

Spatiotemporal forecasting of plant populations and the need to partition forecast uncertainty

Andrew Tredennick

October 13, 2018

University of Georgia



COLLABORATORS

Peter Adler (USU)



Mevin Hooten (CSU)



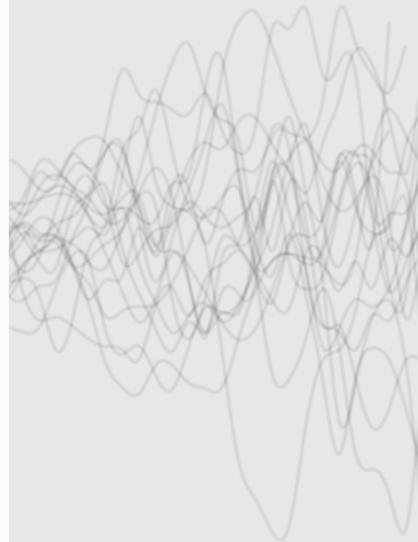


ROAD MAP

1. Plant population forecasts over large spatial extents
2. A plea (and a proposal) to partitioning forecast uncertainty

§ 1

Plant population forecasts



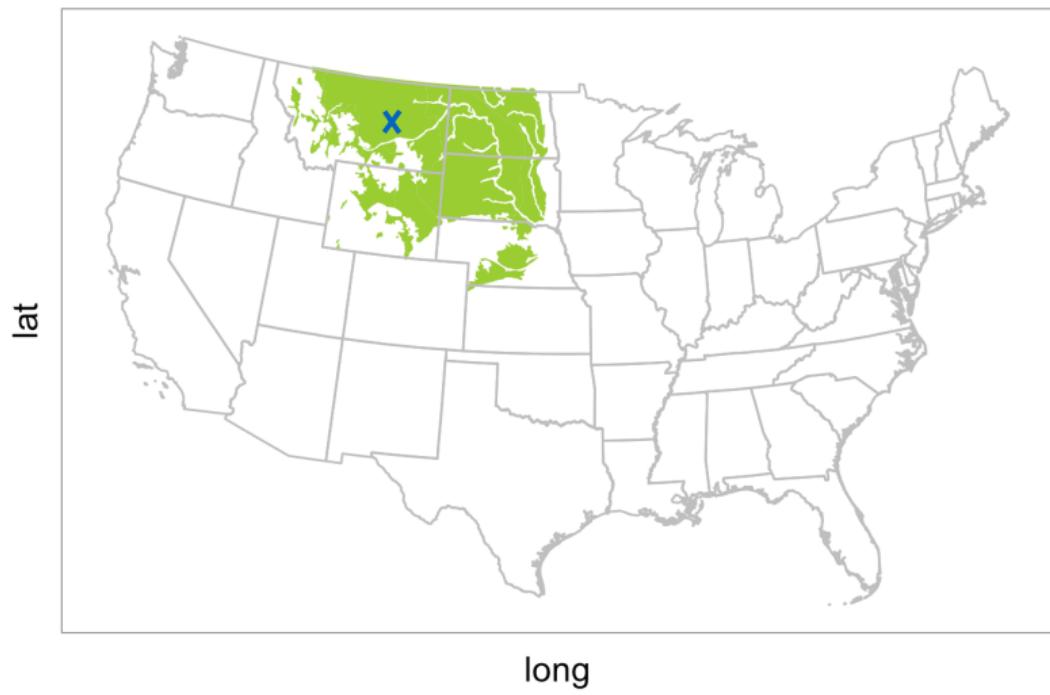


WHAT LAND MANAGERS WANT





WHAT LAND MANAGERS GET



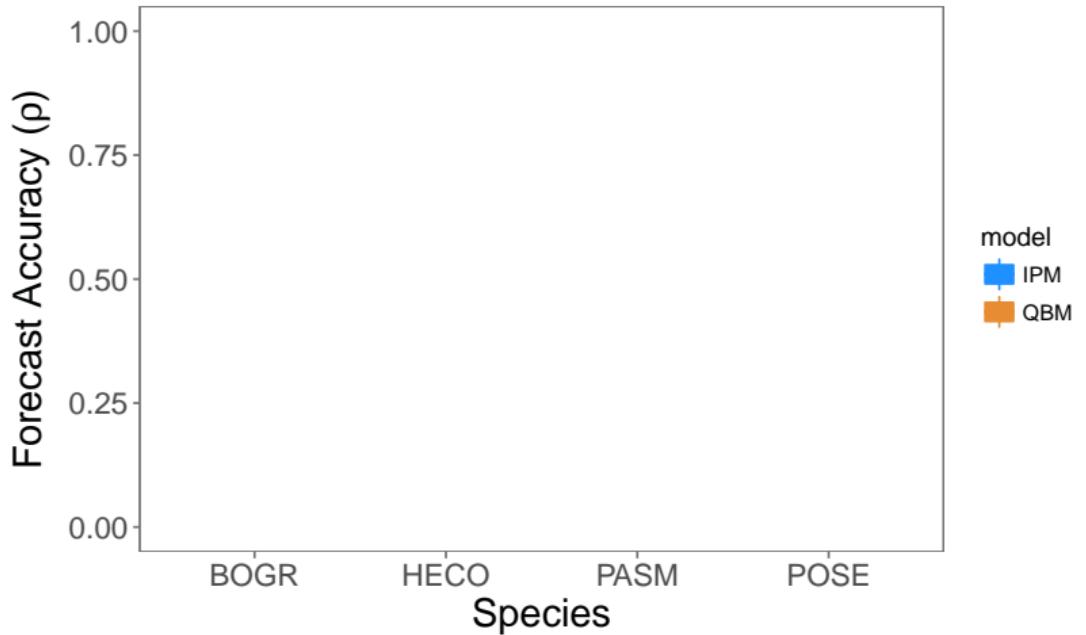


REALLY HARD WORK



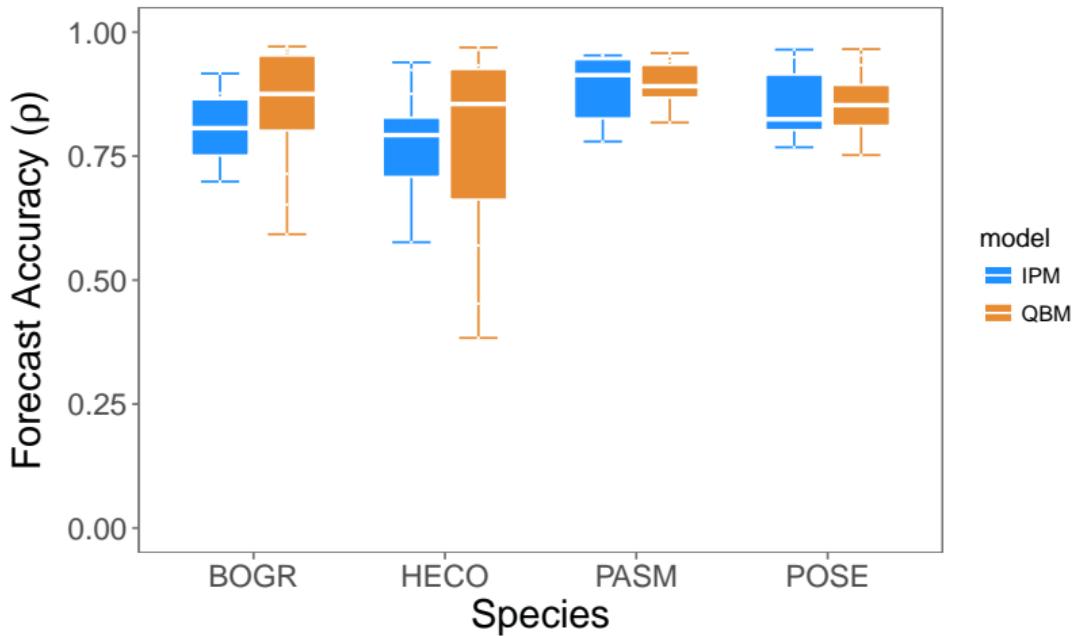


FREE FROM THE TYRANNY OF DEMOGRAPHIC DATA!



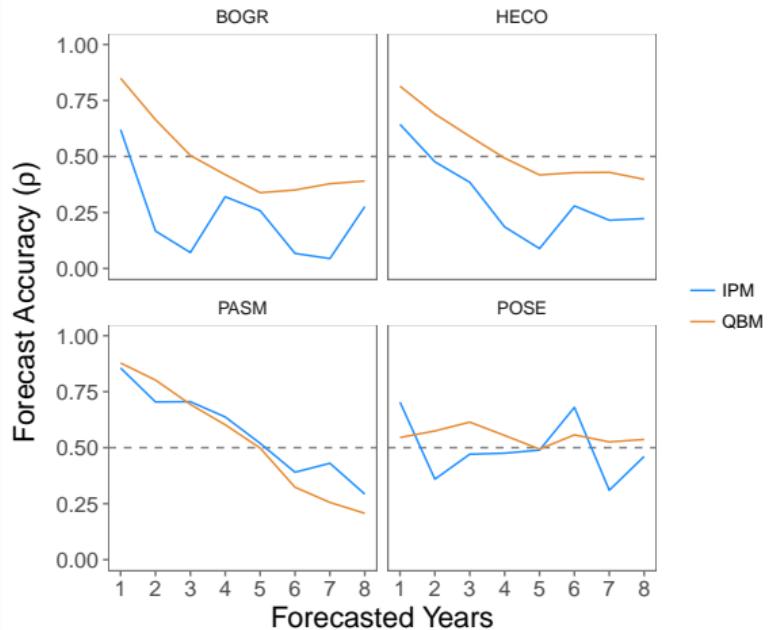


FREE FROM THE TYRANNY OF DEMOGRAPHIC DATA!





