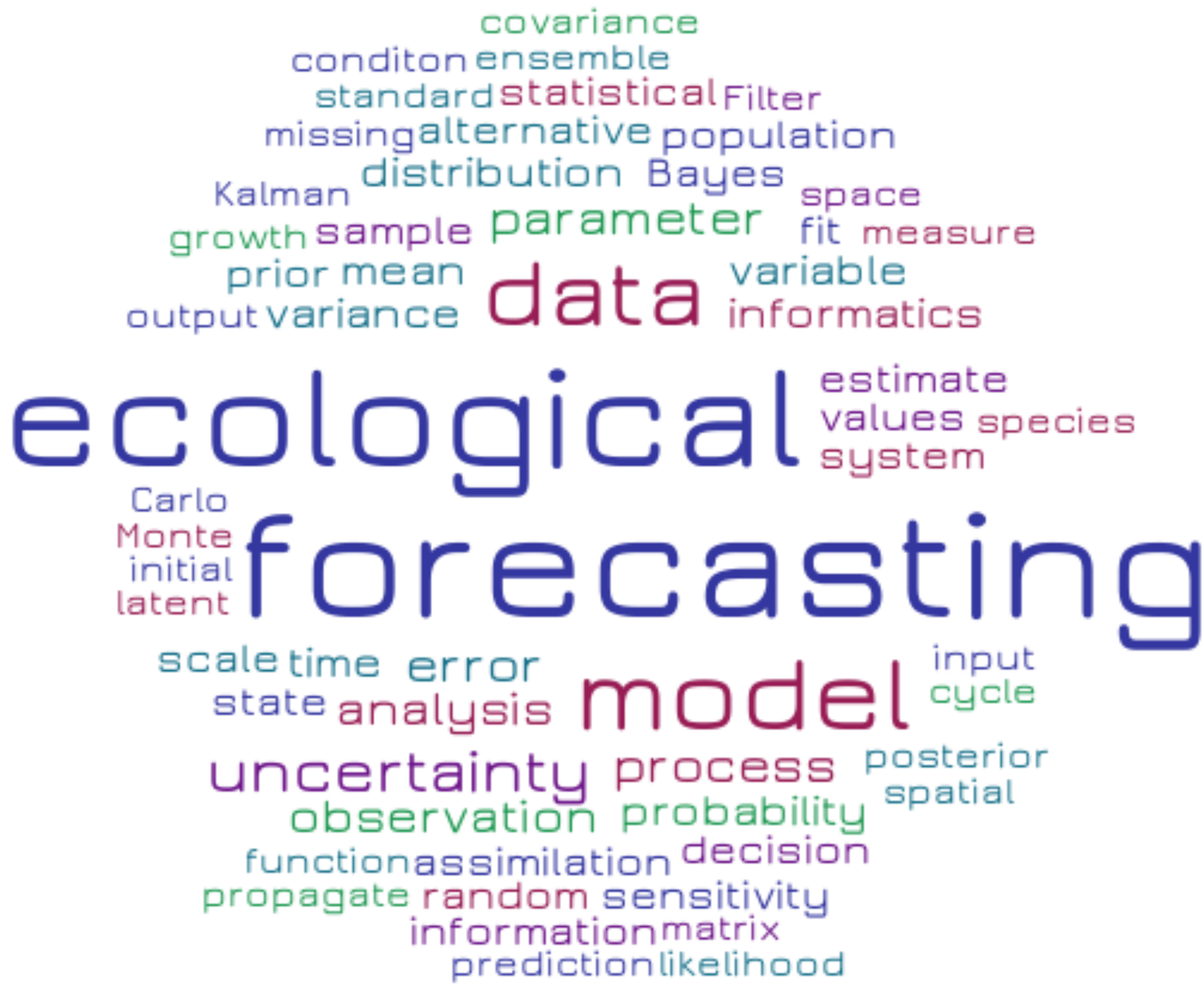
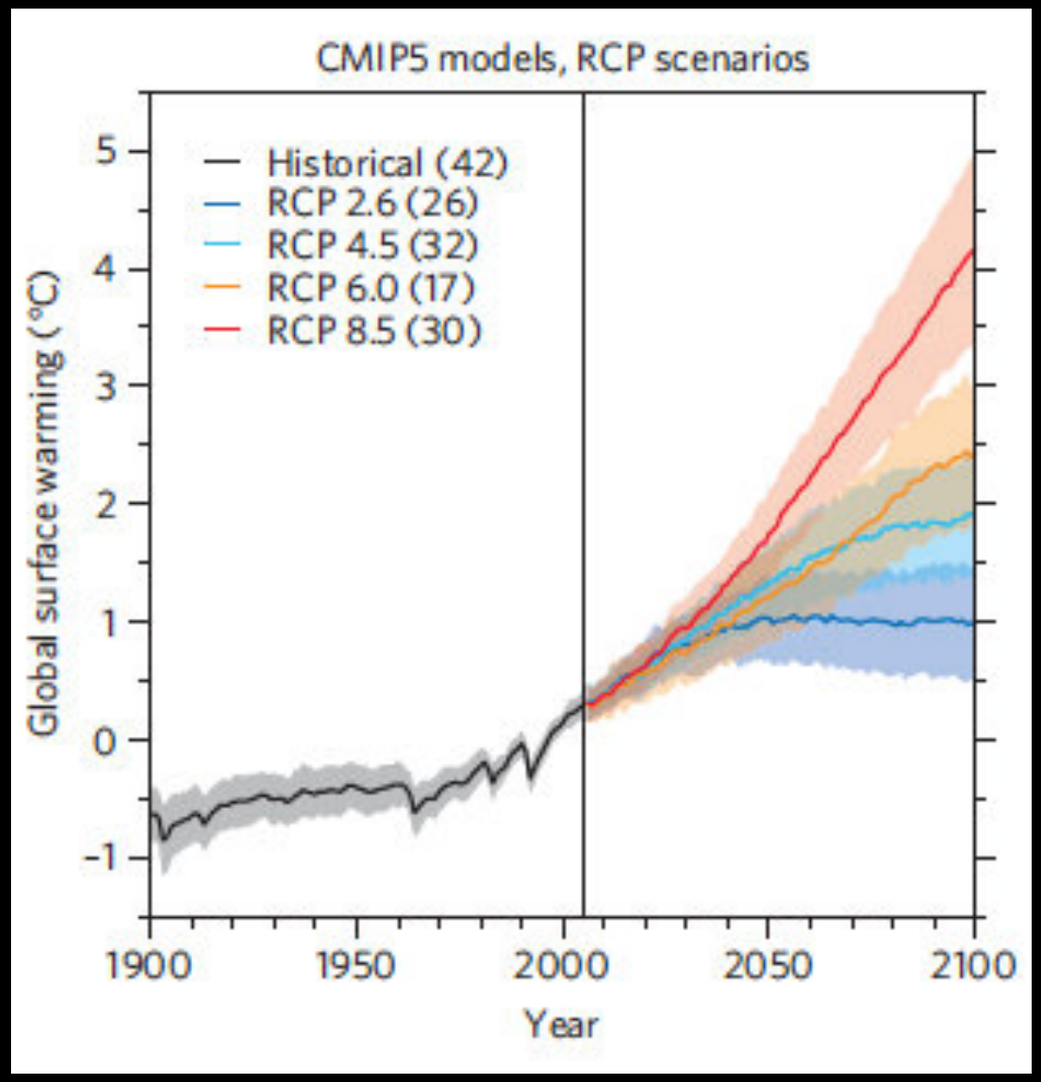

covariance
conditon
ensemble
standard
statistical
Filter
missing
alternative
population
distribution
Bayes
Kalman
space
growth
sample
parameter
fit
measure
prior
mean
variable
data
output
variance
informatics
estimate
ecological
values
species
system
Carlo
Monte
initial
latent
forecasting
scale
time
error
input
state
analysis
model
cycle
uncertainty
process
posterior
spatial
observation
probability
decision
function
assimilation
propagate
random
sensitivity
information
matrix
prediction
likelihood

Forecasting

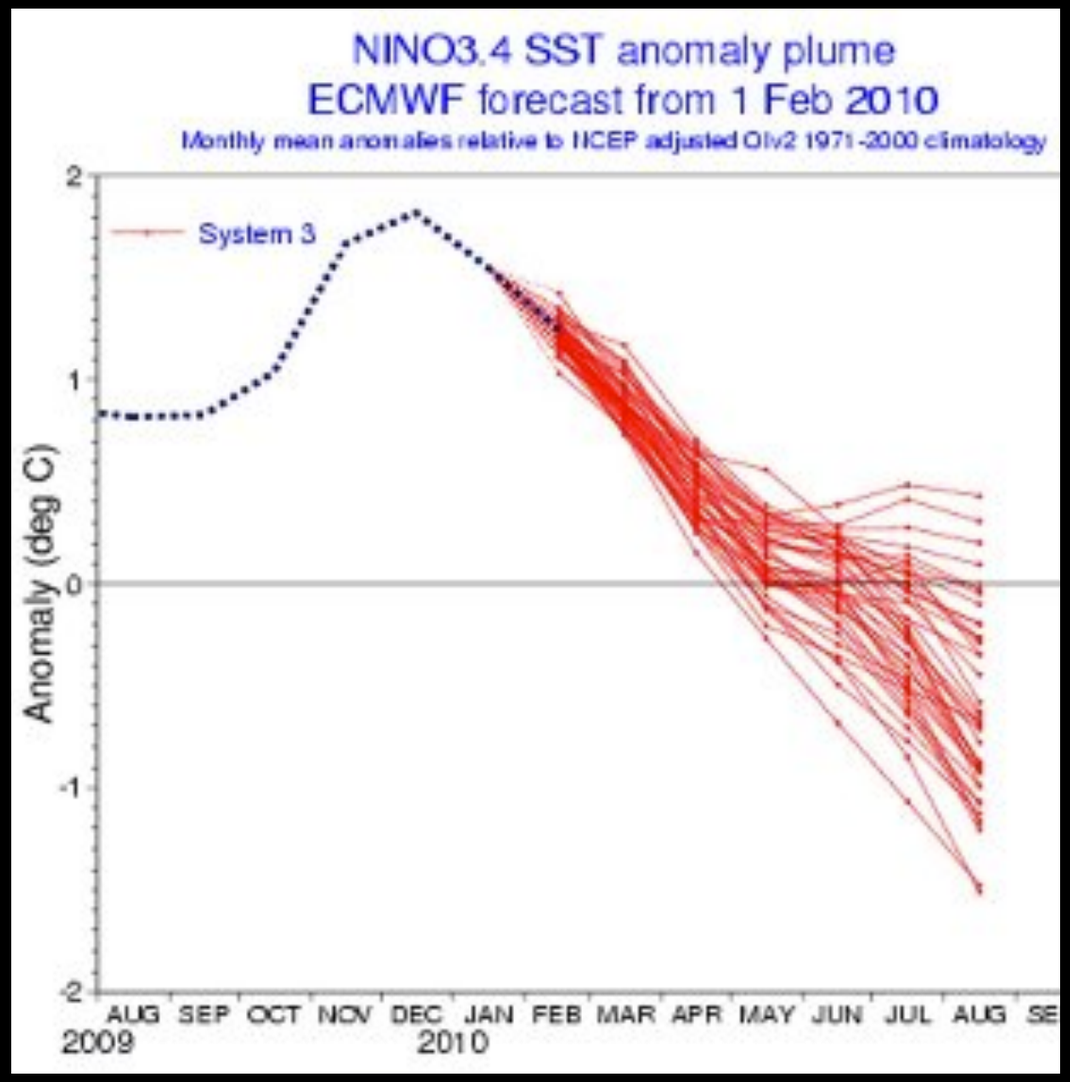
PROJECTION

"**PROBABILISTIC** STATEMENT THAT IT IS POSSIBLE THAT SOMETHING WILL HAPPEN IN THE FUTURE" GIVEN BOUNDARY CONDITION **SCENARIOS**



PREDICTION

"**PROBABILISTIC** STATEMENT THAT SOMETHING WILL HAPPEN IN THE FUTURE **BASED ON WHAT IS KNOWN** TODAY"



MacCracken 2001

WHY FORECAST?

