

Spatiotemporal forecasting of plant populations and a proposal to partition forecast uncertainty

Andrew Tredennick

October 12, 2018

University of Georgia



COLLABORATORS

Peter Adler (USU)



Mevin Hooten (CSU)



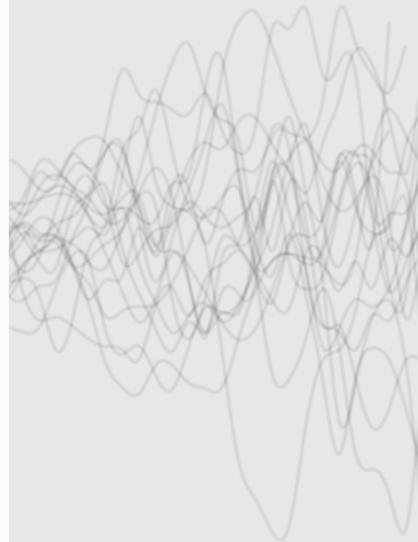


ROAD MAP

1. Plant population forecasts over large spatial extents
2. A proposal for partitioning forecast uncertainty

§ 1

Plant population forecasts



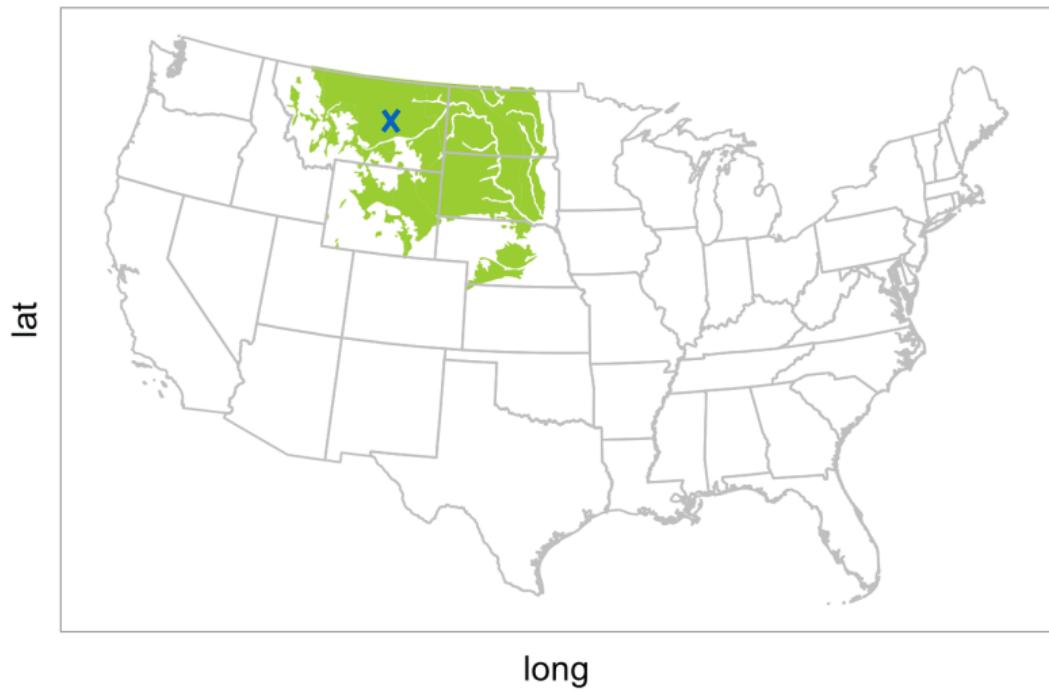


WHAT LAND MANAGERS WANT





WHAT LAND MANAGERS GET



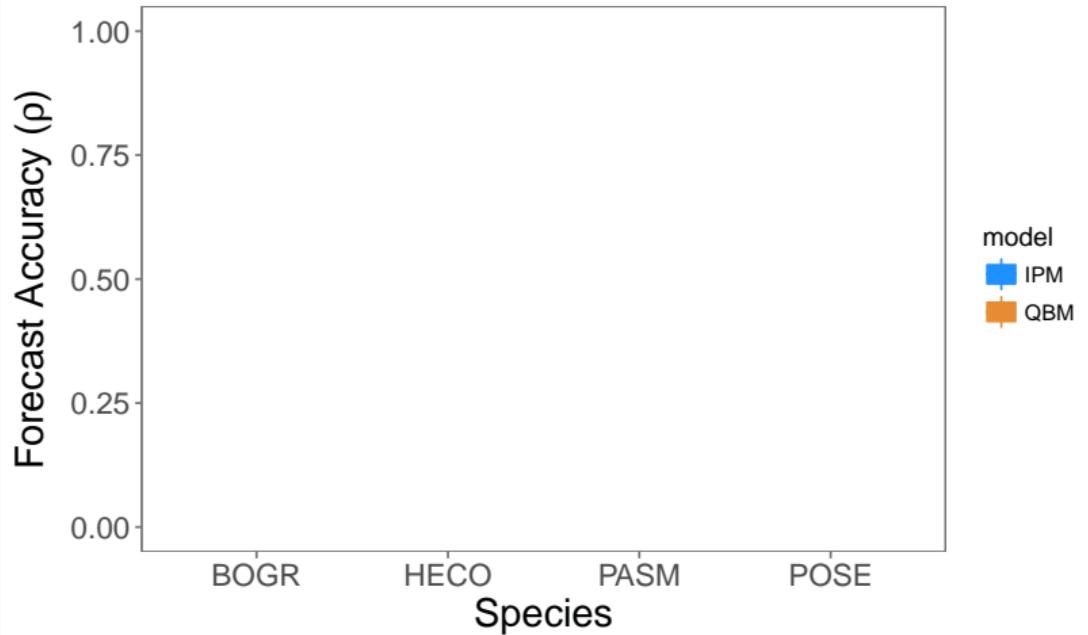


REALLY HARD WORK



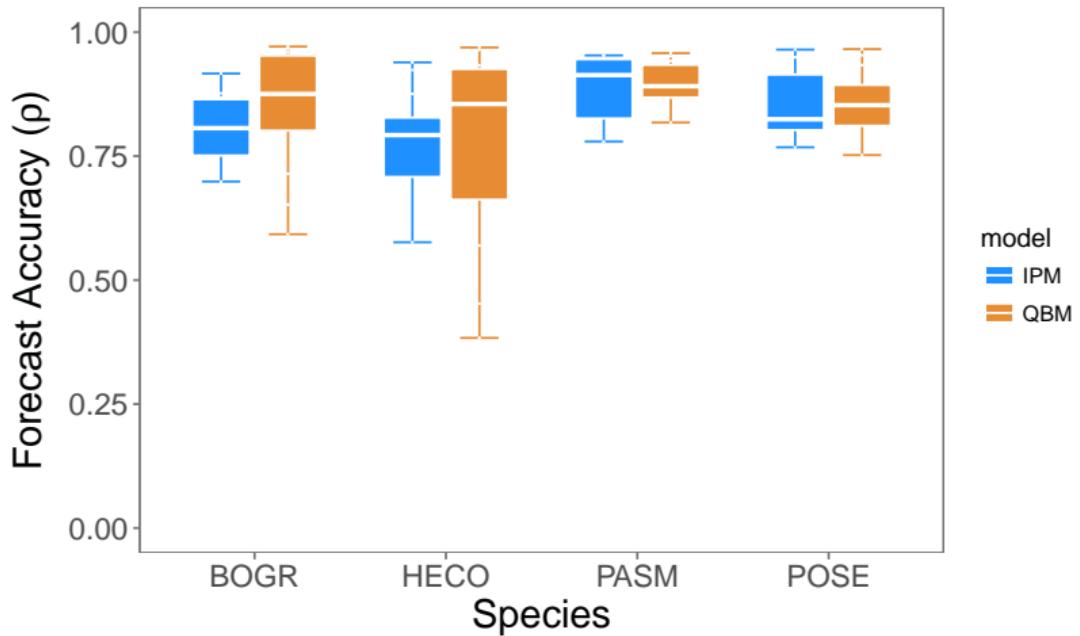


DO WE NEED DEMOGRAPHIC DATA?



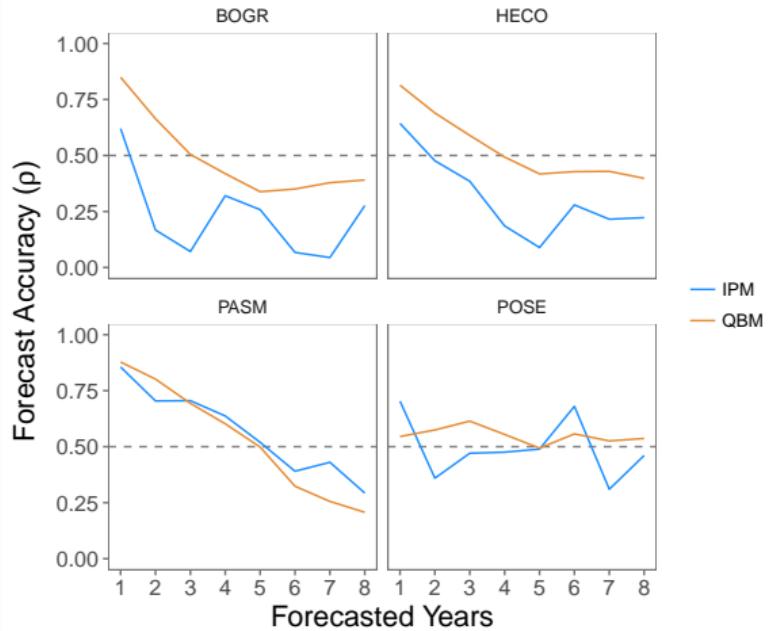


FREE FROM THE TYRANNY OF DEMOGRAPHIC DATA!