FORECAST HORIZONS



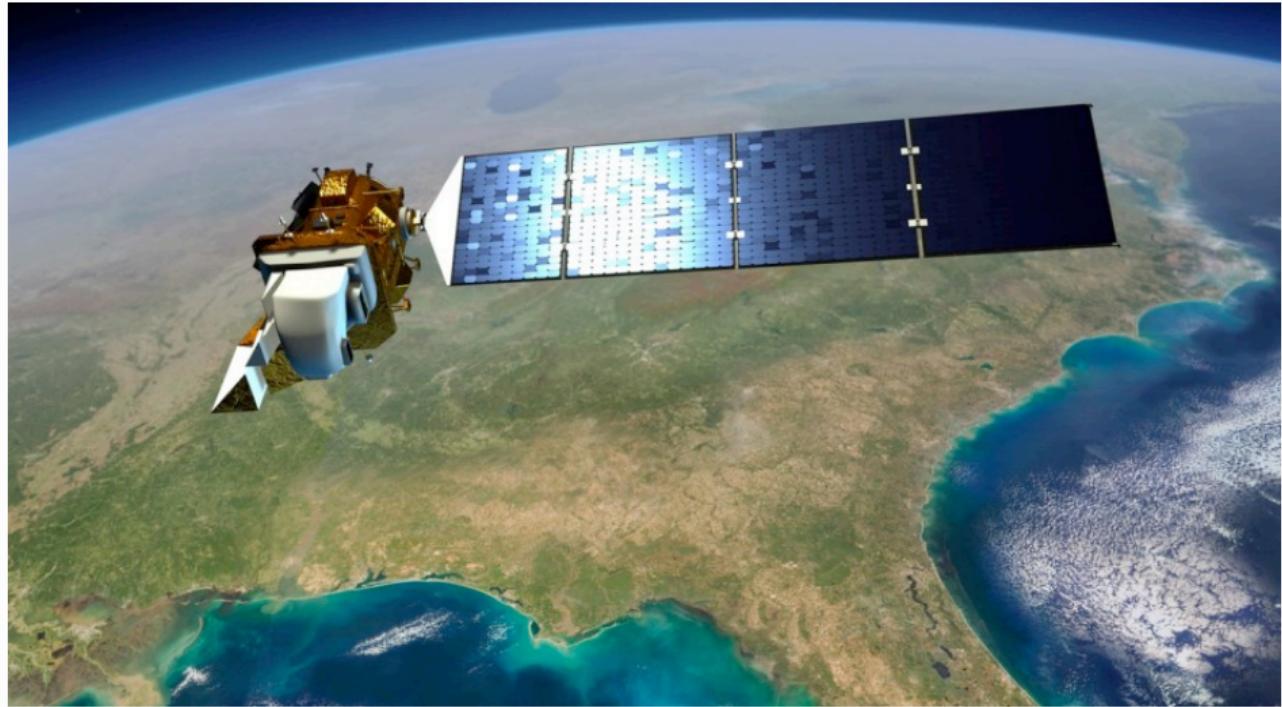


DO WE NEED DEMOGRAPHIC DATA?

Maybe not.