CLIMATE CHANGE

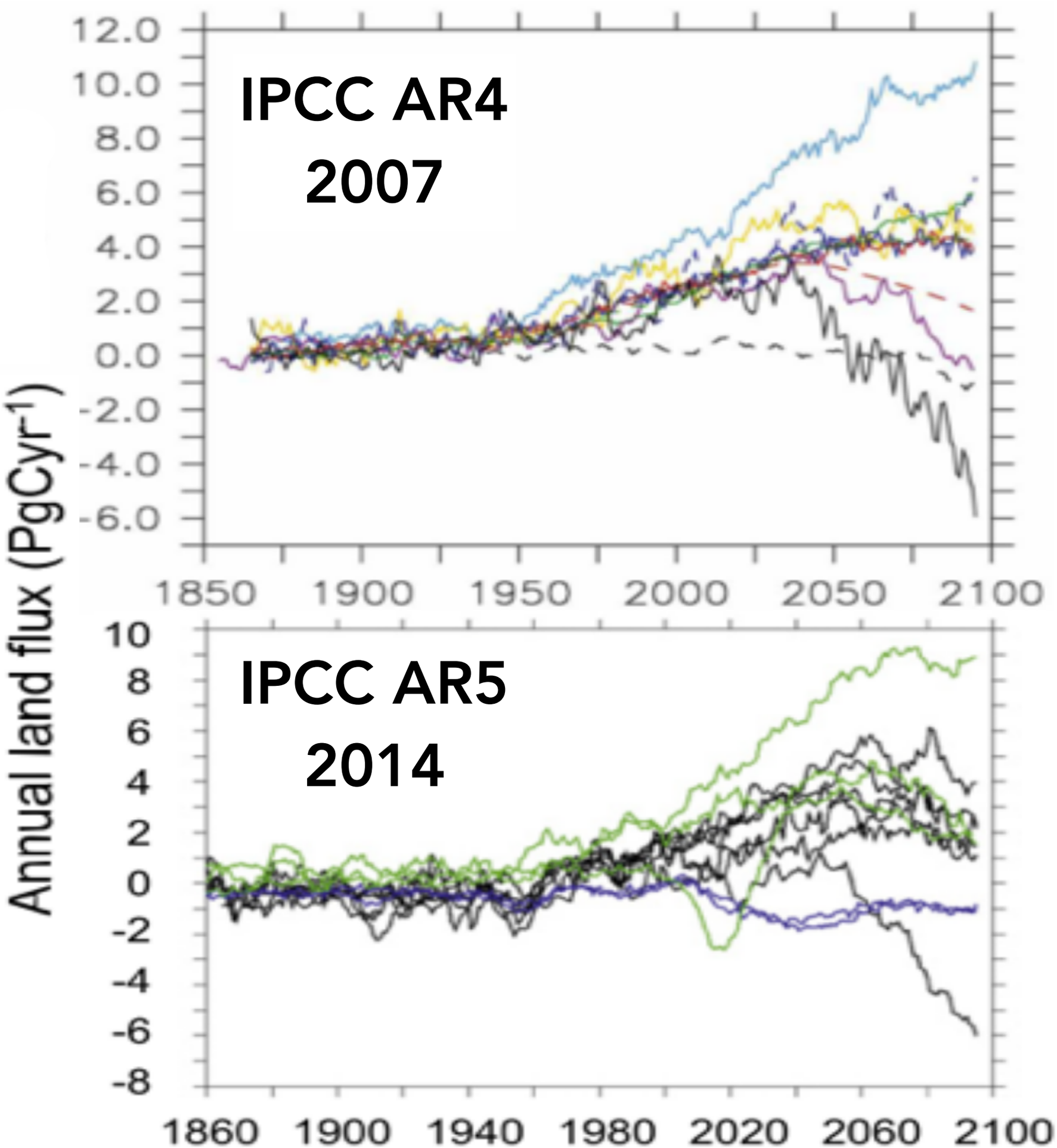
Stationarity Is Dead: Whither Water Management?

Climate change undermines a basic assumption that historically has facilitated management of water supplies, demands, and risks.

P. C. D. Milly,[1*] Julio Betancourt,[2] Malin Falkenmark,[3] Robert M. Hirsch,[4] Zbigniew W. Kundzewicz,[5] Dennis P. Lettenmaier,[6] Ronald J. Stouffer[7]