FORECAST HORIZONS



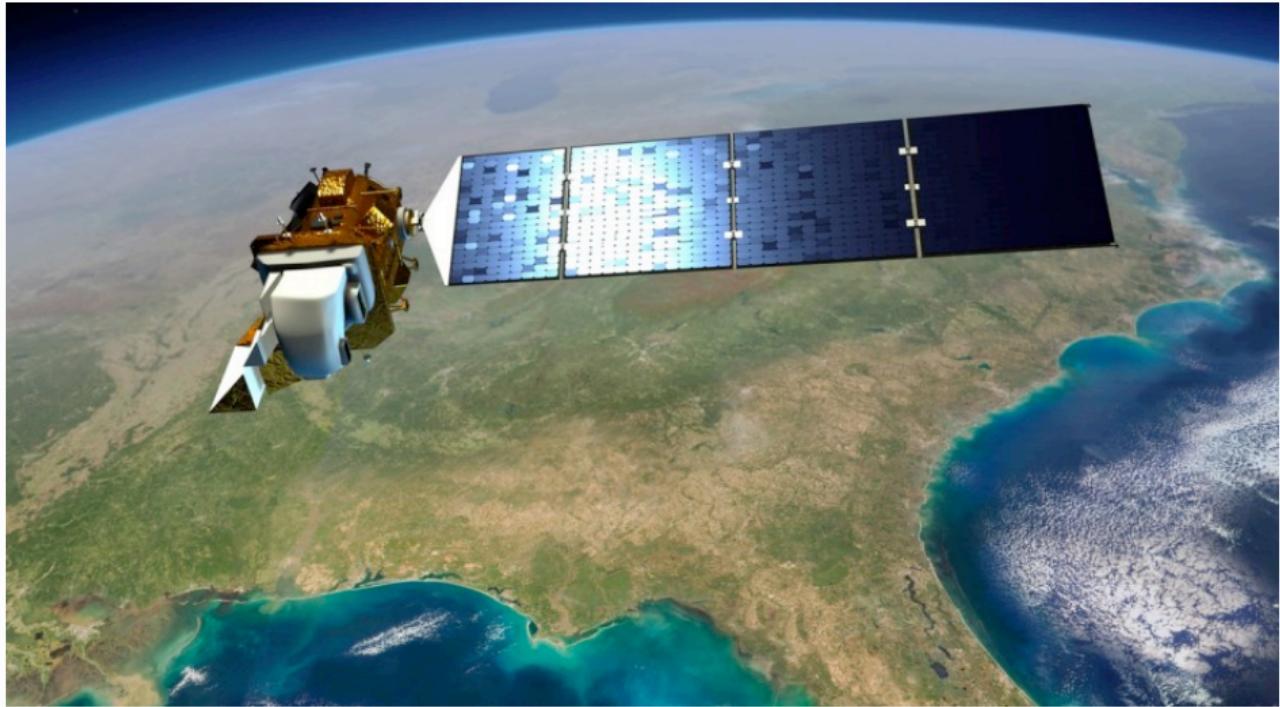


DO WE NEED DEMOGRAPHIC DATA?

Maybe not.



