National Ecological Observatory Network, Inc.

Quantifying Climatically Driven Changes in Ecological Responses: A Simulation Study

Paul Duffy[1], David Schimel[2], Tom Stockton[1], and Tom Hobbs[3]

Abstract

The National Ecological Observatory Network (NEON) will be an NSF-sponsored research facility for the study of long-term, large-scale ecological change. NEON's science mission is to enable understanding and forecasting of the impacts of climate change, land use change and invasive species on continental-scale ecology by providing infrastructure and information to support research in these areas. A necessary component of this mission lies in characterizing the ability of the network to enable detection and quantification of changes in ecological responses that are driven by climatic forcing. There are two main obstacles that currently limit the ability to confidently address quantification of ecosystem change in the context of a shifting climate. First, with respect to key ecosystem responses, there are limited data that have sufficient spatio-temporal resolution and extent. Second, the methods for quantifying changes in ecosystem responses that are driven by climatic forcing have not been well developed. Since the lack of methods is at least partly due to the limited availability of the necessary data, we have developed a simple simulation framework to explore the utility of different quantitative approaches. This simulation framework allows for the general specification of: 1) The functional form of the link between climate drivers and ecosystem response variables of interest, 2) Measurement error and 3) Process error with non-separable space-time covariance structure.

The goal of this simulation study is to characterize an envelope of values for these variables that allows for the detection of trends as well as discrimination among different functional forms that characterize the relationship between climate drivers and ecosystem response variables of interest. This work represents a first step in the ongoing effort to refine quantitative methods that not only identify climatically driven trends in ecological response variables, but also allow for the specification of the functional form of the link between climate forcing and ecosystem responses.

Trend Detection Using Akaike Weights

For each realization, and at each time step, parameters for each of the three candidate models were estimated, and likelihoods were computed. Likelihoods are mapped to an Akaike Weight (AW). Then for a given year, the average AW across all of the ensemble members is computed. Results for the constant model show that after about 20 years the AW reaches 10%. This result is robust across different error levels and is used to define a threshold for the determination of the presence of a trend. Table 1 results show the envelope for the detection of a linear response. These results help quantify the sufficiency of sampling efforts and optimize designs.

Figure 1: Ensemble average (Red dots) of Response Through Time for the Constant Model