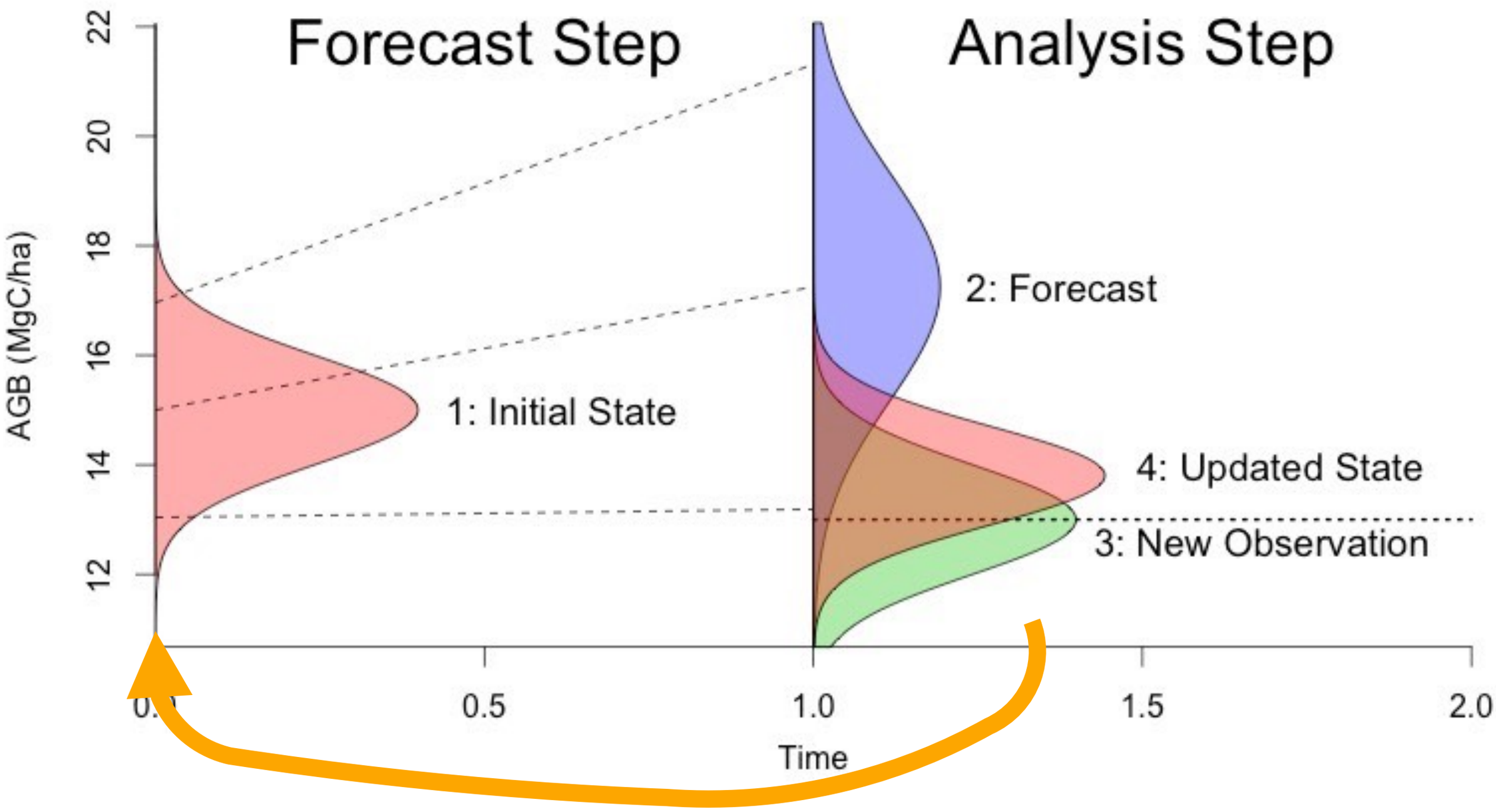
Science 2008

DECISIONS ARE ABOUT THE FUTURE

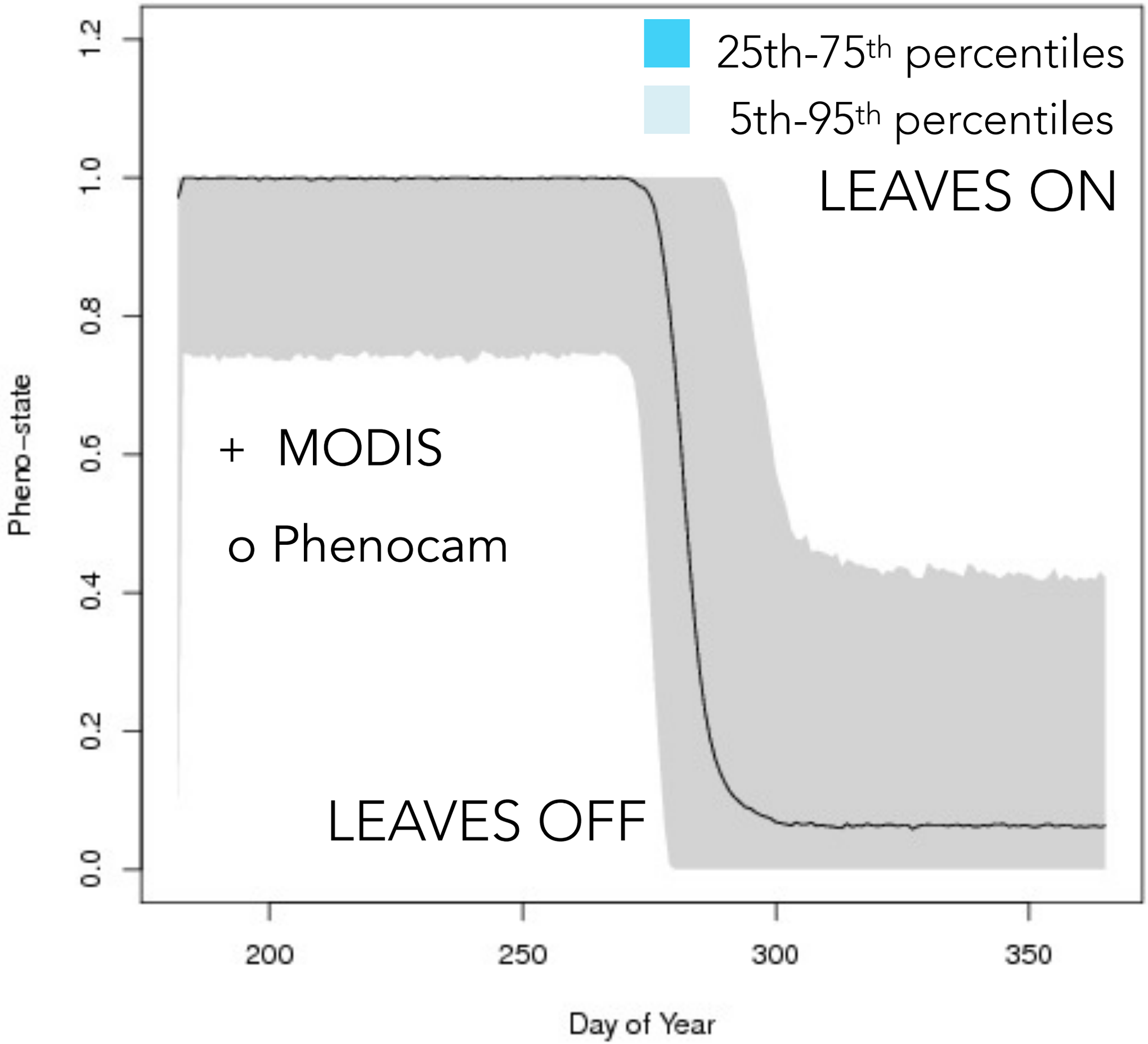


Friedlingstein et al 2006, 2014

Forecasts should be updated when new data becomes available

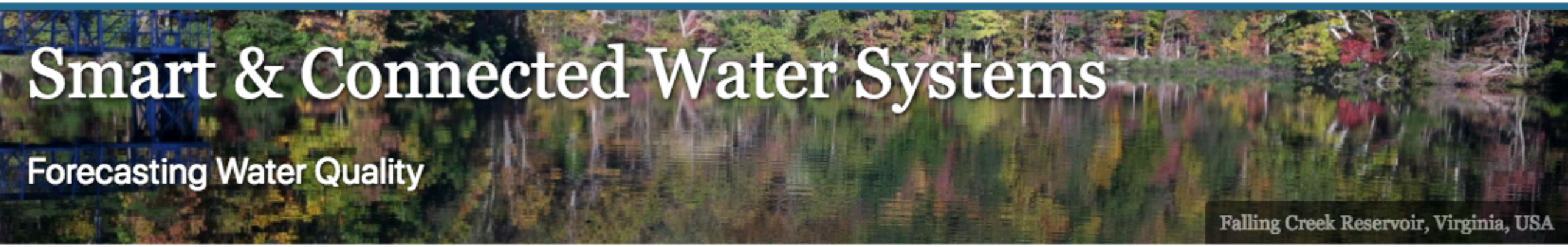


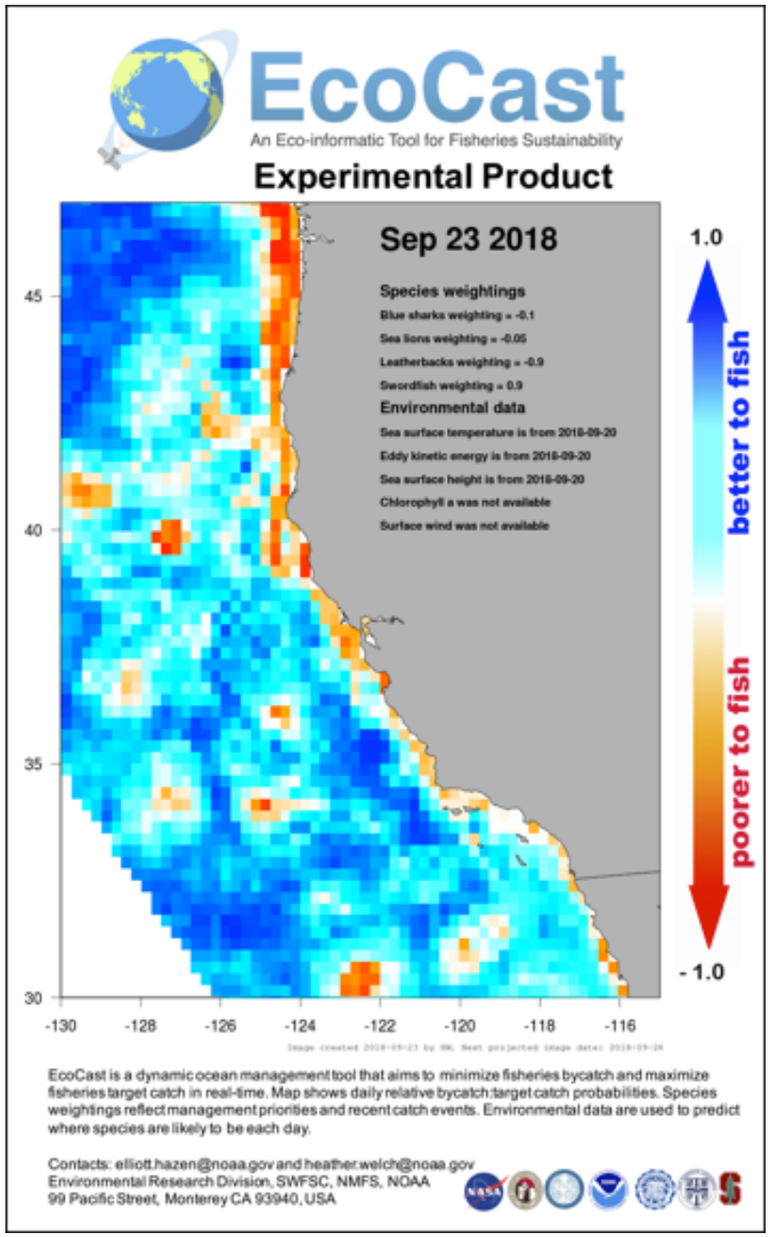
LEAF FALL

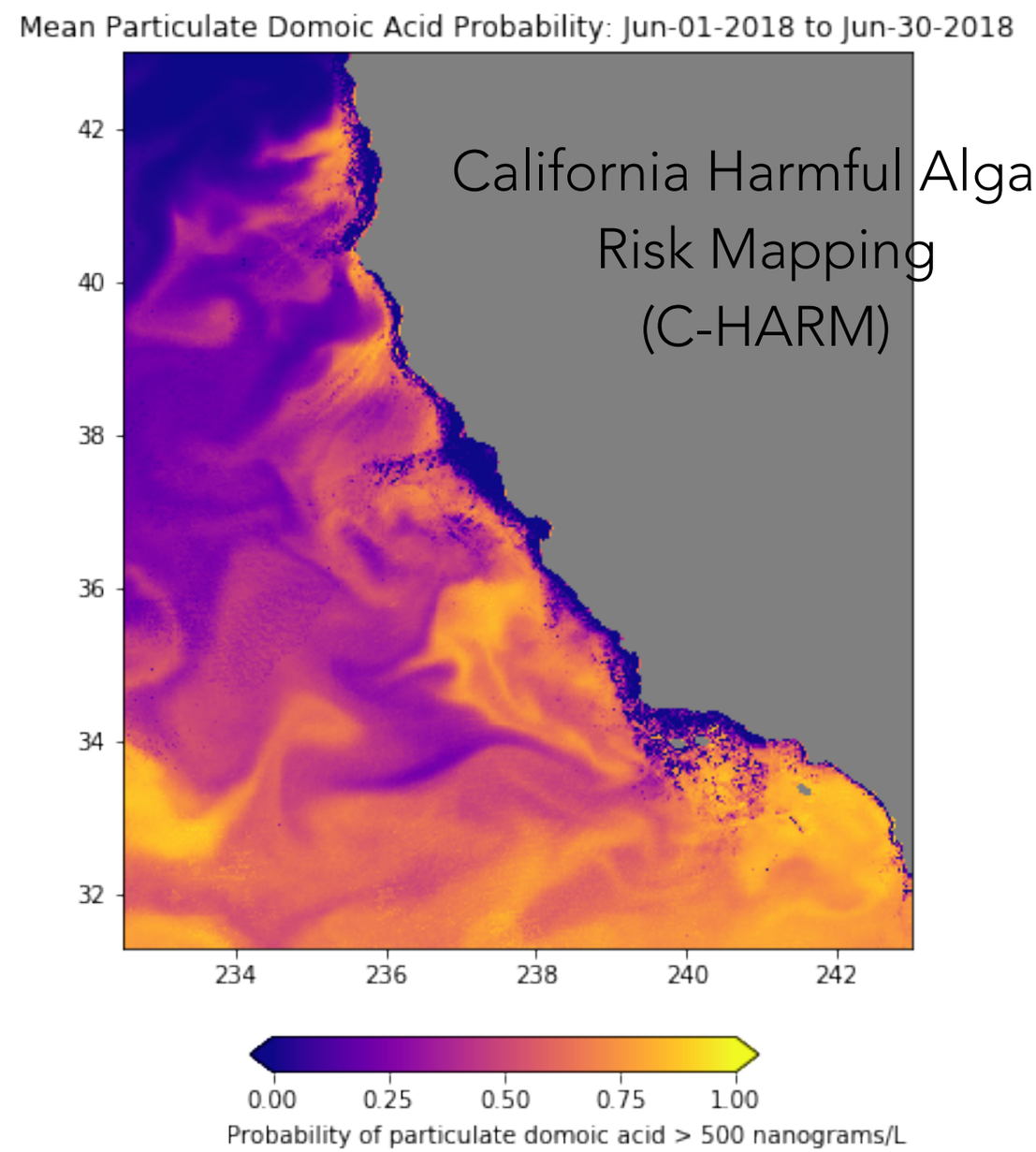

25th-75th percentiles
5th-95th percentiles
LEAVES ON
+ MODIS
o Phenocam
LEAVES OFF
Pheno-state
0.0
0.2
0.4
0.6
0.8
1.0
1.2
200
250
300
350
Day of Year