LET'S USE SATELLITES



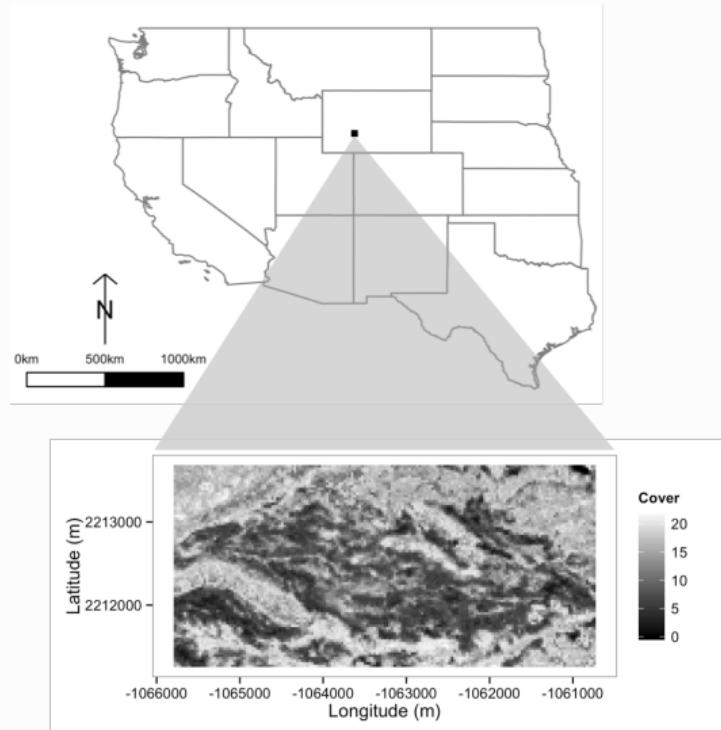


SAGEBRUSH SEA IN WYOMING



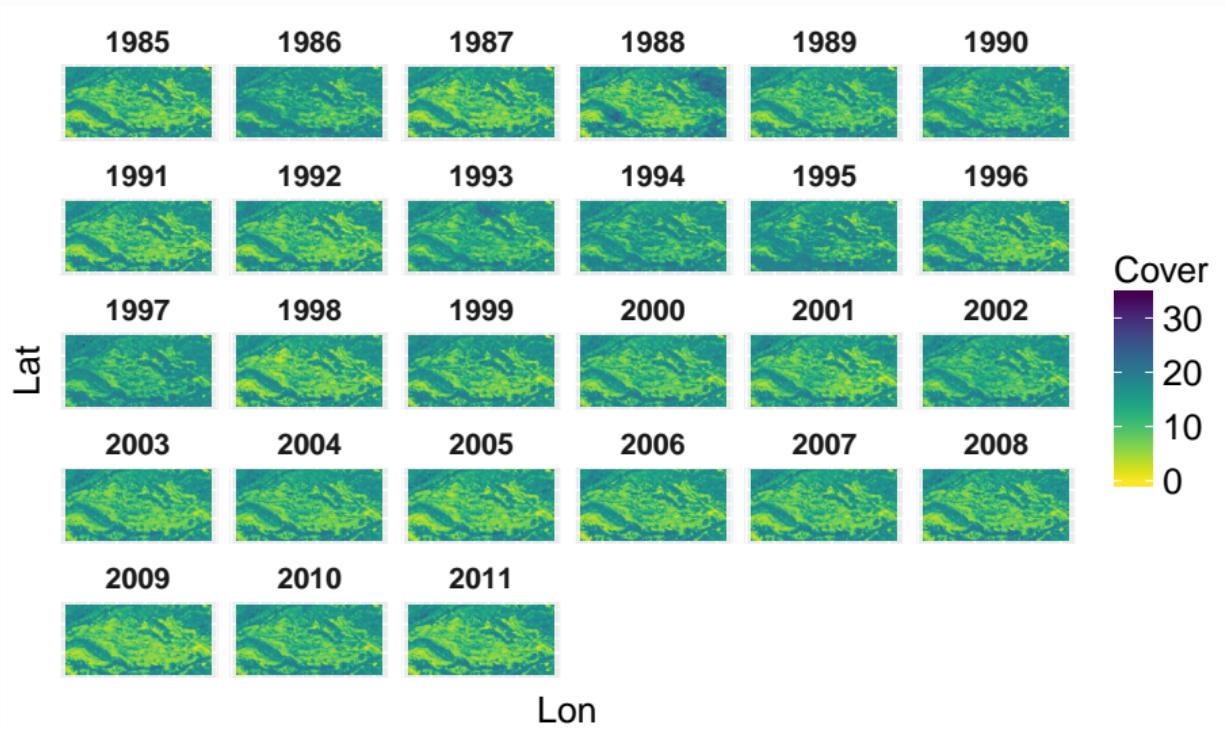


STUDY AREA





LANDSAT TIME SERIES





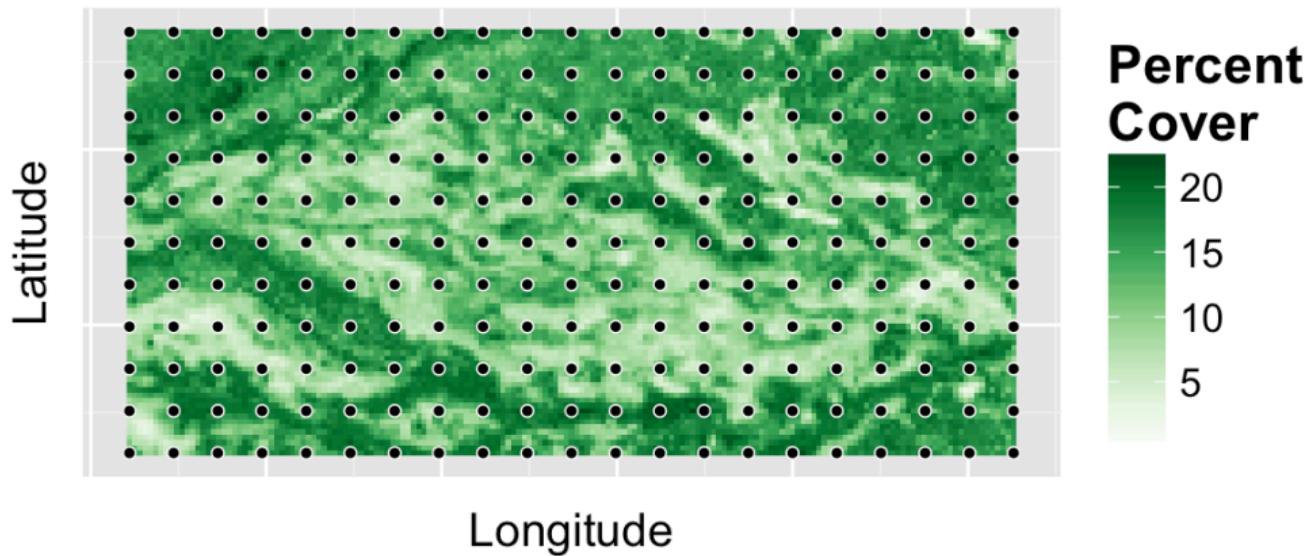
DYNAMIC COVER MODEL

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

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



DIMENSION REDUCTION FOR SPATIAL EFFECT





DYNAMIC COVER MODEL

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

