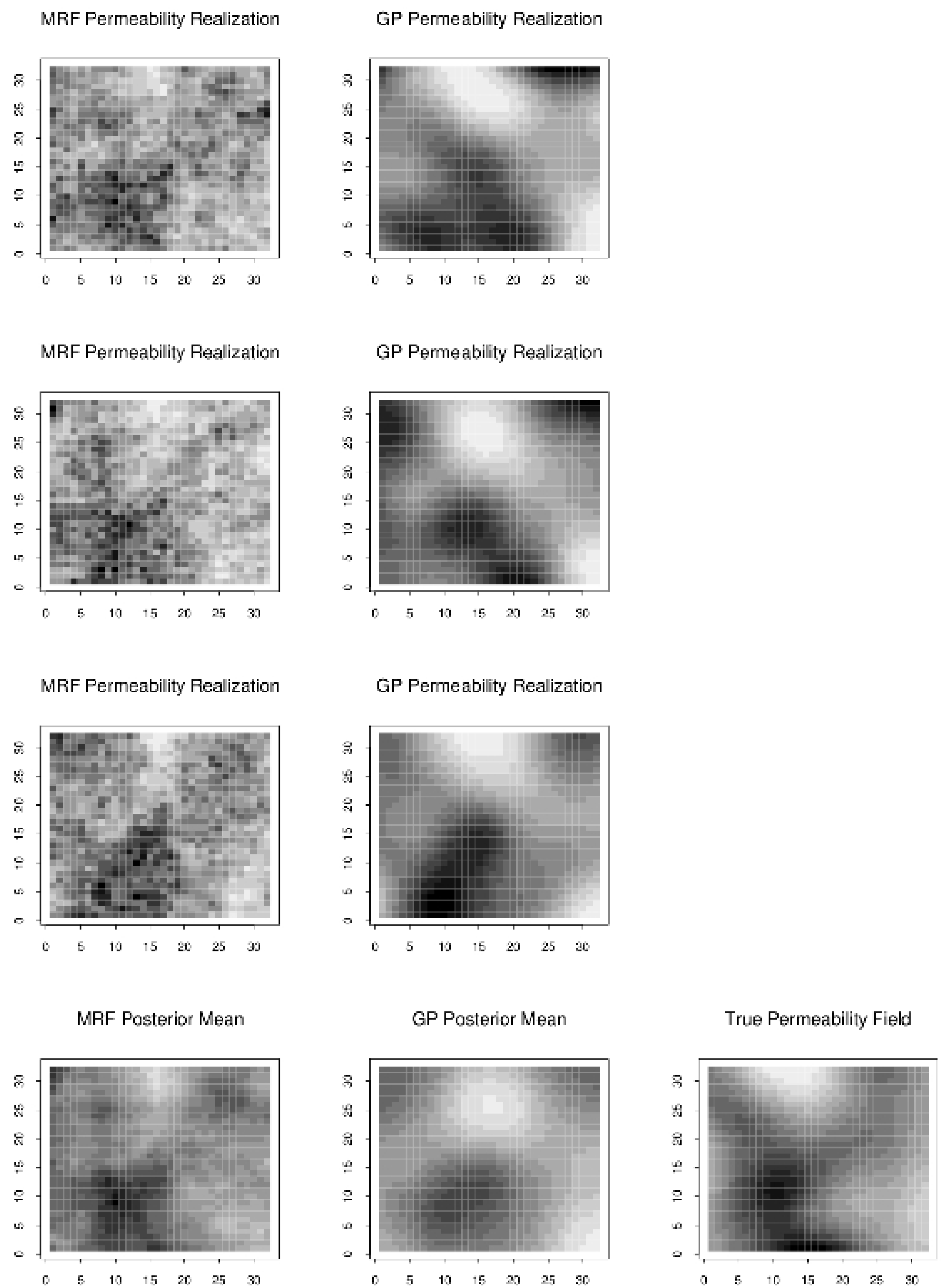


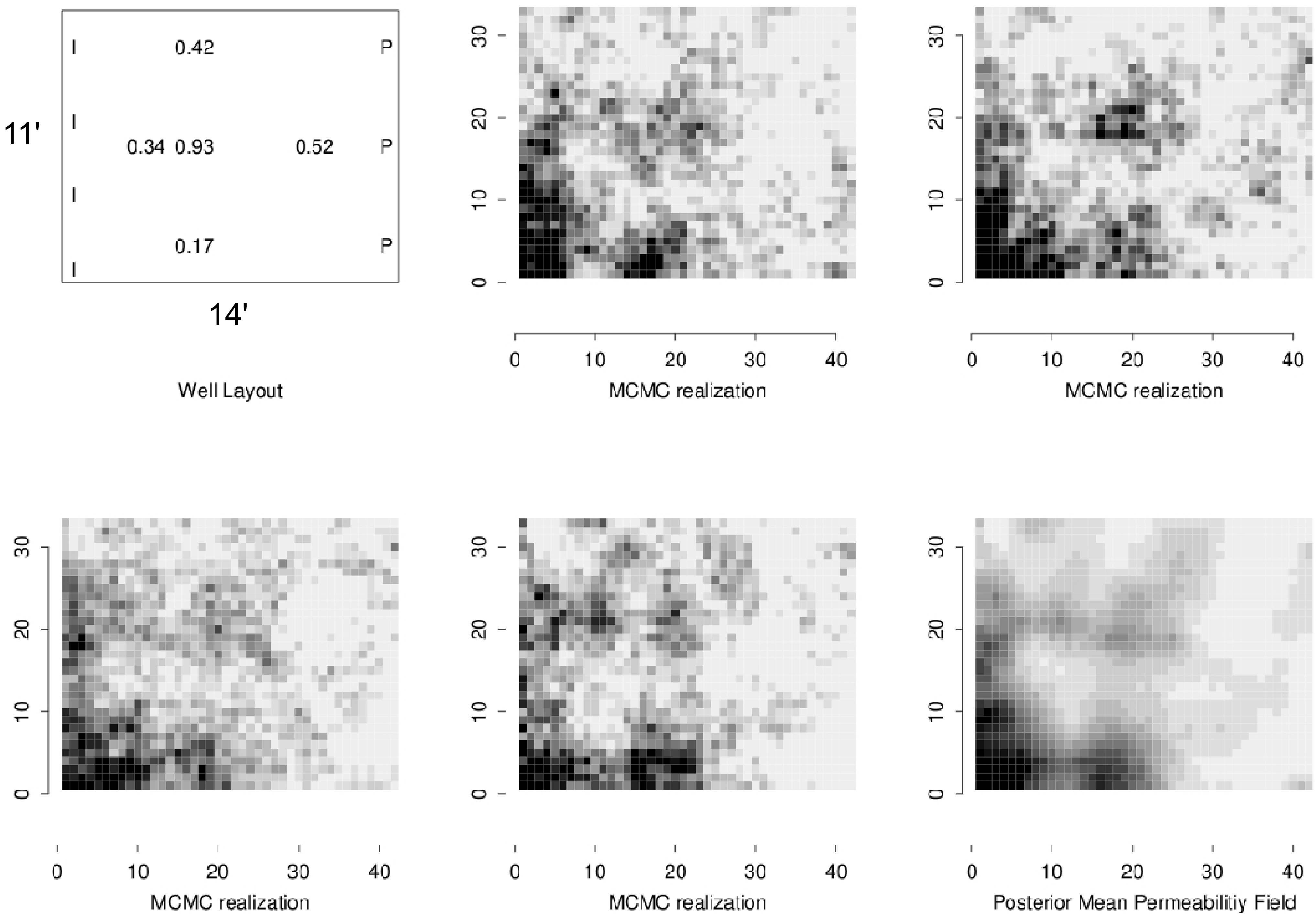
- 2D grid of latent variables
  - State-space model
- Process model connects to nearest neighbors
- Data, Y, only observed at a few points



Lee et al. 2000. MARKOV RANDOM FIELD MODELS FOR HIGH-DIMENSIONAL PARAMETERS IN SIMULATIONS OF FLUID FLOW IN POROUS MEDIA


0 10 20 30 40 50 60
True Permeability Field
MCMC realization
MCMC realization
MCMC realization
MCMC realization
Posterior Mean Permeabiltiy Field


MRF Permeability Realization
GP Permeability Realization
MRF Permeability Realization
GP Permeability Realization
MRF Permeability Realization
GP Permeability Realization
MRF Posterior Mean
GP Posterior Mean
True Permeability Field



11'
14'
I
I
I
I
P
P
P
0.42
0.34
0.93
0.52
0.17
Well Layout
MCMC realization
MCMC realization
MCMC realization
MCMC realization
Posterior Mean Permeabilitiy Field