Smart & Connected Water Systems

About Team In the News Partners Contact Us





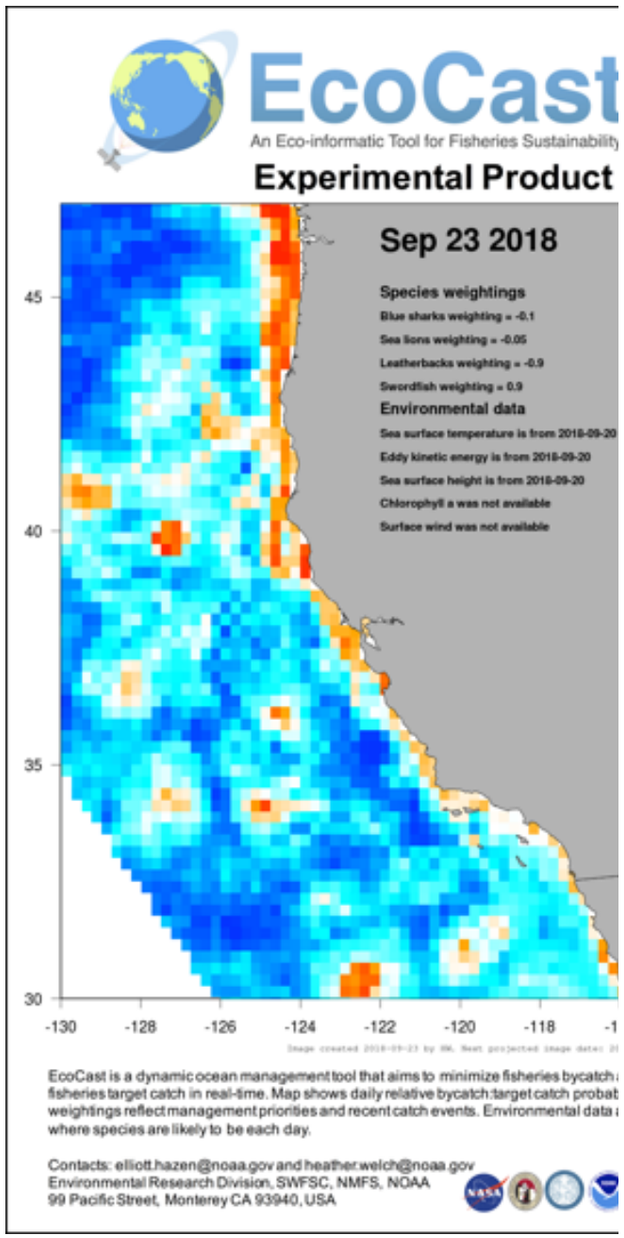


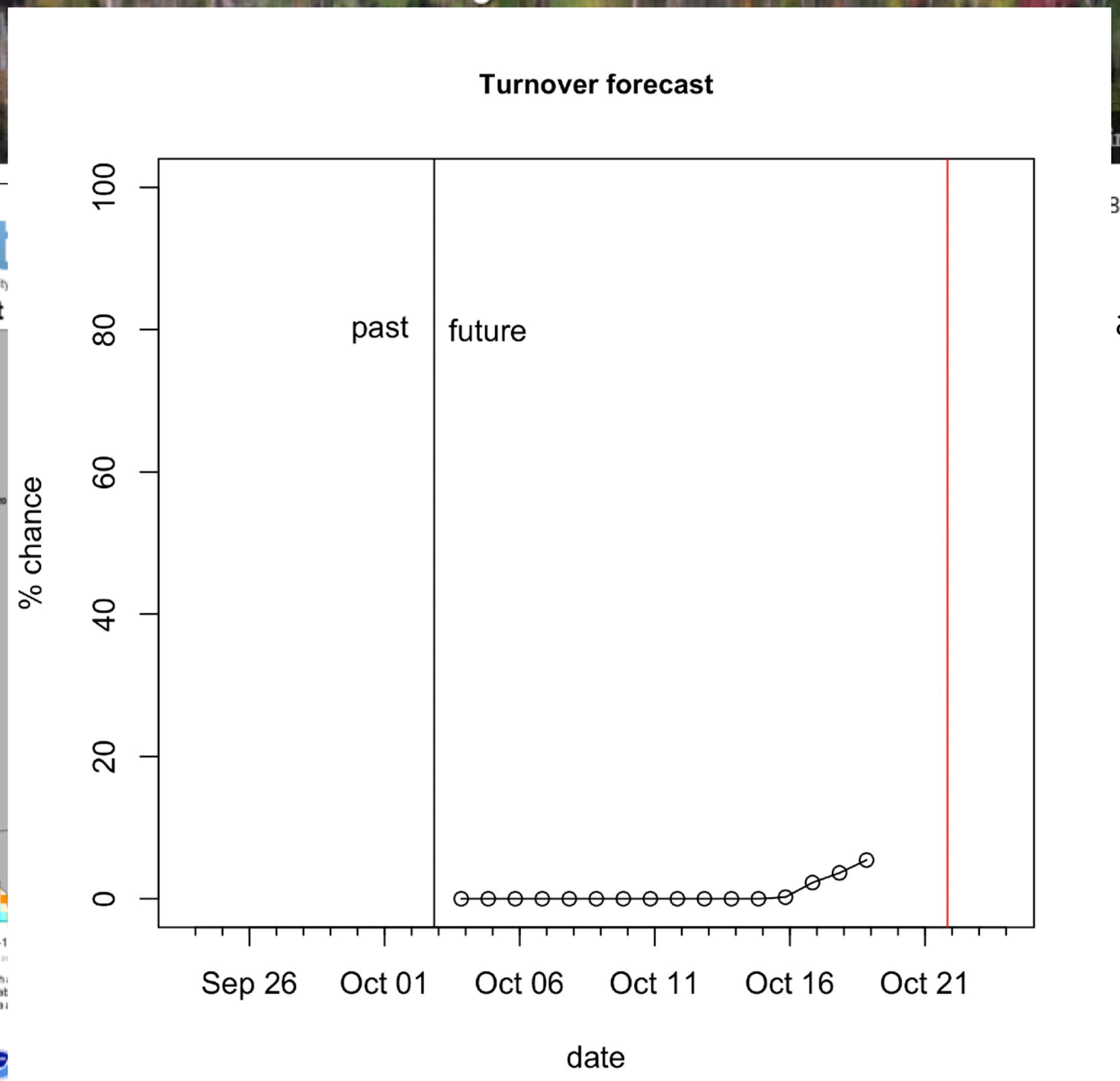
Smart & Connected Water Systems

About Team In the News Partners Contact Us

Smart & Connected Water Systems

Forecasting Water Quality

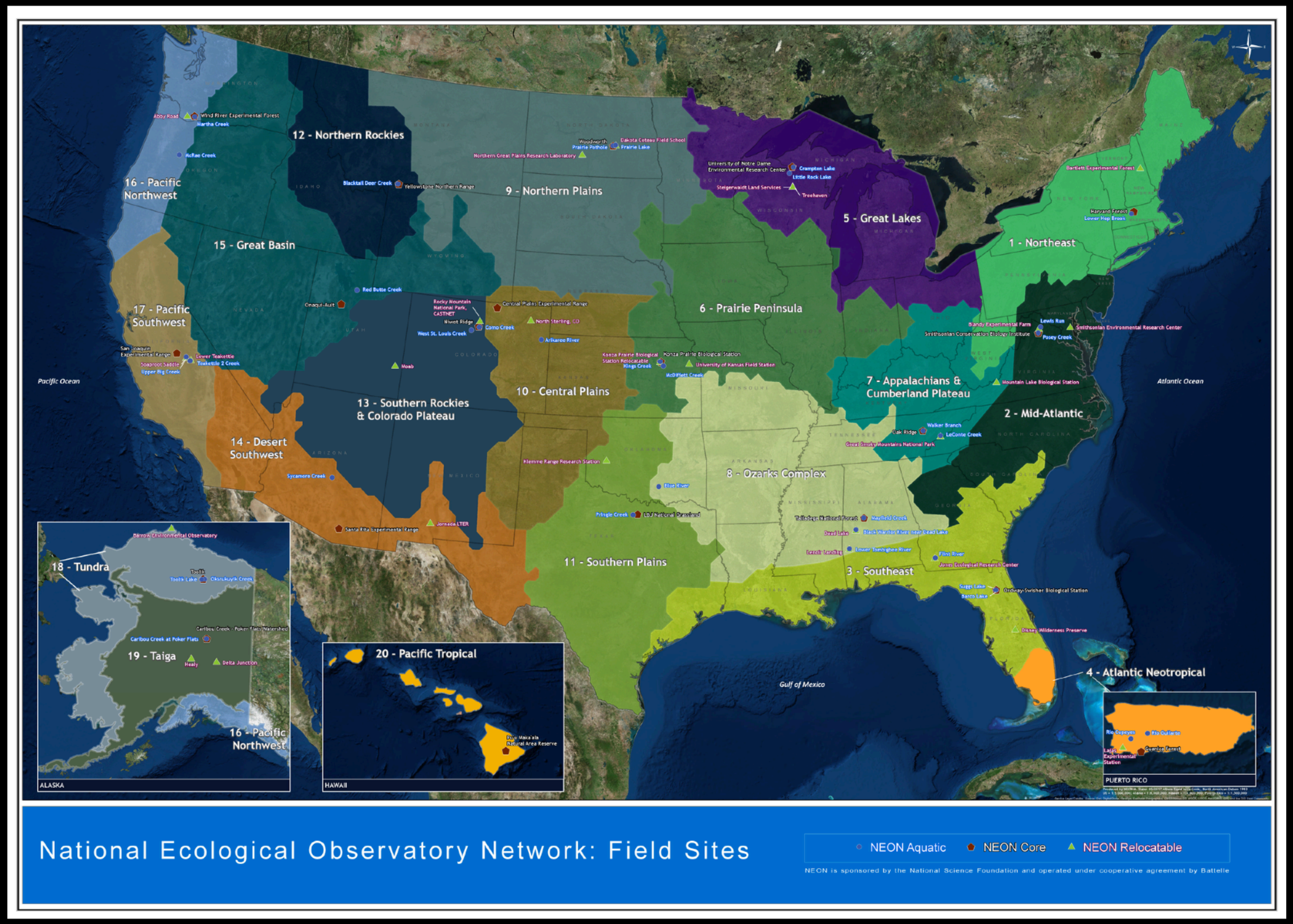




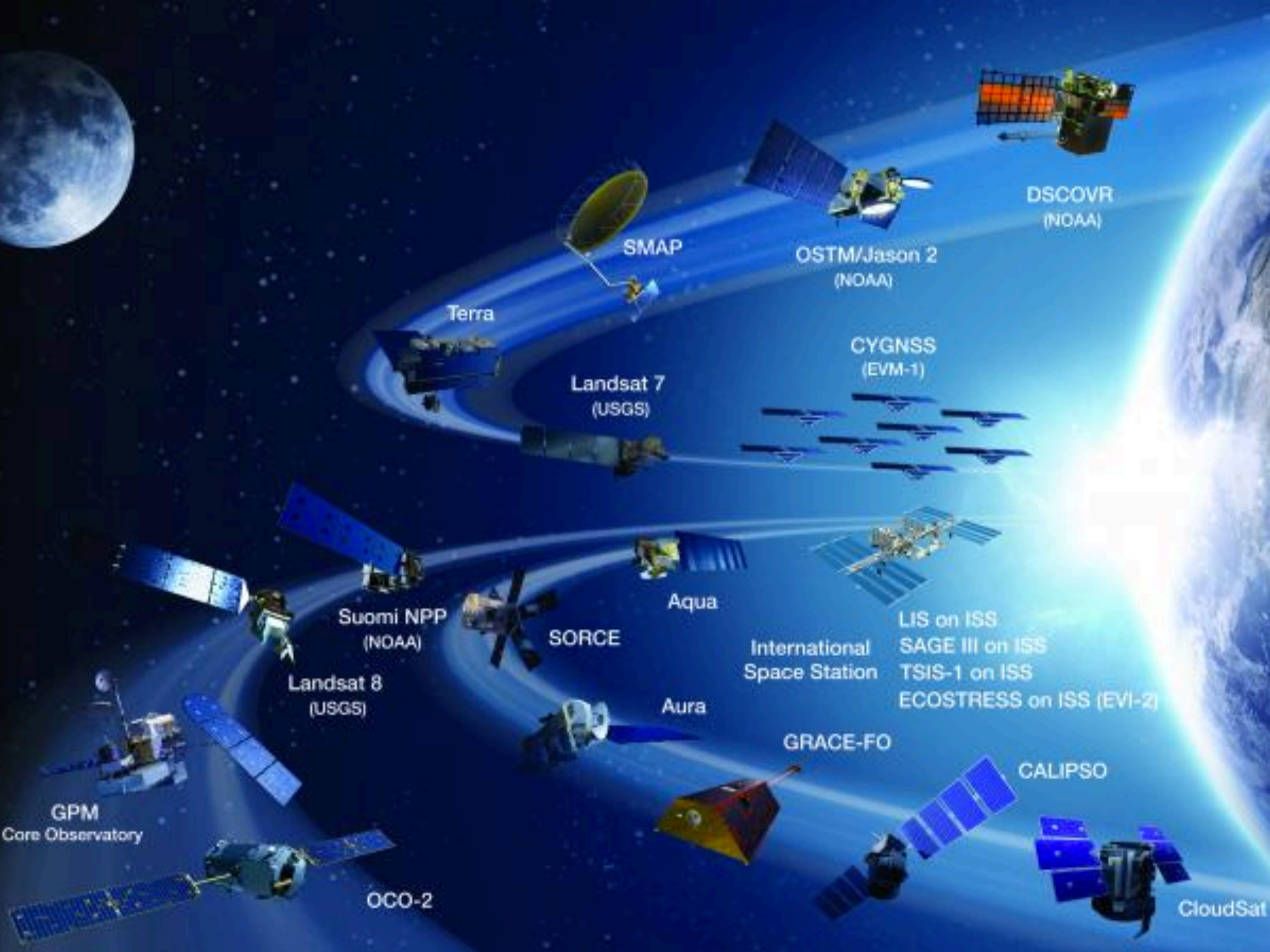
REAL-TIME MONITORING

= REAL-TIME SCIENCE

NETWORK SCIENCE



TERN, LTER, CZO, FluxNET, GLEON, NutNet, etc.


DSCOVR
(NOAA)
OSTM/Jason 2
(NOAA)
SMAP
Terra
CYGNSS
(EVM-1)
Landsat 7
(USGS)
Suomi NPP
(NOAA)
SORCE
Aqua
International
Space Station
LIS on ISS
SAGE III on ISS
TSIS-1 on ISS
ECOSTRESS on ISS (EVI-2)
Landsat 8
(USGS)
Aura
GRACE-FO
CALIPSO
GPM
Core Observatory
OCO-2
CloudSat