LET'S USE SATELLITES



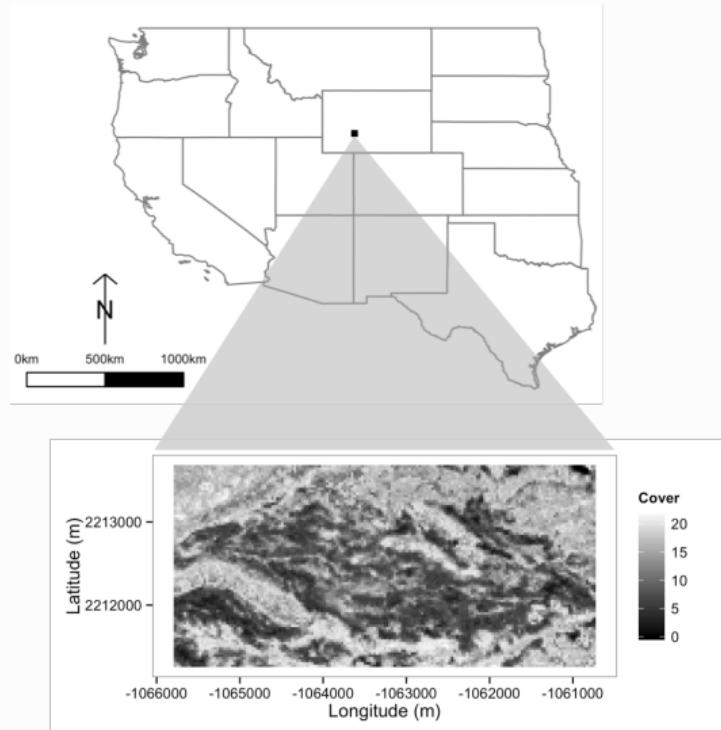


SAGEBRUSH SEA IN WYOMING



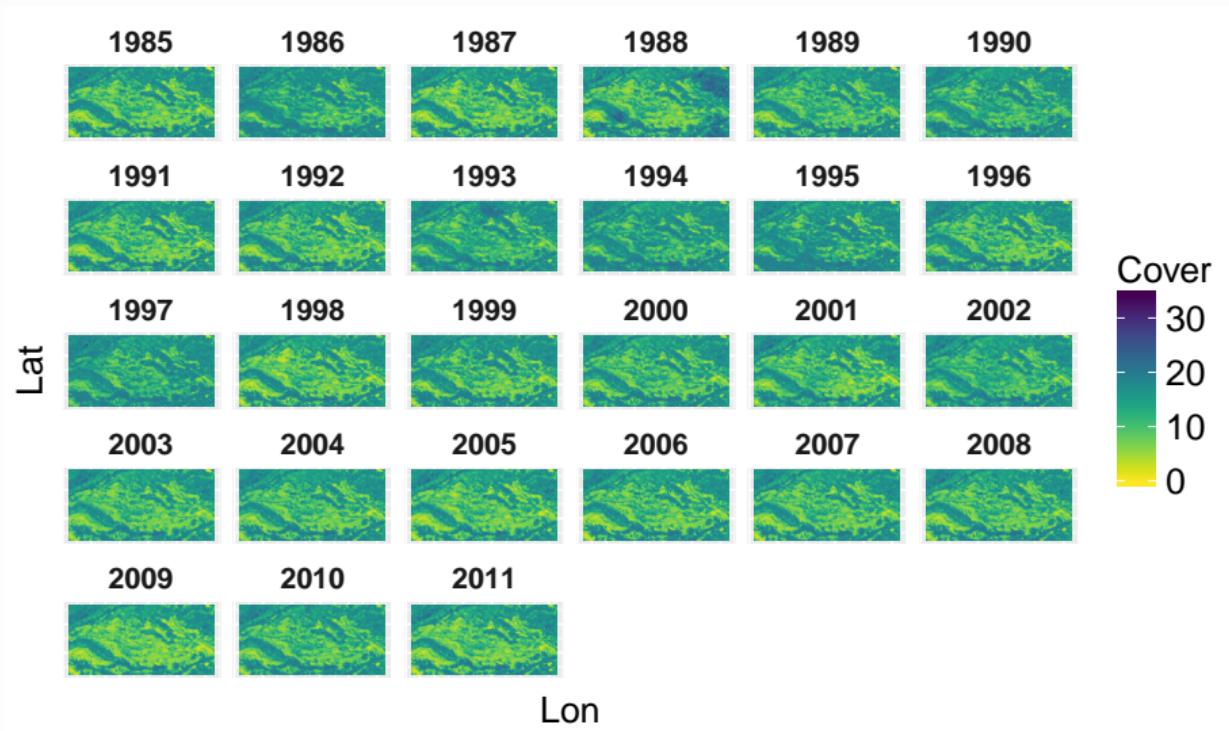


STUDY AREA





LANDSAT TIME SERIES





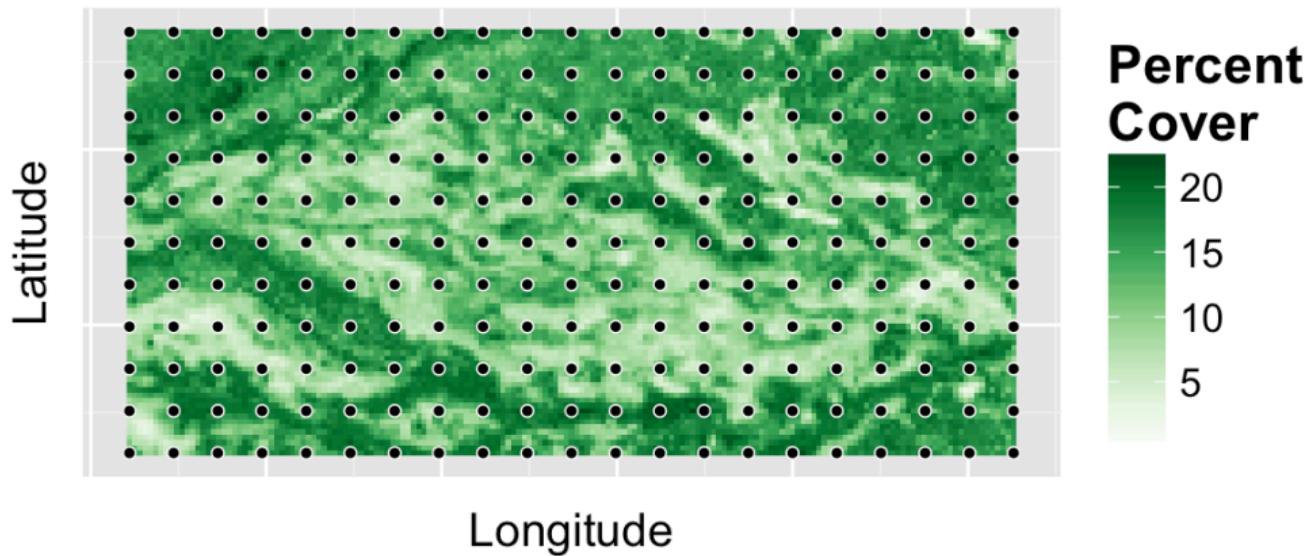
DYNAMIC COVER MODEL

$$y_{i,t} \sim \text{Poisson}(\mu_{i,t})$$

$$\log(\mu_{i,t}) = \underbrace{\beta_{0,t} + \beta_1 y_{i,t-1}}_{\text{temporal + dens. dep}} + \underbrace{\mathbf{x}'_t \gamma}_{\text{climate}} + \underbrace{\eta_i}_{\text{spatial}}$$