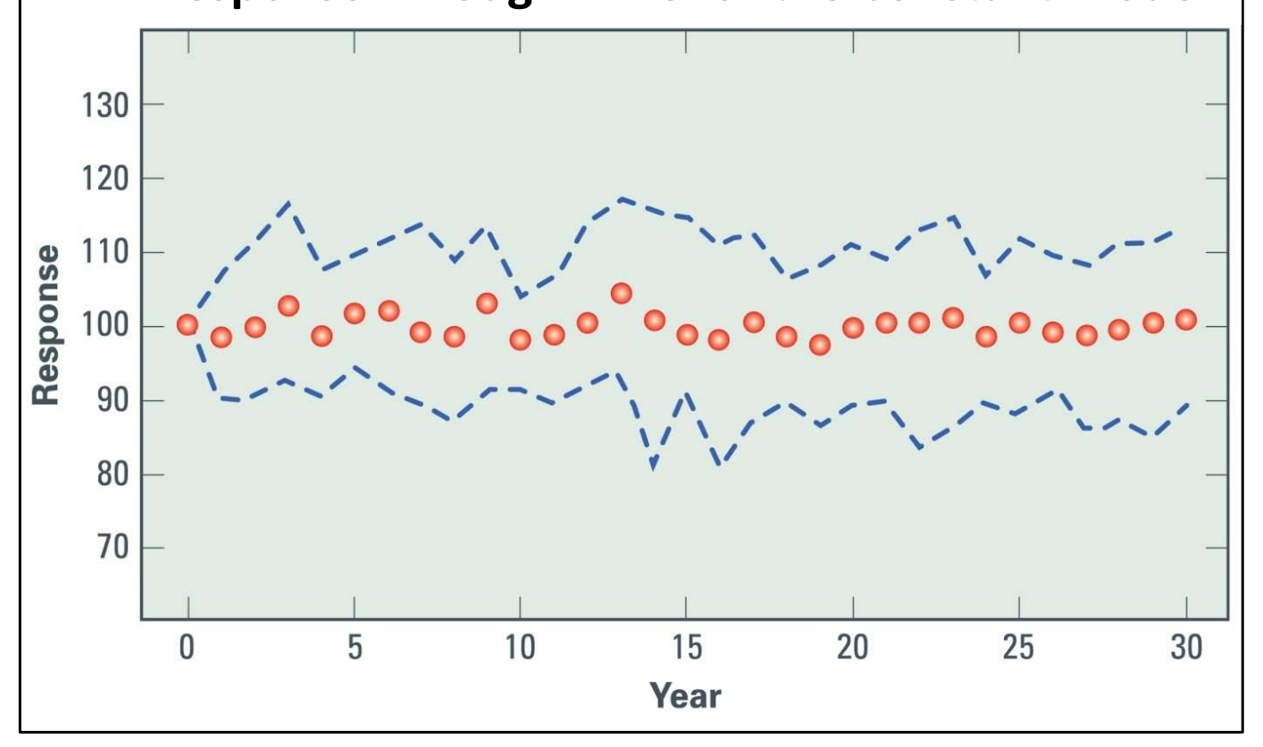


Figure 2: Ensemble Average (Red dots) of Akaike Weights Through Time for the Constant Model