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

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

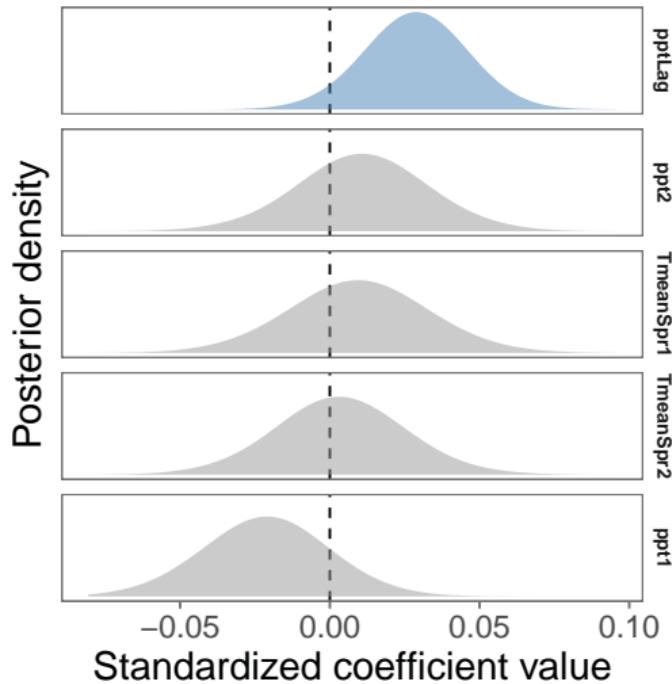
$$\mathbf{K} = \mathbf{w}_{s,m} / \sum_{s=1}^S \mathbf{w}_{s,m}$$

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

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



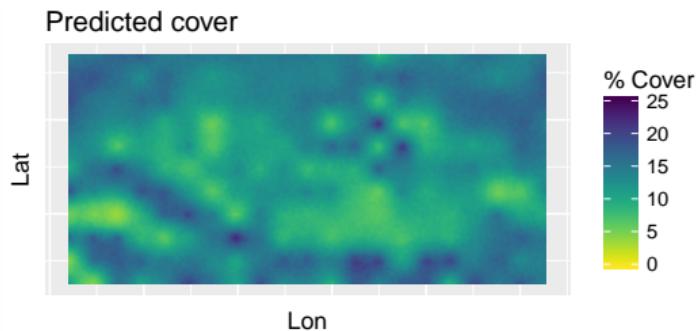
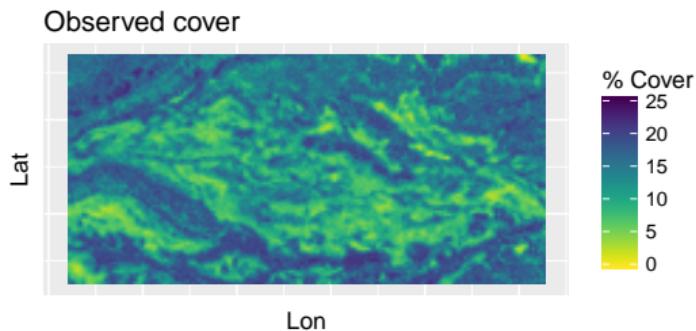
CLIMATE EFFECTS