DIMENSION REDUCTION FOR SPATIAL EFFECT





DYNAMIC COVER MODEL

$$y_{i,t} \sim \text{Poisson}(\mu_{i,t})$$

$$\log(\mu_{i,t}) = \underbrace{\beta_{0,t} + \beta_1 y_{i,t-1}}_{\text{temporal + dens. dep.}} + \underbrace{\mathbf{x}'_t \gamma}_{\text{climate}} + \underbrace{\eta_i}_{\text{spatial}}$$

$$\eta \approx \mathbf{K}\alpha,$$

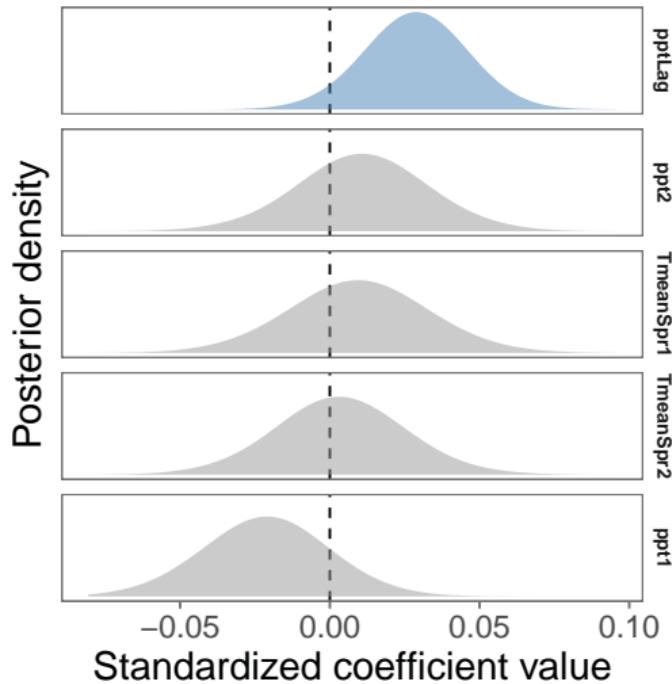
$$\mathbf{K} = \mathbf{w}_{s,m} / \sum_{s=1}^S \mathbf{w}_{s,m}$$