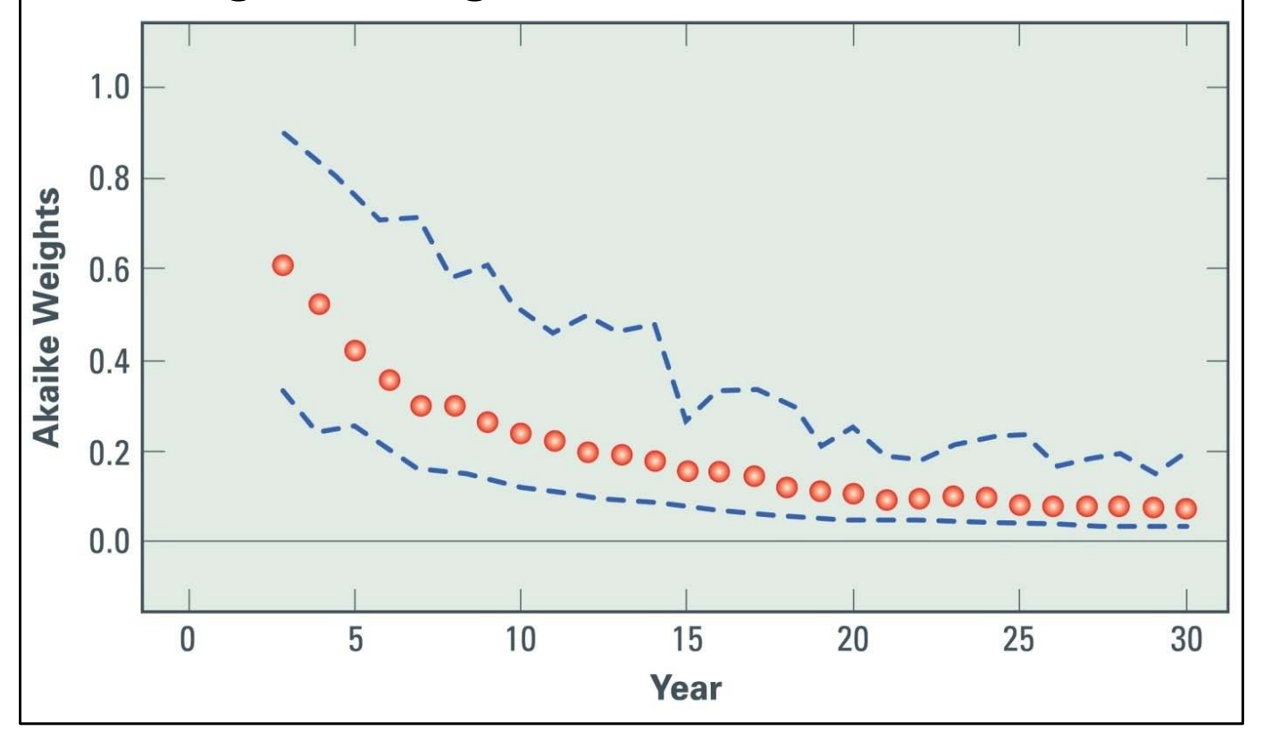
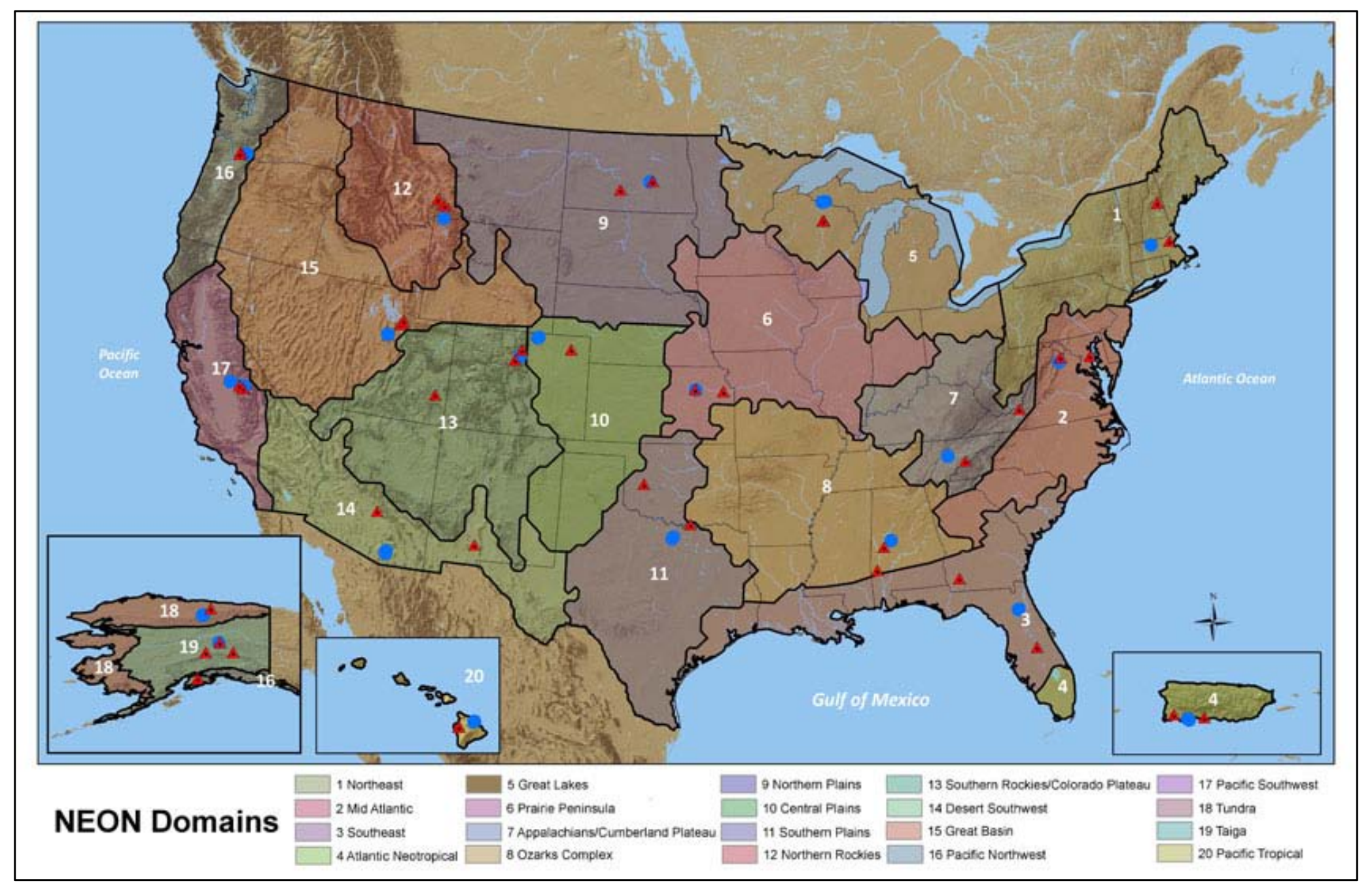


Table 1: Years Until Akaike Weight Detection of Linear Trend for [Low/High] Measurement Error