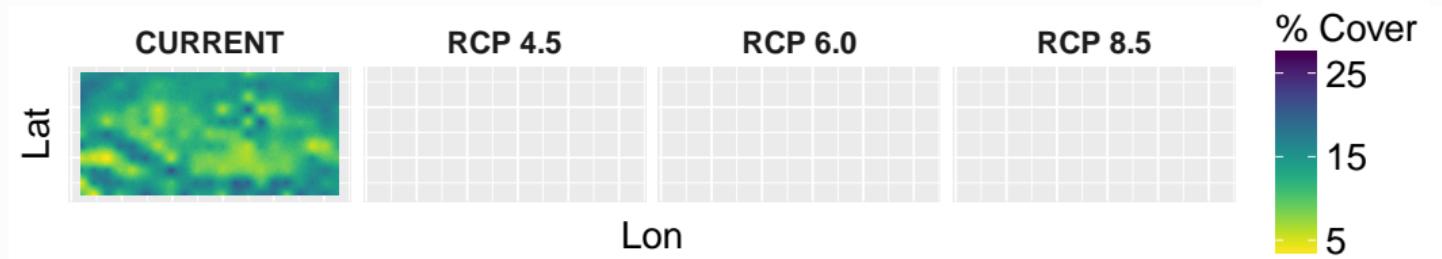
MODEL PERFORMANCE

In-sample RMSE $\approx 4\%$



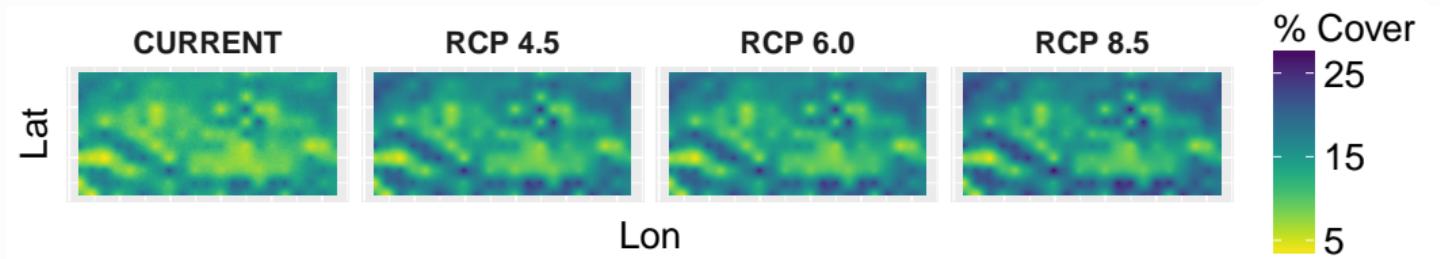


FORECASTS UNDER CLIMATE CHANGE: SPATIAL



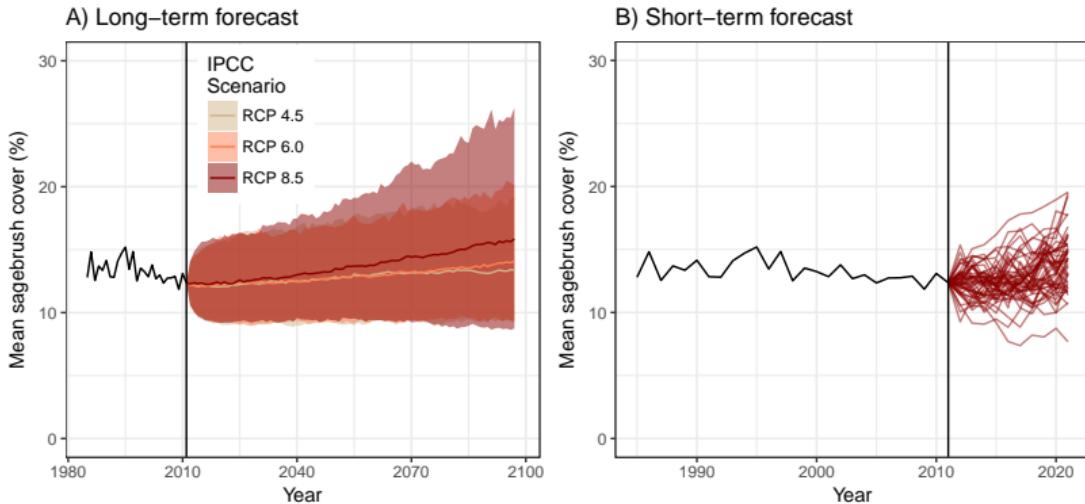


FORECASTS UNDER CLIMATE CHANGE: SPATIAL



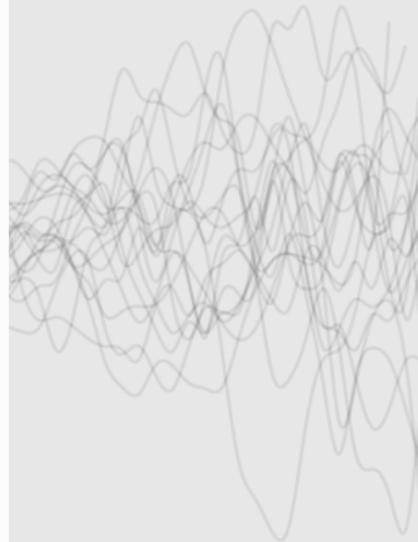


FORECASTS UNDER CLIMATE CHANGE: TEMPORAL



§ 2

Partitioning forecast uncertainty





FORECAST UNCERTAINTY, TO A FIRST APPROXIMATION

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

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



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

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



INTERACTION (COVARIANCES) CANNOT BE IGNORED

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

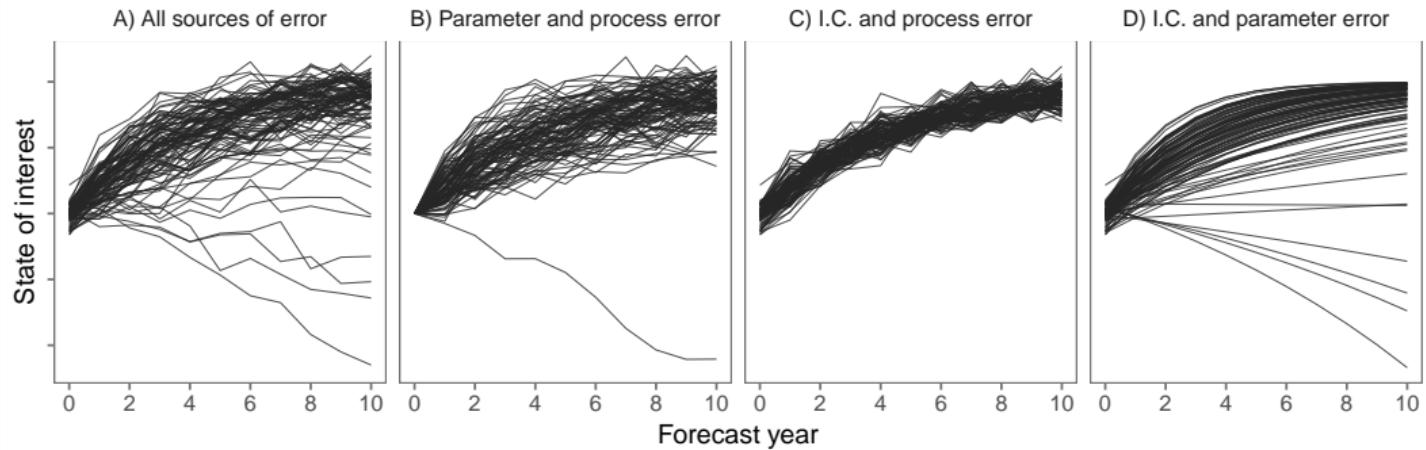
$$z_{t=1} \sim \text{Normal}(z_0, \sigma_{\text{init.}}^2), \quad \text{initial conditions uncertainty}$$

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

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



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

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



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 $T + q, \dots, T + Q$