$$\mathbf{w}_{s,m} = \exp(-d_{s,m}/\sigma)$$

$$\alpha_m \sim \text{Normal}(0, \sigma_\eta^2)$$



CLIMATE EFFECTS

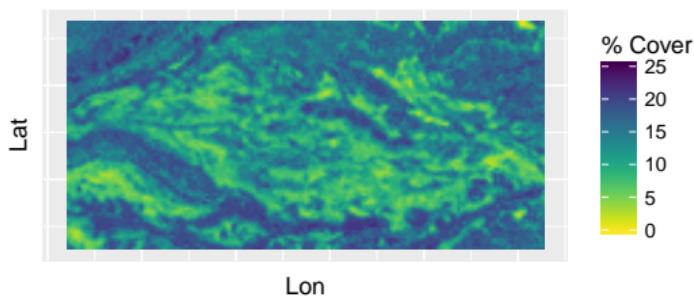




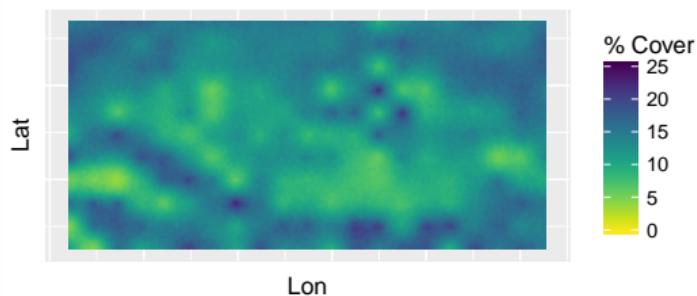
MODEL PERFORMANCE

RMSE $\approx 4\%$

Observed cover

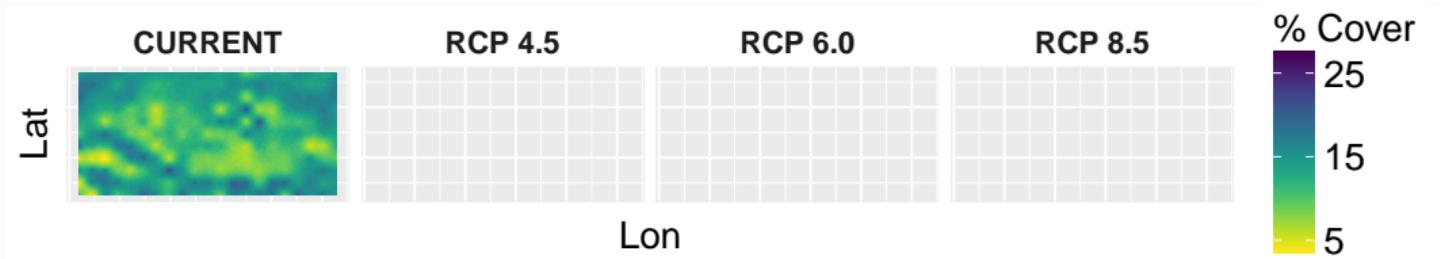


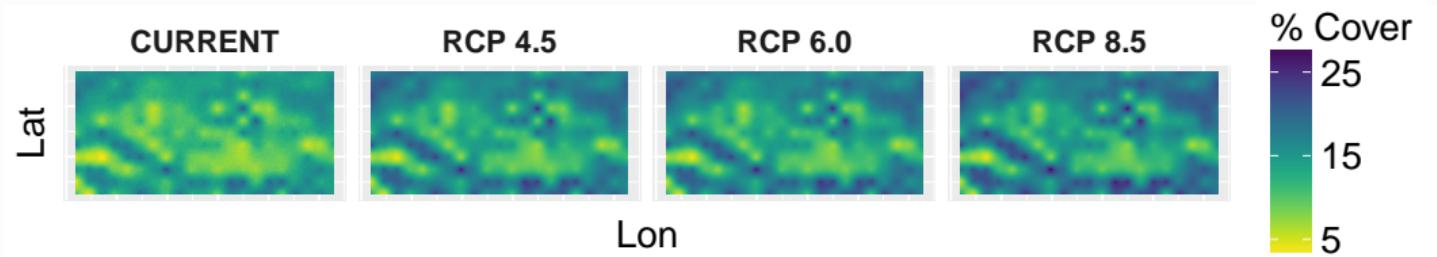
Predicted cover





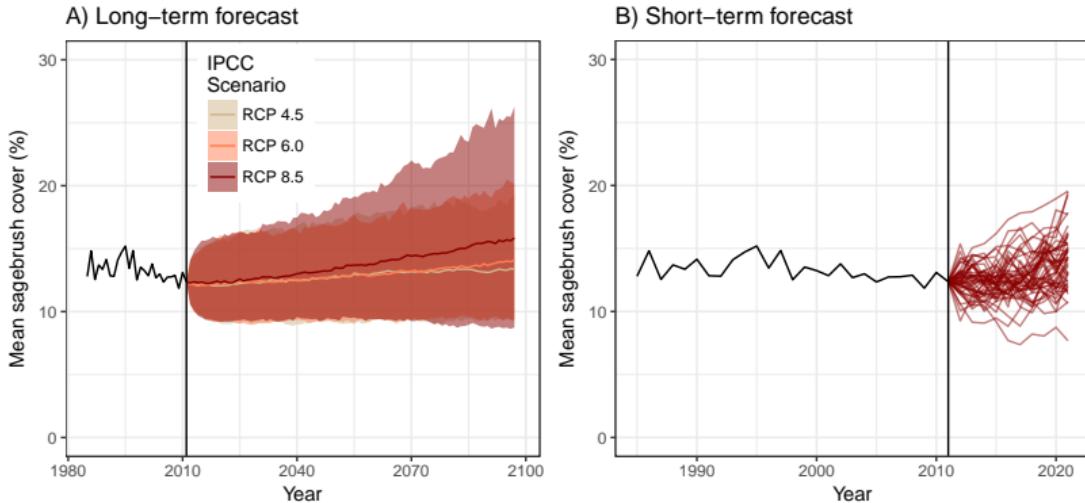
FORECASTS UNDER CLIMATE CHANGE: SPATIAL



 FORECASTS UNDER CLIMATE CHANGE: SPATIAL

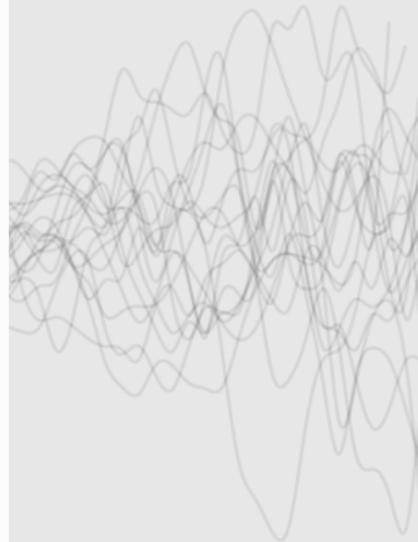


FORECASTS UNDER CLIMATE CHANGE: TEMPORAL



§ 2

Partitioning forecast uncertainty





AN (INVERSE) ERROR PROPAGATION PROBLEM

For some function q : $q = f(x_1, x_2, \dots, x_n)$

$$\begin{aligned}\sigma_q^2 &= \left(\frac{\delta q}{\delta x_1} \sigma_{x_1} \right)^2 + \left(\frac{\delta q}{\delta x_2} \sigma_{x_2} \right)^2 + \cdots + \left(\frac{\delta q}{\delta x_n} \sigma_{x_n} \right)^2 \\ &= \sum_{i=1}^n \left(\frac{\delta q}{\delta x_i} \sigma_{x_i} \right)^2\end{aligned}$$