iNaturalist

Explore Learn Record

USA npn

National Phenology Network

Taking the Pulse of Our Planet

COMMUNITY SCIENCE

naturewatch

ENGAGING CITIZENS IN SCIENCE

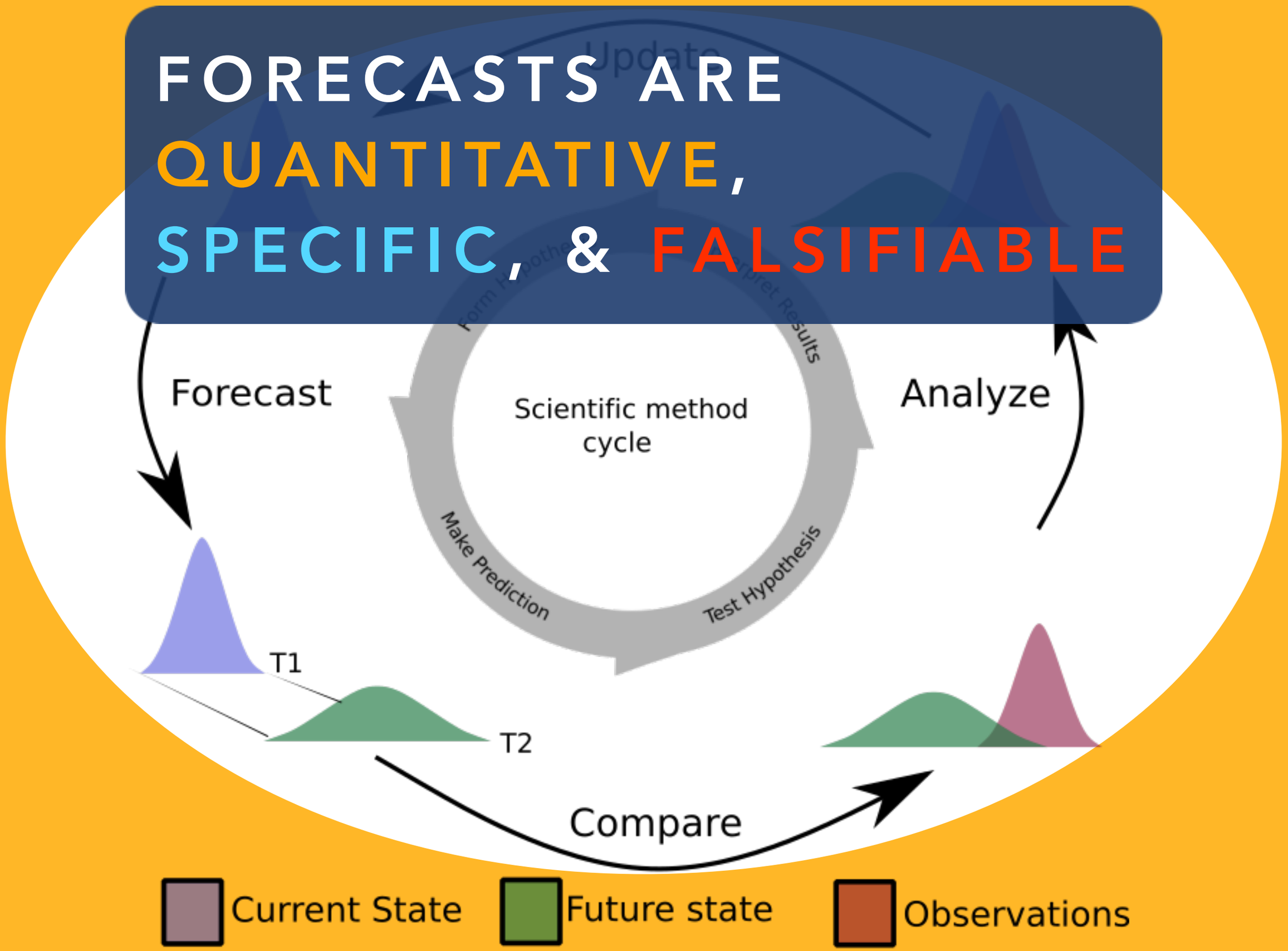
cyano TRACKER

eBird

GLOBAL SOUNDSCAPES

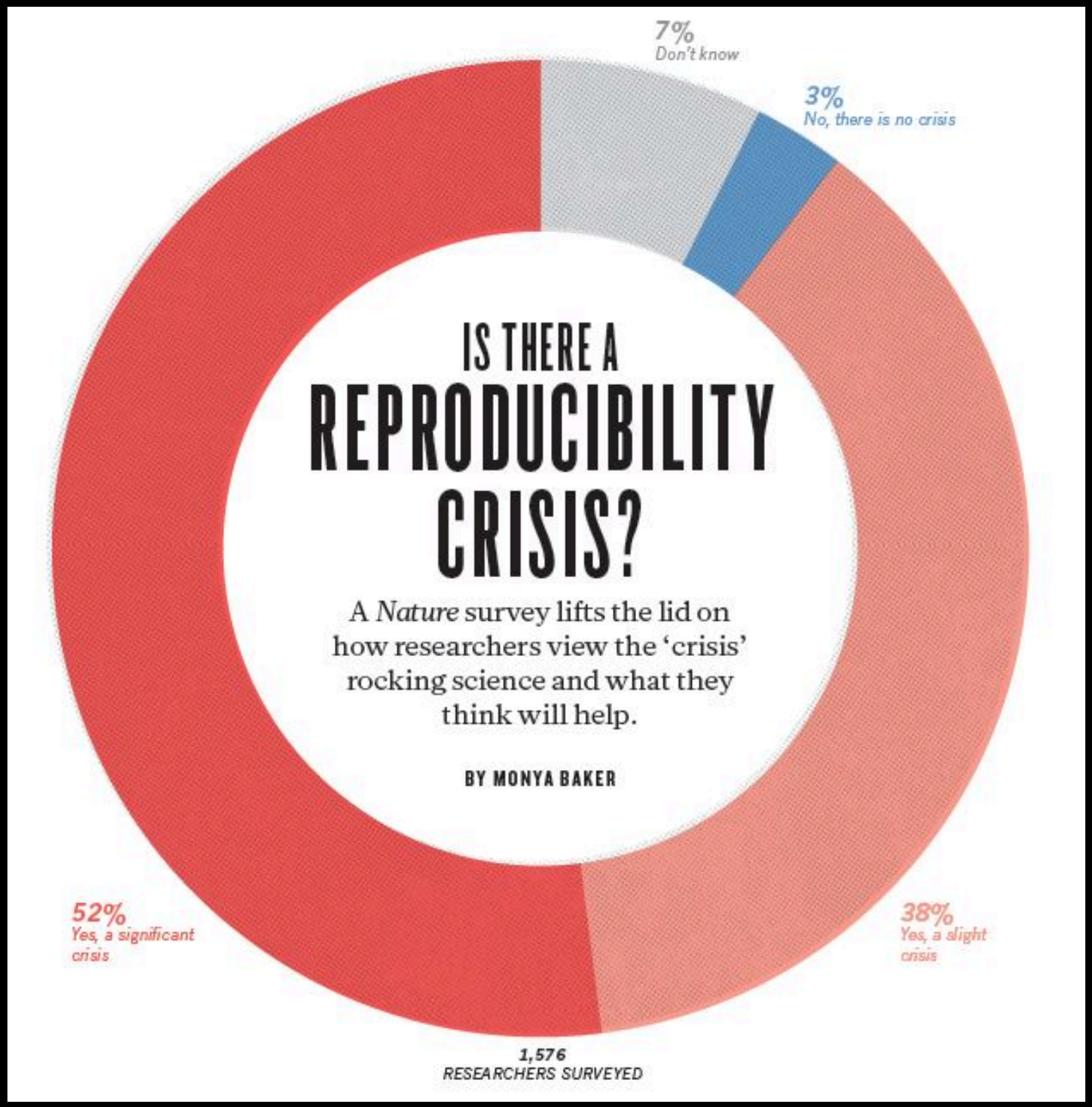
NEAR-TERM ITERATIVE FORECASTING IS WIN-WIN

FORECASTS ARE
QUANTITATIVE,
SPECIFIC, & FALSIFIABLE



Dietze et al 2018 PNAS

Forecasts are <u>a priori</u>

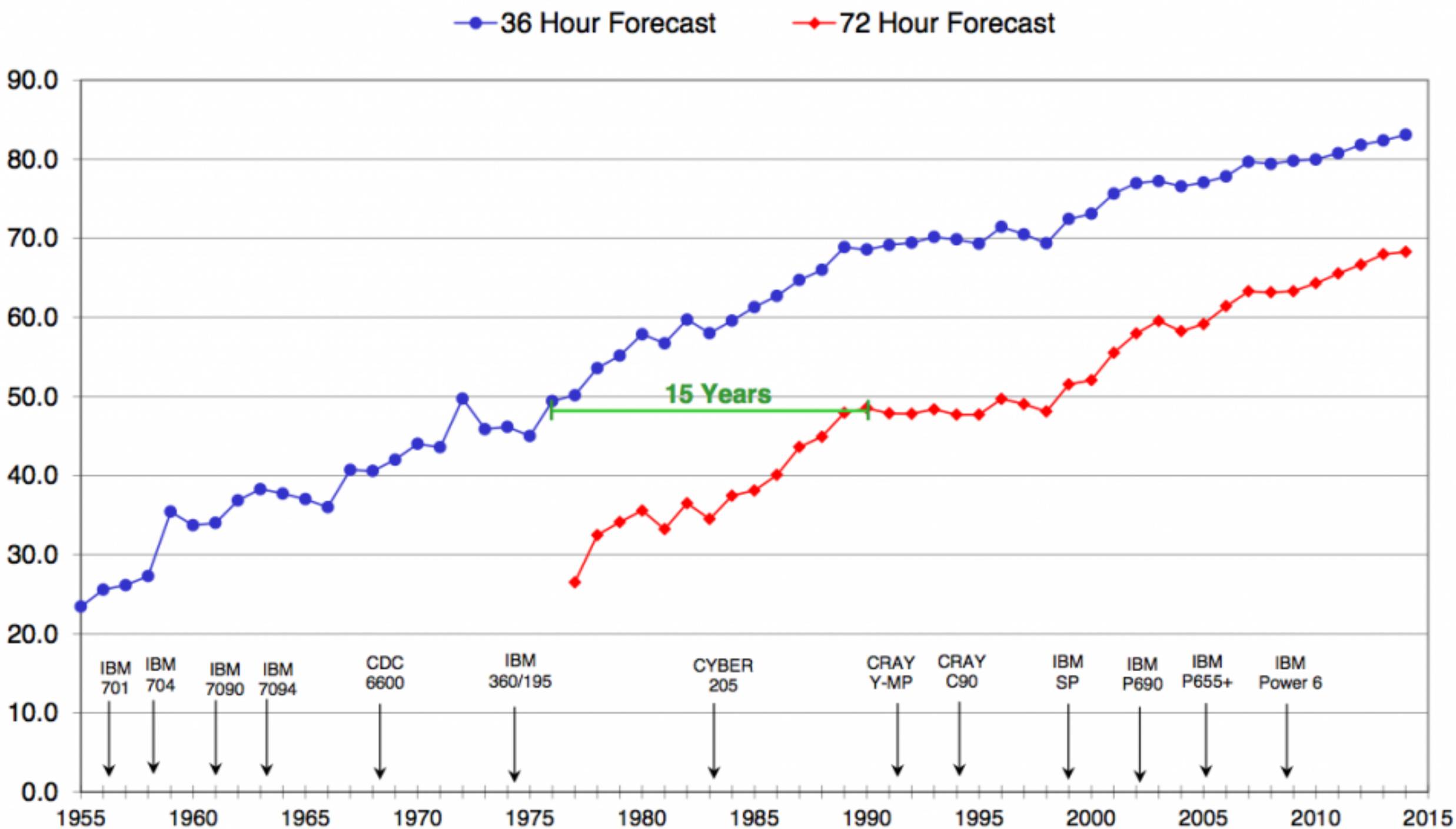


and out-of-sample

NCEP Operational Forecast Skill

36 and 72 Hour Forecasts @ 500 MB over North America

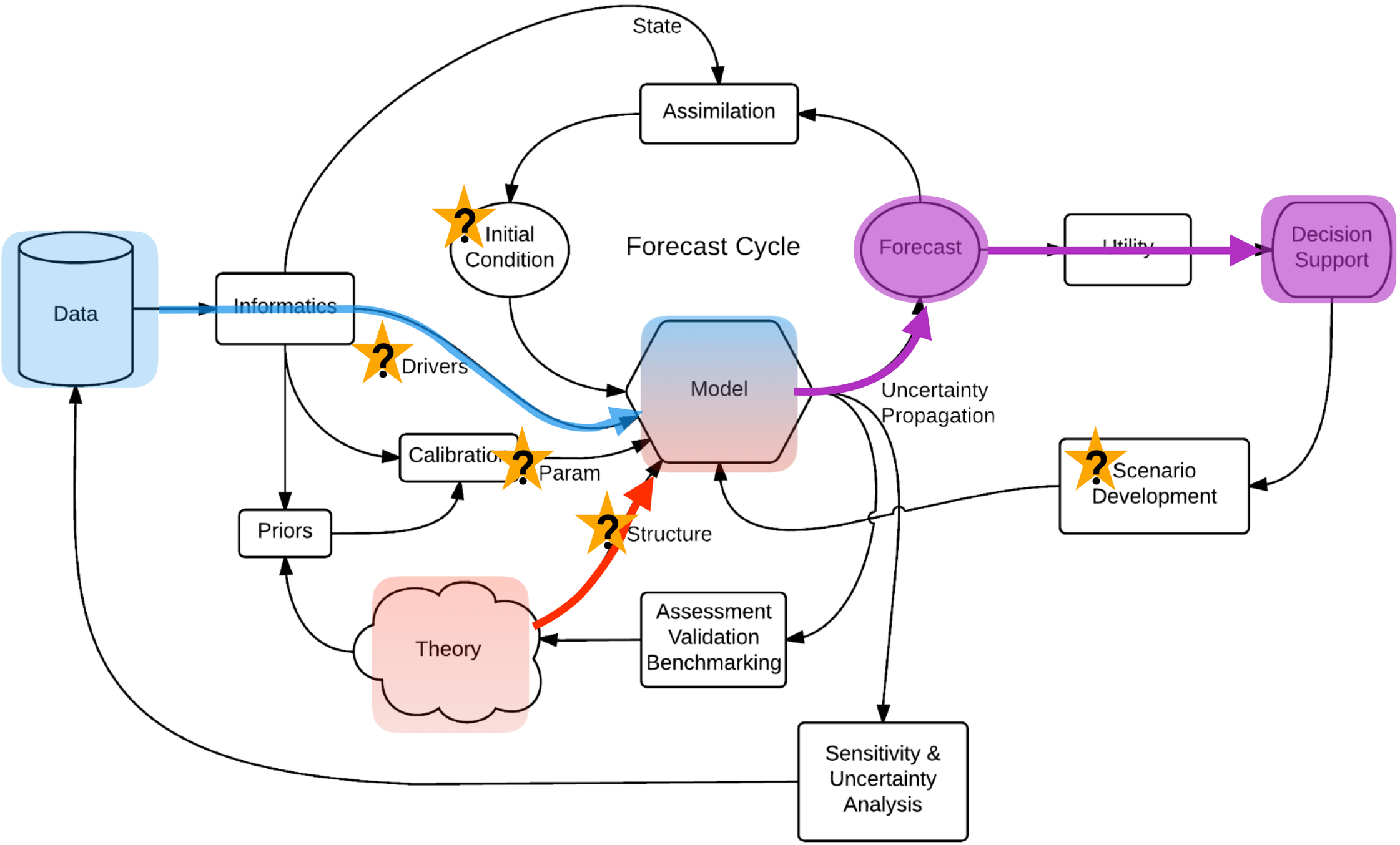
[100 * (1-S1/70) Method]