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



POST HOC PARTITIONING FROM MCMC SAMPLES

Ignore initial conditions uncertainty

$$z_T^{(*)} = E(z_T | y_1, \dots, y_T) \approx \frac{\sum_{k=1}^K z_T^{(k)}}{K}$$

$$z_{T+q} \sim \begin{cases} \left[z_{T+q} | g(z_{T+q-1}^{(k)}, \mathbf{x}_T^{(j(k))}, \theta^{(k)}), \sigma_p^{2(k)} \right], & q > 1 \\ \left[z_{T+q} | g(z_T^{(*)}, \mathbf{x}_T^{(j(k))}, \theta^{(k)}), \sigma_p^{2(k)} \right], & q = 1. \end{cases}$$



POST HOC PARTITIONING FROM MCMC SAMPLES

k	z_T	θ_1	\dots
1	$z_T^{(*)}$	$\theta_1^{(1)}$	\dots
2	$z_T^{(*)}$	$\theta_1^{(2)}$	\dots
3	$z_T^{(*)}$	$\theta_1^{(3)}$	\dots
4	$z_T^{(*)}$	$\theta_1^{(4)}$	\dots
5	$z_T^{(*)}$	$\theta_1^{(5)}$	\dots
\vdots	\vdots	\vdots	\vdots
K	$z_T^{(*)}$	$\theta_1^{(K)}$	\dots



POST HOC PARTITIONING FROM MCMC SAMPLES

$$\mathbf{z}^{(l)} = \mathbf{z}^{(l, \overline{PA}, \overline{D}, \overline{PS})}$$

$$\mathbf{z}^{(l)} \approx \left[z_{T+q} \mid g(z_{T+q-1}^{(k)}, \mathbf{x}_T^{(*)}, \theta^{(*)}), 0 \right]$$



POST HOC PARTITIONING FROM MCMC SAMPLES

$$\mathbf{z}^{(l)} = \mathbf{z}^{(l, \overline{PA}, \overline{D}, \overline{PS})}$$

$$\mathbf{z}^{(l)} \approx \left[z_{T+q} \mid g(z_{T+q-1}^{(k)}, \mathbf{x}_T^{(*)}, \theta^{(*)}), 0 \right]$$