AN (INVERSE) ERROR PROPAGATION PROBLEM

$$\sigma_q^2 = \underbrace{\sum_{i=1}^n \left(\frac{\delta q}{\delta x_i} \sigma_{x_i} \right)^2}_{\text{variances}} + \underbrace{\sum_{i=1}^n \sum_{j(j \neq i)}^n 2\sigma_{ij} \left(\frac{\delta q}{\delta x_i} \right) \left(\frac{\delta q}{\delta x_j} \right)}_{\text{covariances}}$$



RECAST AS FORECAST UNCERTAINTY

Forecast of state z at $t + 1$ from function q : $q = f(z_t, x_t, \theta, \varepsilon_{t+1})$

$$\text{Var}[y_{t+1}] \approx \underbrace{\left(\frac{\delta q}{\delta y} \right)^2}_{\text{stability}} \text{Var}[y_t] + \underbrace{\left(\frac{\delta f}{\delta x} \right)^2}_{\text{driver sens.}} \underbrace{\text{Var}[x_t]}_{\text{driver uncert.}} + \underbrace{\left(\frac{\delta f}{\delta \theta} \right)^2}_{\text{param sens.}} \underbrace{\text{Var}[\theta]}_{\text{param. uncert.}} + \underbrace{\text{Var}[\varepsilon_{t+1}]}_{\text{process error}}$$