NCEP Central Operations January 2015

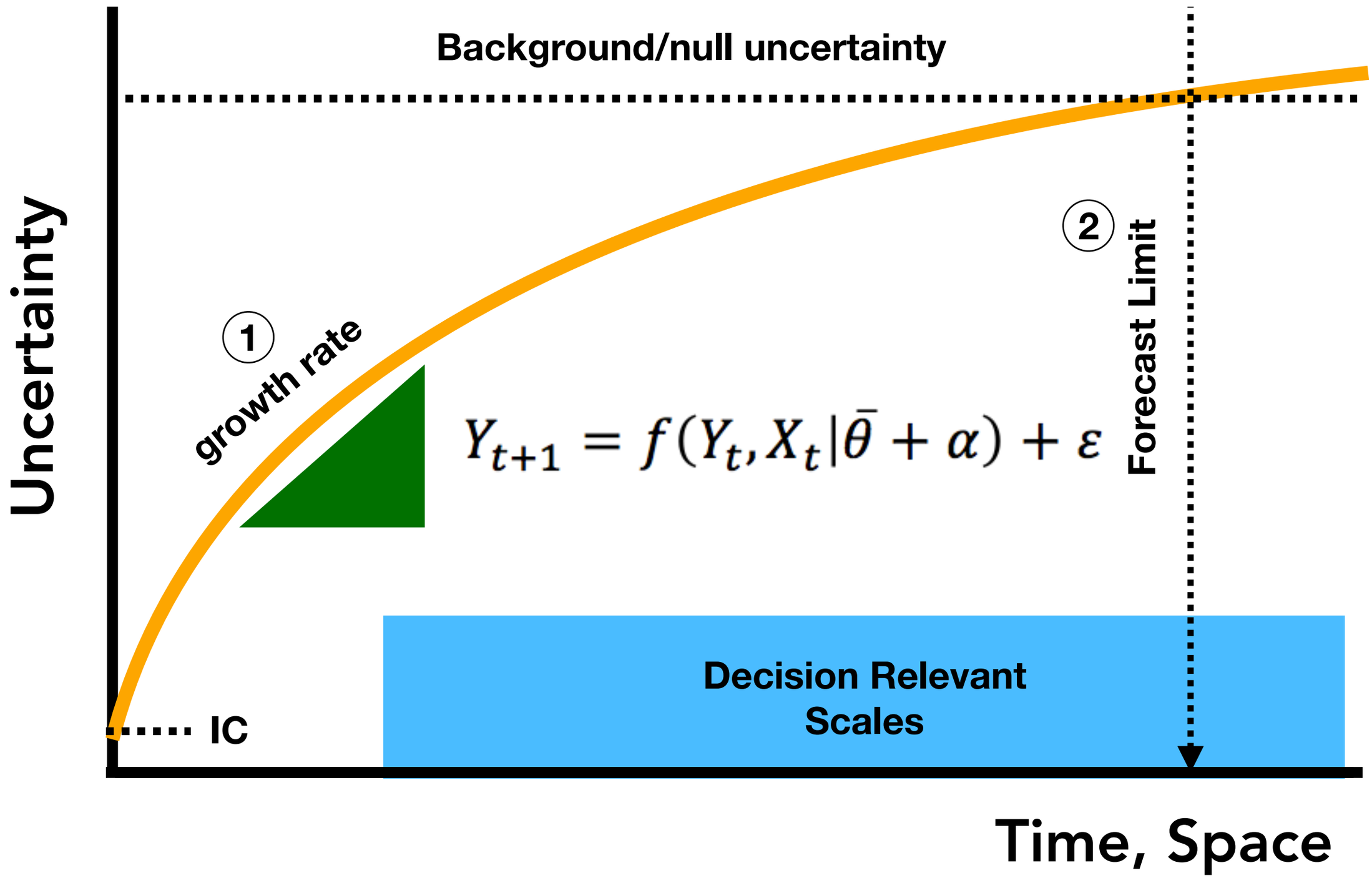
Can we forecast ecology like we forecast weather?

- Theory
- Methods
- Think Probabilistic



Dietze 2017 “Ecological Forecasting” Princeton University Press

HOW DO WE MEASURE PREDICTABILITY?

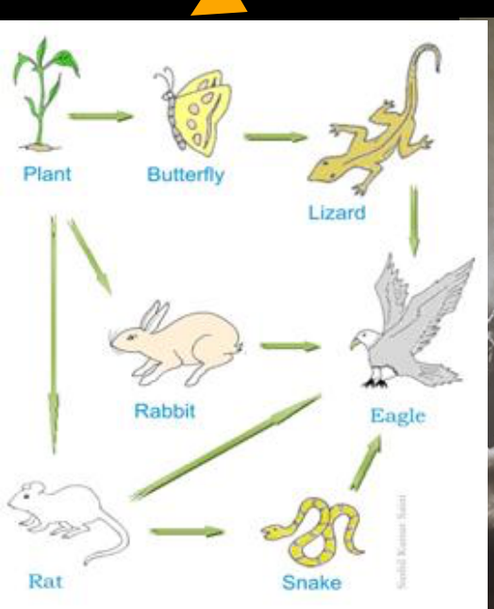


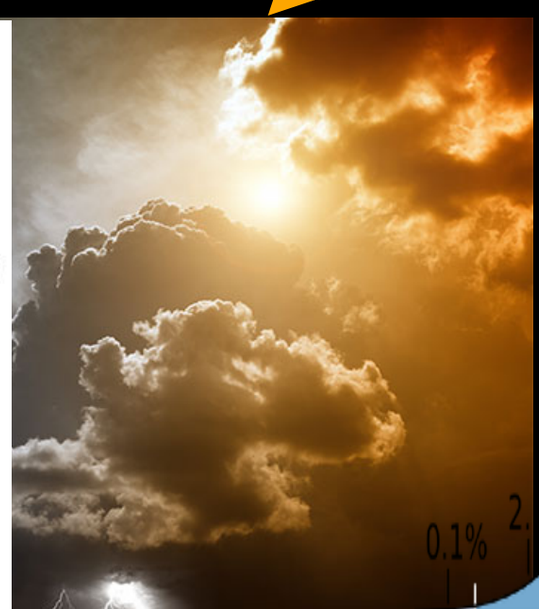
Dietze *2017 Ecological Applications*

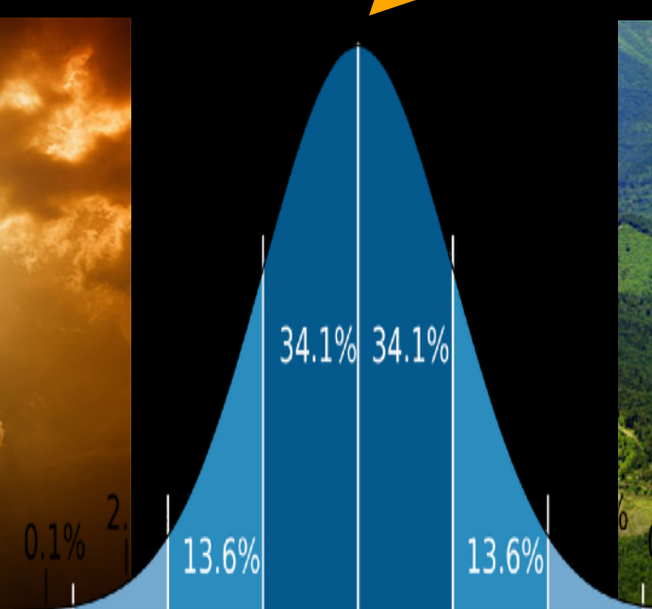
WHAT CAUSES VAR TO INCREASE WITH TIME?

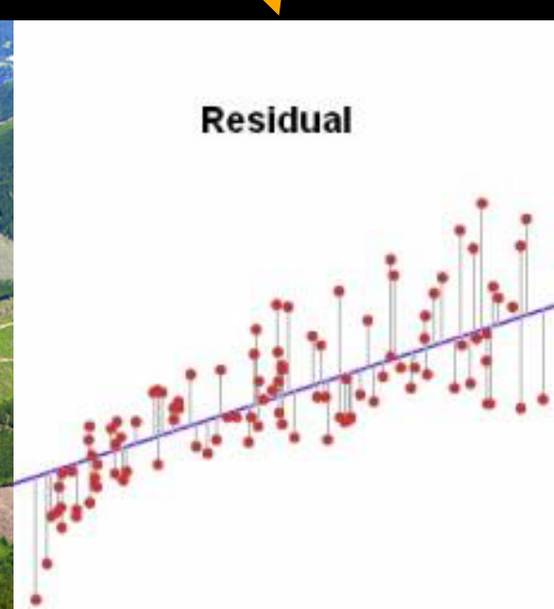
$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{\partial f}{\partial Y}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC\\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial X}\right)^2}_{\substack{driver\\ sens}} \underbrace{Var[X]}_{\substack{driver\\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial \theta}\right)^2}_{\substack{param\\ sens}} \left(\underbrace{Var[\bar{\theta}]}_{\substack{param\\ uncert}} + \underbrace{Var[\alpha]}_{\substack{param\\ variability}} \right) + \underbrace{Var[\varepsilon]}_{\substack{process\\ error}}$$

$$= INTERNAL + EXTERNAL + PARAMETERS + \begin{matrix} RANDOM \\ EFFECTS \end{matrix} + \begin{matrix} PROCESS \\ ERROR \end{matrix}$$