POST HOC PARTITIONING FROM MCMC SAMPLES

$$\mathbf{z}^{(l)} = \mathbf{z}^{(l, \overline{PA}, \overline{D}, \overline{PS})}$$

$$\mathbf{z}^{(l)} \approx \left[z_{T+q} \mid g(z_{T+q-1}^{(k)}, \mathbf{x}_T^{(*)}, \boldsymbol{\theta}^{(*)}), 0 \right]$$

$$V^{(l)} = \text{var}(\mathbf{z}^{(l)})$$



POST HOC PARTITIONING FROM MCMC SAMPLES

Source of Uncertainty	Notation
Initial conditions	$V^{(I)} = V^{(I, \bar{PA}, \bar{D}, \bar{PS})}$
Parameter uncertainty	$V^{(PA)} = V^{(\bar{I}, PA, \bar{D}, \bar{PS})}$
Driver uncertainty	$V^{(D)} = V^{(\bar{I}, \bar{PA}, D, \bar{PS})}$
Process uncertainty	$V^{(PS)} = V^{(\bar{I}, \bar{PA}, \bar{D}, PS)}$



PARTITION FORECAST UNCERTAINTY: ANOVA

$$V_{T+q}^{(F)} = V_{T+q}^{(I)} + V_{T+q}^{(PA)} + V_{T+q}^{(D)} + V_{T+q}^{(PS)}$$



PARTITION FORECAST UNCERTAINTY: ANOVA

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



PARTITION FORECAST UNCERTAINTY: ANOVA

Example where forecast is influenced by initial conditions (I) and parameter uncertainty (PA):

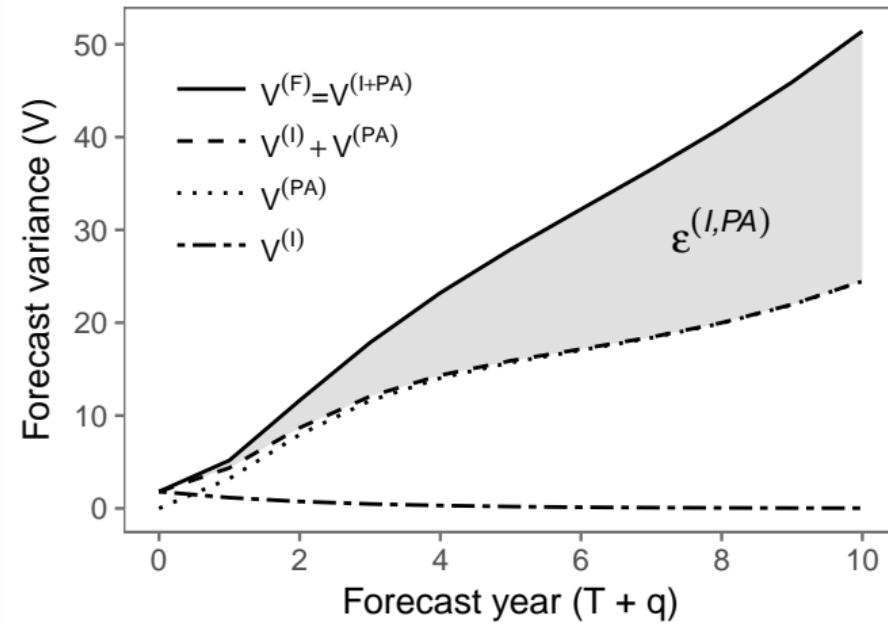
$$V_{T+q}^{(F)} = V_{T+q}^{(I)} + V_{T+q}^{(PA)} + \varepsilon_{T+q}^{(I,PA)}$$

so,

$$\varepsilon_{T+q}^{(I,PA)} = V_{T+q}^{(F)} - \left[V_{T+q}^{(I)} + V_{T+q}^{(PA)} \right]$$

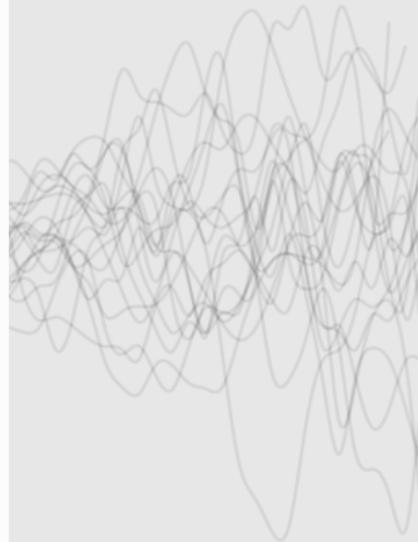


RETURN TO EXAMPLE OF AR(1) PROCESS



§ 3

Conclusions





CONCLUSIONS

1. Partition uncertainty to **advance ecological forecasting** – how do we get better?
2. Partition uncertainty to **advance scientific progress** – what don't we know?
3. Hierarchical Bayesian models ideally suited for partitioning uncertainty because they allow us to fully specify the inclusion of uncertainty.