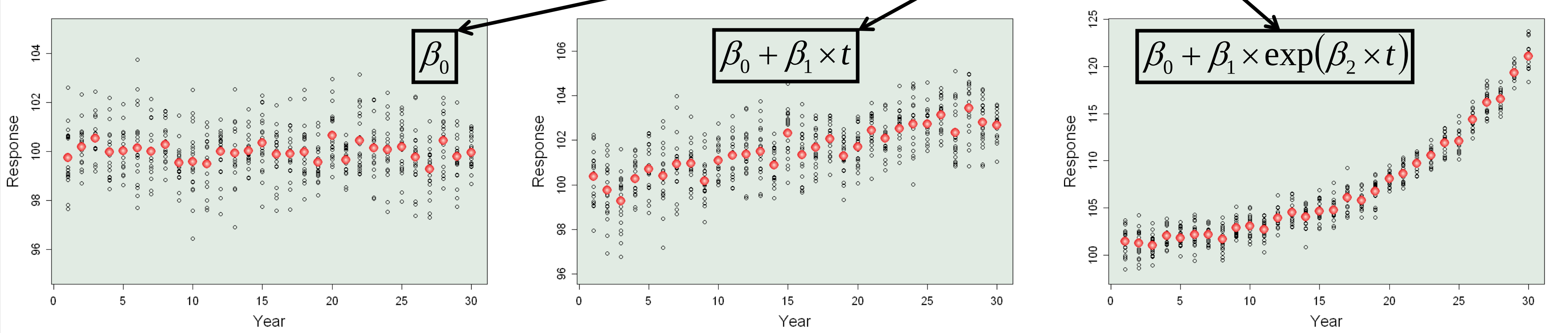
		Interannual variability		
		Low = 0.1	Medium = 0.50	High = 1.00
Magnitude of Trend	Low = 0.10	[29/29]	[>30]	[>30]
	Medium = 0.25	[14/16]	[20/21]	[23/24]
	High = 0.50	[9/10]	[12/12]	[15/15]

Simulation Setup



The simulation assigns deterministic and random components to ecological responses that are considered as samples from the 20 core sites. Data are assumed to be normally distributed and simulated annually for 30 years. Multiple realizations are generated to produce an ensemble of results for analysis. Ecological responses are simulated with one of three functional forms: 1) Constant, 2) Linear, 3) Exponential. Trend magnitude is bounded using a high and low case. The error structure allows for both measurement error, which is independent, and process error, which can have non-separable spatial and temporal components. Error cases are bounded using high and low designations, with high error corresponding to 100% process, 20% measurement error, exponential spatial correlations, AR(1) temporal autocorrelation and a space-time interaction parameter. Simulations are performed in the R statistical computing language (R core team 2009). Estimation is performed in both a classical and hierarchical setting with the former using maximum likelihood estimation (MLE) and the latter using a Markov Chain Monte Carlo (MCMC) approach. Sequential annual updating is performed to generate posterior distributions for model parameters. This is an explicit example of the data assimilation approach that will be enabled by NEON, where ecological models and forecasts can be continually updated as new data become available.