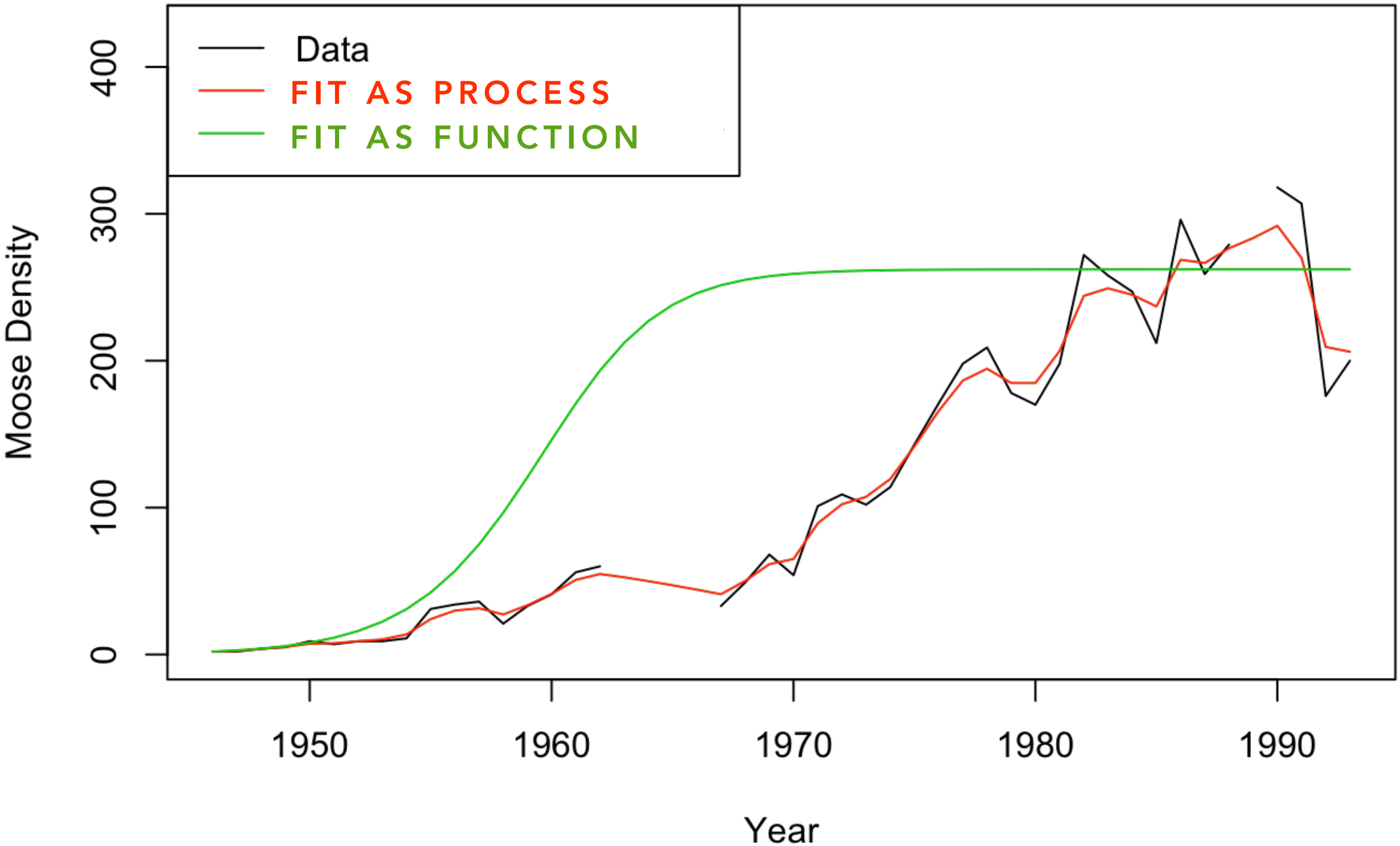






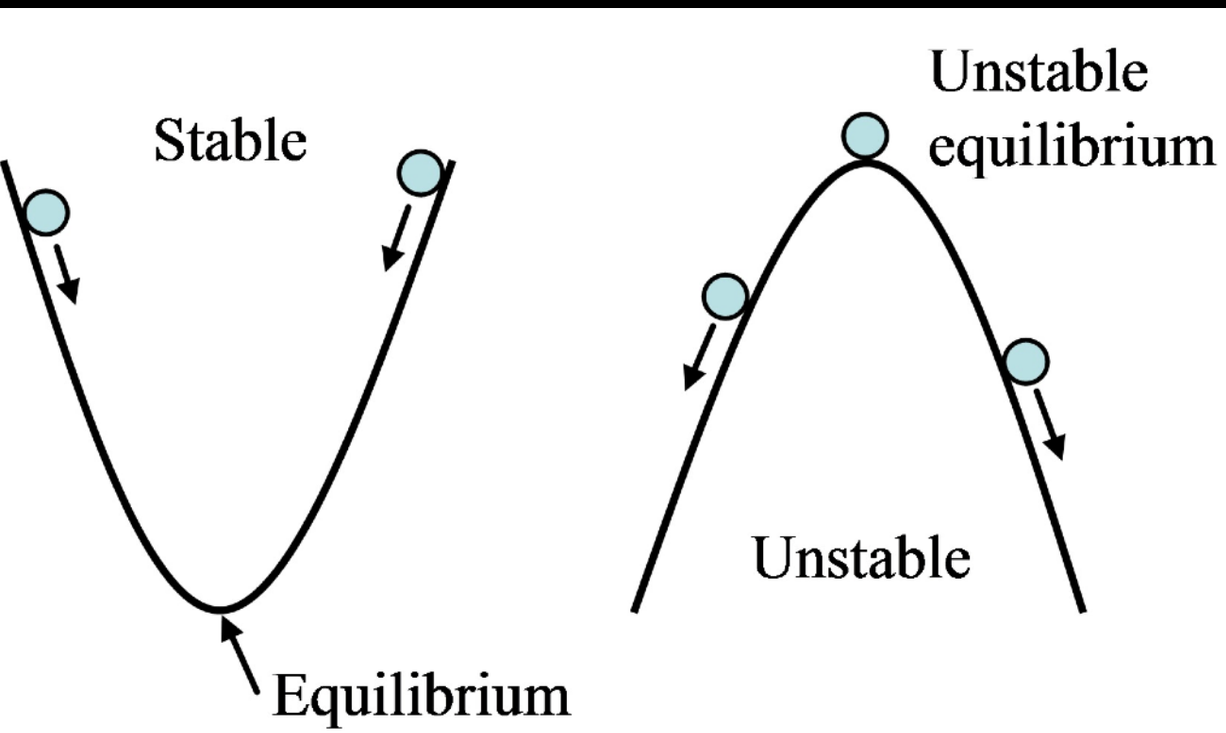


Logistic Growth



INTERNAL STABILITY

$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{\partial f}{\partial Y}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC \\ uncert}} +$$