INTERACTION (COVARIANCES) CANNOT BE IGNORED

$$z_{t+1} = z_t \beta_0 + x_t \beta_1 + \varepsilon_{t+1},$$

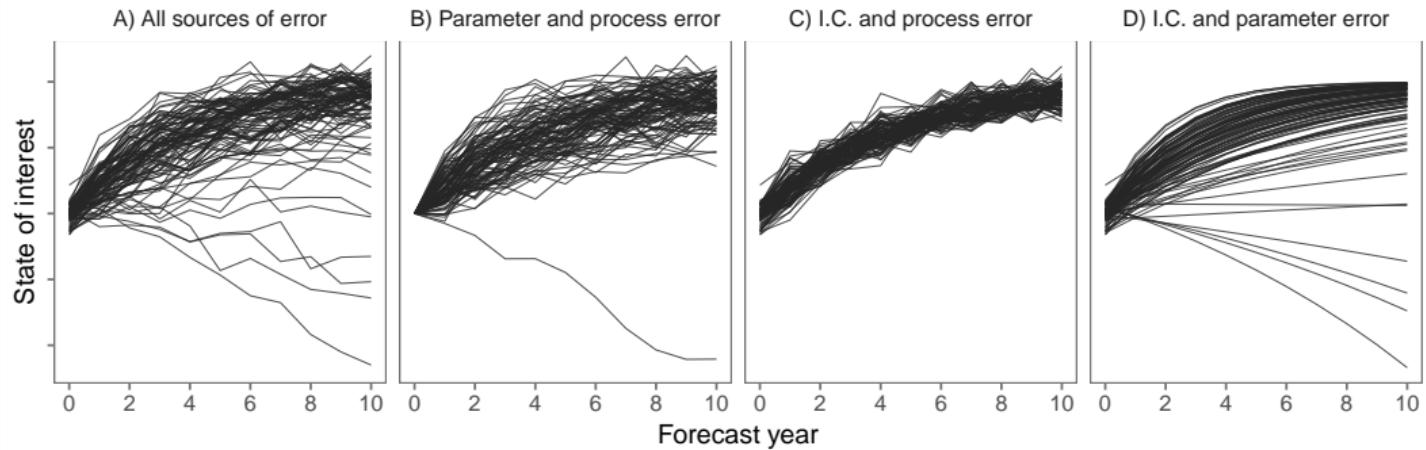
$$z_{t=1} \sim \text{Normal}(z_0, \sigma_{\text{init.}}^2), \quad \text{initial conditions uncertainty}$$

$$\boldsymbol{\beta} \sim \text{MVN}(0, \sigma_{\text{param.}}^2 \mathbf{I}), \quad \text{parameter uncertainty}$$

$$\varepsilon_t \sim \text{Normal}(0, \sigma_{\text{proc.}}^2) \quad \text{process uncertainty}$$



INTERACTION (COVARIANCES) CANNOT BE IGNORED





HIERARCHICAL BAYESIAN MODELS PROPAGATE UNCERTAINTY FOR US

Data Model: $y_t \sim [y_t | z_t, \sigma_o^2], \quad t = 1, \dots, T,$

Process Models: $z_t \sim [z_t | \mu_t, \sigma_p^2],$

$\mu_t = g(z_{t-1}, \mathbf{x}'_t, \theta), \quad t = 2, \dots, T,$

Parameter Models: $\varphi \sim [\theta, \sigma_p^2, \sigma_o^2, z_{t=1}],$



THE FORECAST DISTRIBUTION

$$\begin{aligned} [z_{T+1}|y_1, \dots, y_T] &= \int \int \dots \int [z_{T+1}|z_T, \mathbf{x}_T, \theta, \sigma_p^2] \\ &\times [z_1, \dots, z_{T+1}, \theta, \sigma_p^2|y_1, \dots, y_T] d\theta d\sigma_p^2 dz_1 \dots dz_T. \end{aligned} \tag{1}$$

 THE FORECAST DISTRIBUTION, VIA MCMC

We have:

- $k = 1, \dots, K$ MCMC iterations
- $j = 1, \dots, J$ realizations of the covariate, resampled to match K
- Forecasts at times $T + q, \dots, T + Q$

$$z_{T+q}^{(k)} \sim \left[z_{T+q} | g(z_{T+q-1}^{(k)}, \mathbf{x}_{T+q}^{(j(k))}, \theta^{(k)}), \sigma_p^{2(k)} \right]$$



PARTITION FORECAST UNCERTAINTY

ANOVA

$$\begin{aligned} V_{T+q}^{(F)} = & V_{T+q}^{(I)} + V_{T+q}^{(PA)} + V_{T+q}^{(D)} + V_{T+q}^{(PS)} \\ & + \varepsilon_{T+q}^{(I,PA)} + \varepsilon_{T+q}^{(I,D)} + \varepsilon_{T+q}^{(I,PS)} + \varepsilon_{T+q}^{(PA,PS)} + \varepsilon_{T+q}^{(PA,D)} + \varepsilon_{T+q}^{(D,PS)} \\ & + \varepsilon_{T+q}^{(I,PA,D)} + \varepsilon_{T+q}^{(I,PA,PS)} + \varepsilon_{T+q}^{(I,D,PS)} + \varepsilon_{T+q}^{(PA,D,PS)} \\ & + \varepsilon_{T+q}^{(I,PA,D,PS)} \end{aligned}$$