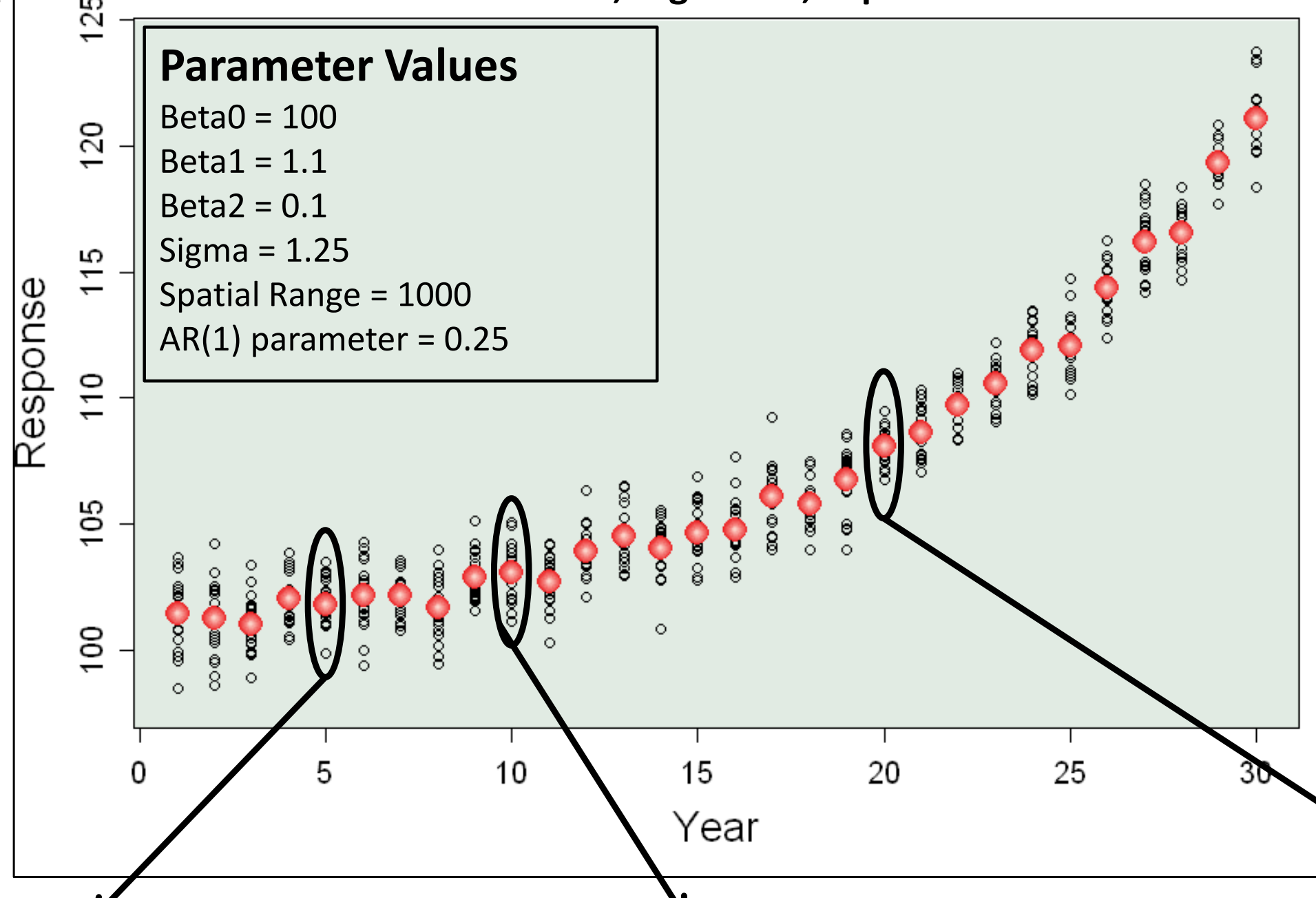
$$y = \underbrace{f(\text{climate})}\; + \varepsilon(\text{space}, \text{time})$$



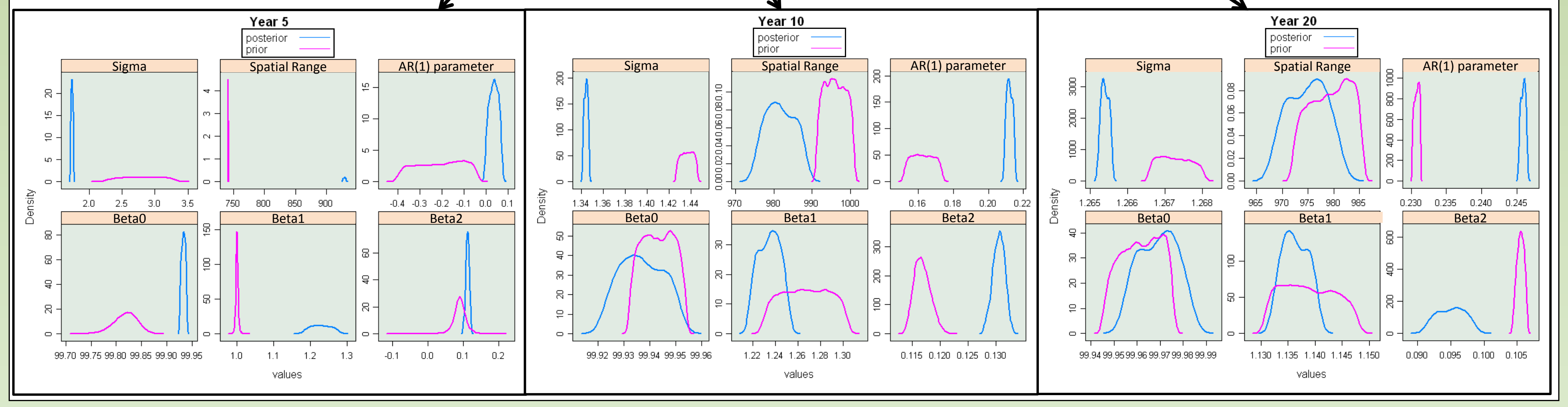
Bayesian Data Assimilation and Posterior Parameter Distributions

Initial parameter estimates for the model are made using exploratory data analysis and prior information. These can be used to inform numerical optimization algorithms for MLEs or initial parameter distributional modeling for priors in the hierarchical model. Prior distributions are updated via MCMC sampling at each time step to generate posterior distributions.

Figure 3: Average (Red dots) Across Core Sites of a Single Realization for the Low Trend, High Error, Exponential Case



Parameter Values
Beta0 = 100
Beta1 = 1.1
Beta2 = 0.1
Sigma = 1.25
Spatial Range = 1000
AR(1) parameter = 0.25



Ecological variables often respond to climatic forcing in ways that are complex both spatially and temporally. Assumptions about independence of covariance structures in space and time are often unrealistic. This simulation framework allows for the specification of non-separable spatial and temporal covariance components.

Deviance Information Criterion (DIC)

The Deviance Information Criteria (DIC) is a hierarchical modeling generalization of the Akaike Information Criteria (AIC). Like the AIC, the DIC is based on the likelihood of the data and a penalty for the number of parameters in the model. One of the main benefits of the DIC is that it can easily be computed from a collection of MCMC samples. The formula for the DIC is given by $DIC = D(\bar{\theta}) + 2P_D$. Where, $P_D = \bar{D}(\theta) - D(\bar{\theta})$, which can be thought of as the mean of the deviance minus the deviance of the means (Speigelhalter et al 2002). This provides a quantitative metric that can be used for model selection among a class of hierarchical models that are fit using MCMC methods.

Table 2: Measures of Model Fit for Linear and Exponential Functional Forms. Data were simulated from an Exponential Response with Low Trend and High Error

Year	Model	Mean Posterior	P_D	DIC
3	Linear	-75.57	4.63	160.40
	Exp	-72.57	7.30	159.75
5	Linear	-26,673.11	5.72	53,357.64
	Exp	-26,553.42	7.10	53,121.02
10	Linear	-81,488.66	5.94	162,989.21
	Exp	-81,648.50	7.69	163,312.37
20	Linear	-190,920.00	5.64	381,851.30
	Exp	-185,267.90	7.88	370,551.60
30	Linear	-328,811.80	5.07	657,633.70
	Exp	-273,504.60	8.07	547,025.30

Implications for the NEON Design

This simulation framework provides a method to assess the sufficiency of the NEON network design in the context of the mission to enable understanding and forecasting of the impacts of climate change, land use change and invasive species on continental-scale ecology. Investigators that have estimates of the potential deterministic forms and covariance structures of their ecological response of interest can assess where their data fit into the simulation framework. This framework can also be used to determine if the functional form of a linkage between climate and an ecological response variable is changing through time.

These results also help demonstrate the necessity of the NEON network with respect to the goal of enabling forecasting at the continental scale. In general, ecological data with the sufficient spatio-temporal resolution and extent necessary to resolve complex covariance structures are currently unavailable. As such, the envelope of errors used in this simulation study is assembled from disparate sources in the literature.

The most challenging scenario is the high error, low trend case for the exponential model. The low signal to noise ratio, coupled with the largest number of parameters results in the longest time needed to distinguish both the functional form of the response and obtain good estimates for the posterior distributions of the model parameters. In this case, the simulation results show that the network will provide sufficient data after roughly 20 years (Table 2). Although the representation of covariance structures in this simulation work is relatively complex, there are likely to be ecological responses that require additional model complexity. Research needs to move forward in a way that ensures models are constantly challenged and revised based on the data collected.

1: Neptune and Company, Inc., Paul Duffy, email: paul.duffy@neptuneinc.org
2: NEON, Inc., David Schimel, email: dschimel@neoninc.org
NEON, Inc. website: www.neoninc.org
3: Colorado State University