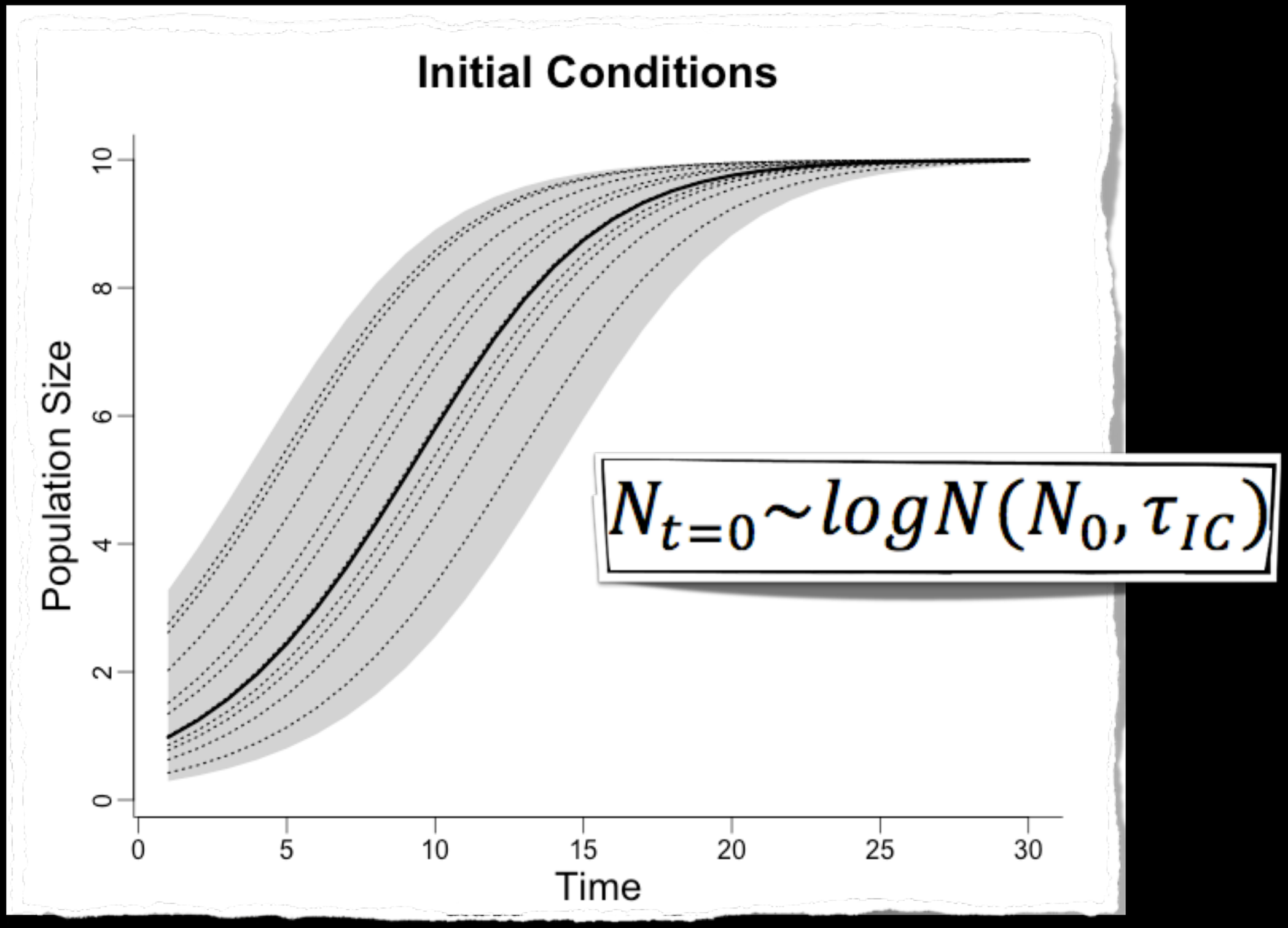


INTERNAL STABILITY

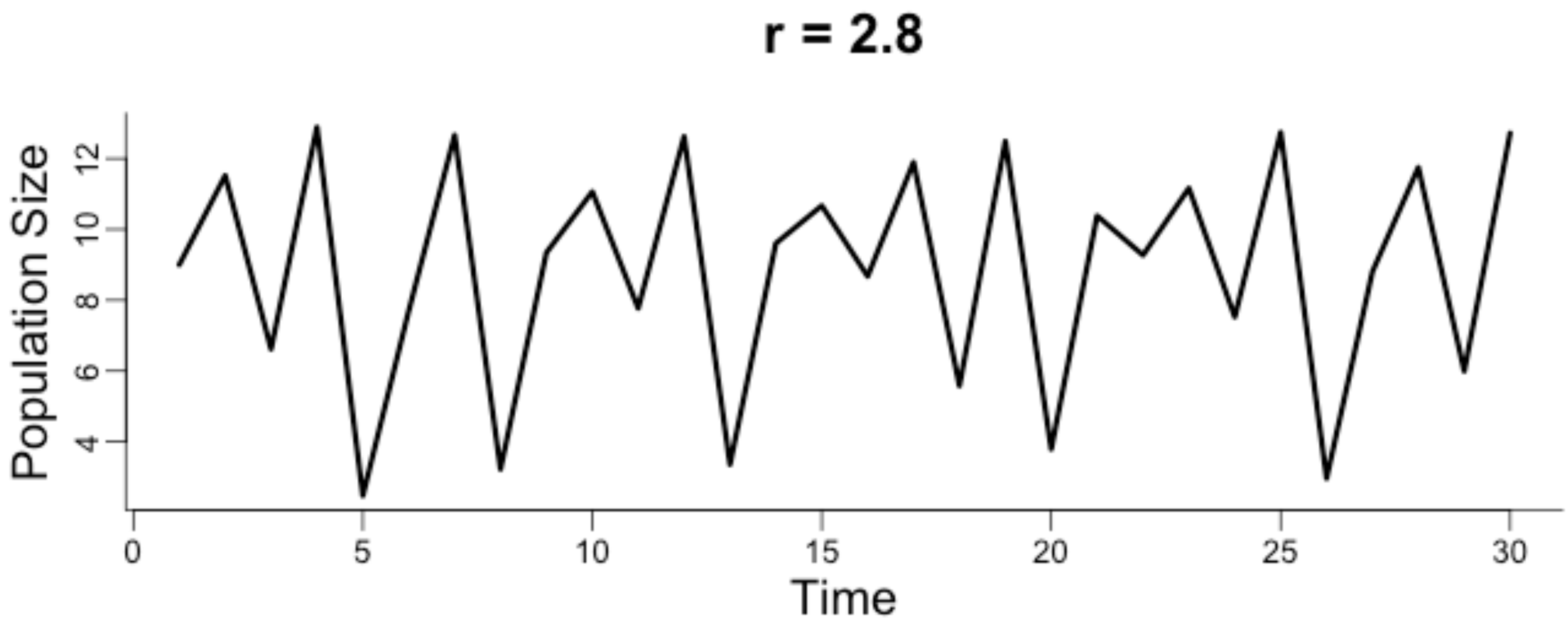
chaos
weather
epidemics

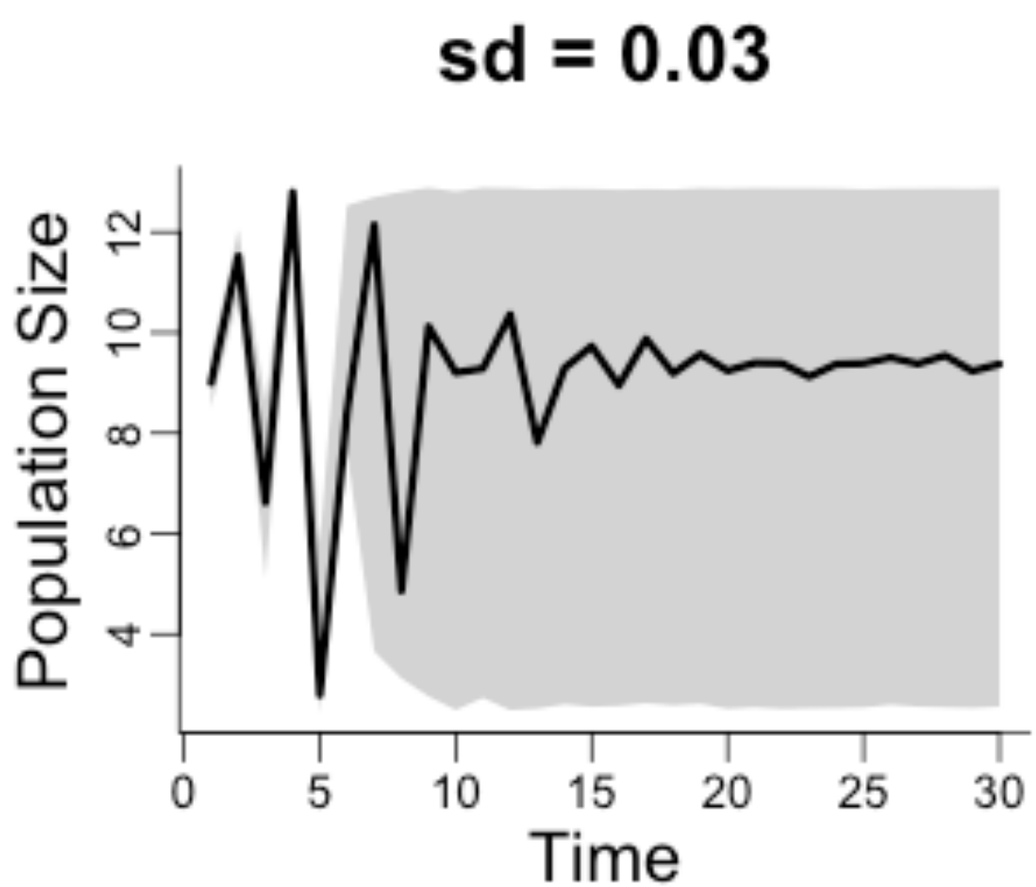
most
ecology

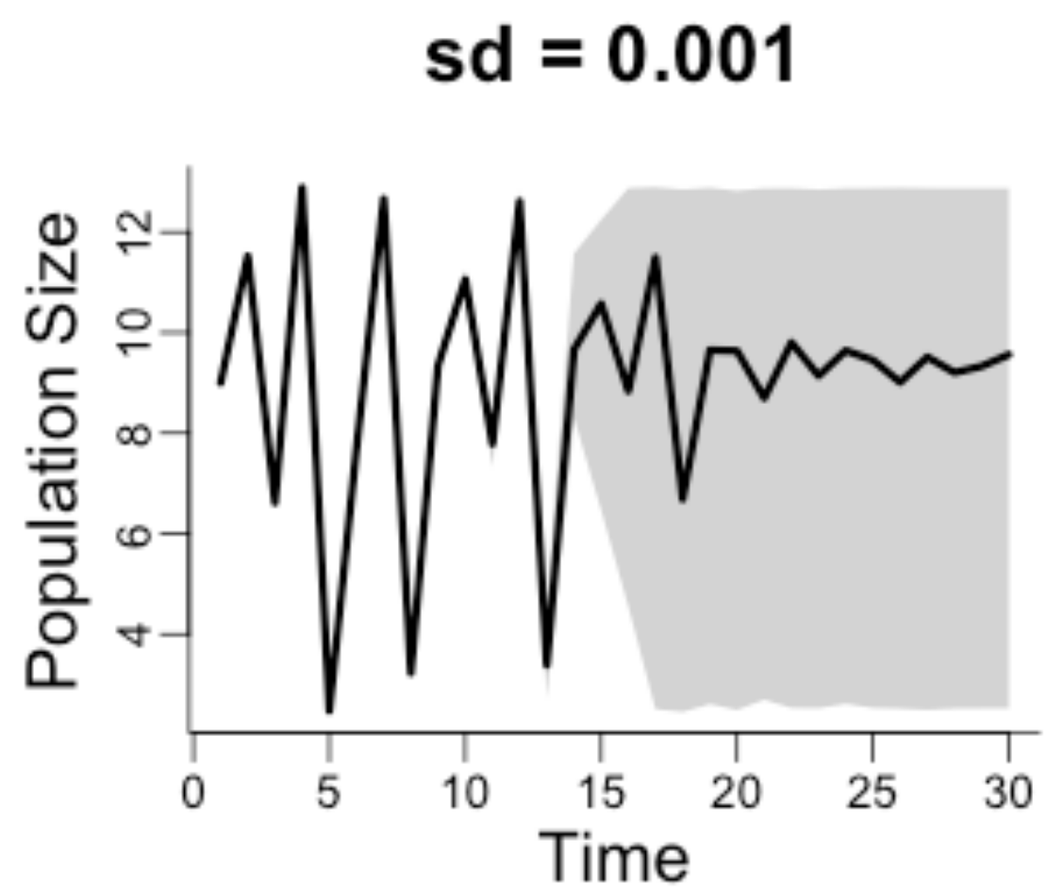
All other terms <u>grow linearly</u>



INITIAL CONDITIONS

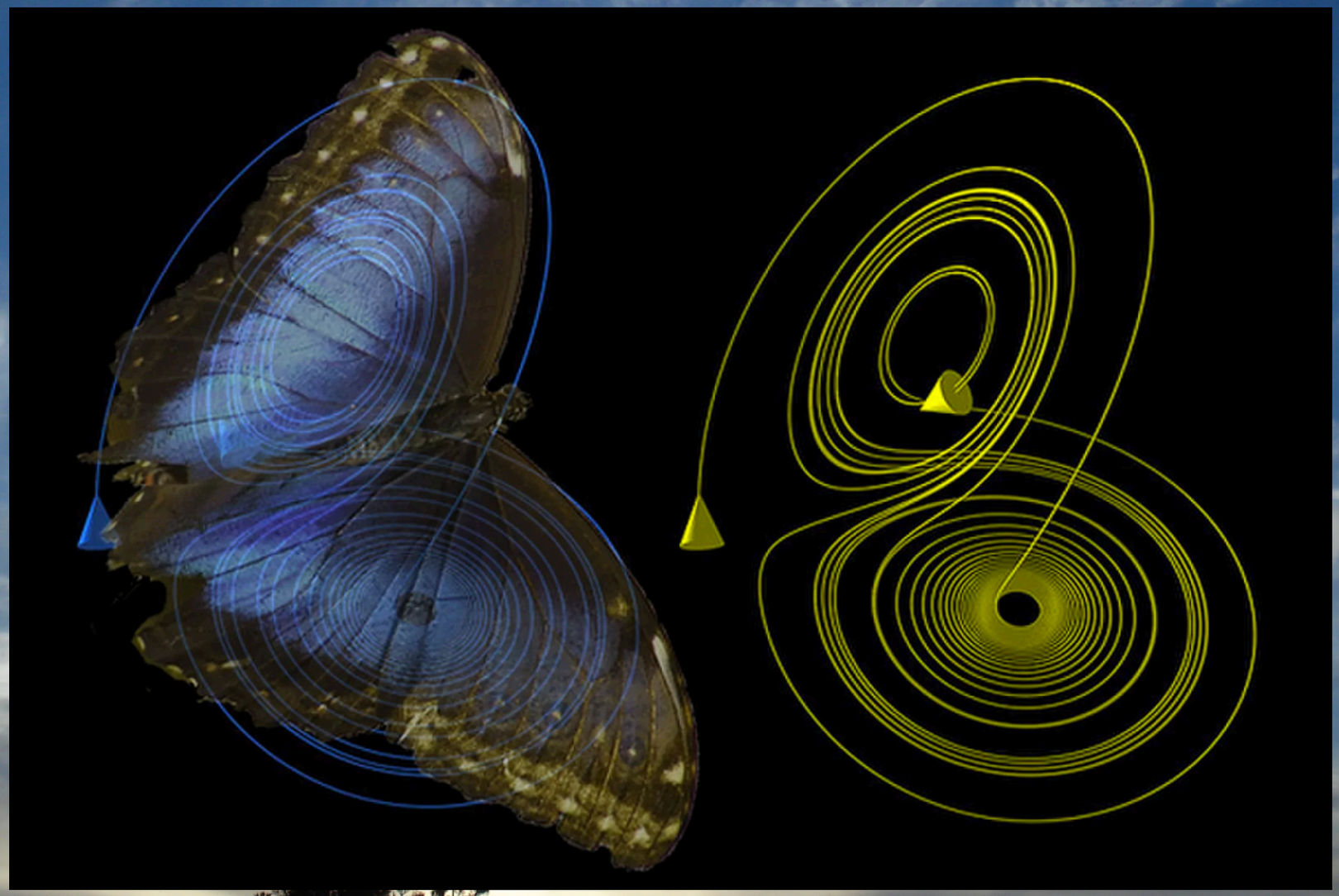

r = 2.8
Population Size
4
6
8
10
12
0
5
10
15
20
25
30
Time


sd = 0.03
Population Size
4
6
8
10
12
0
5
10
15
20
25
30
Time


sd = 0.001
Population Size
4
6
8
10
12
0
5
10
15
20
25
30
Time

WEATHER FORECASTING:
AN INITIAL CONDITIONS PROBLEM

$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{\partial f}{\partial Y}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC \\ uncert}}$$



EXOGENOUS STABILITY

$$+\underbrace{\left(\frac{\partial f}{\partial X}\right)^2}_{\substack{driver\\sens}}\underbrace{Var[X]}_{\substack{driver\\uncert}}+$$

- **Predictable if low sensitivity or low uncertainty**

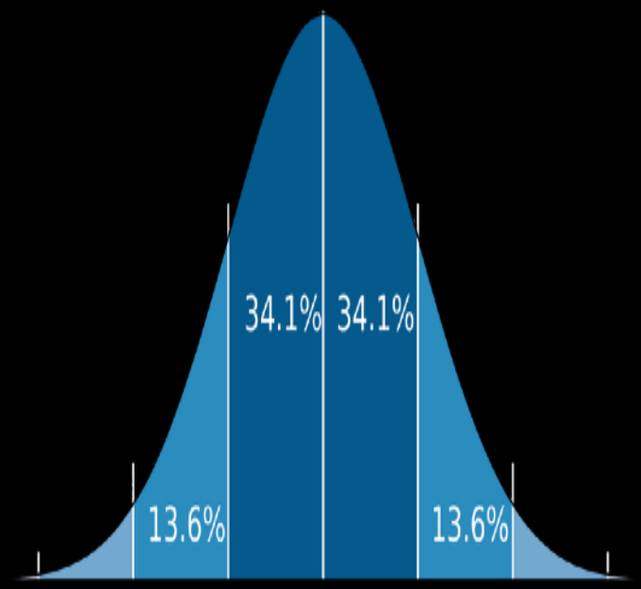
- Var[x] also needs to be forecast
  - **Rel. importance increases with time**
  - Different X for forecast than explain?
  - Not in model select, over complex

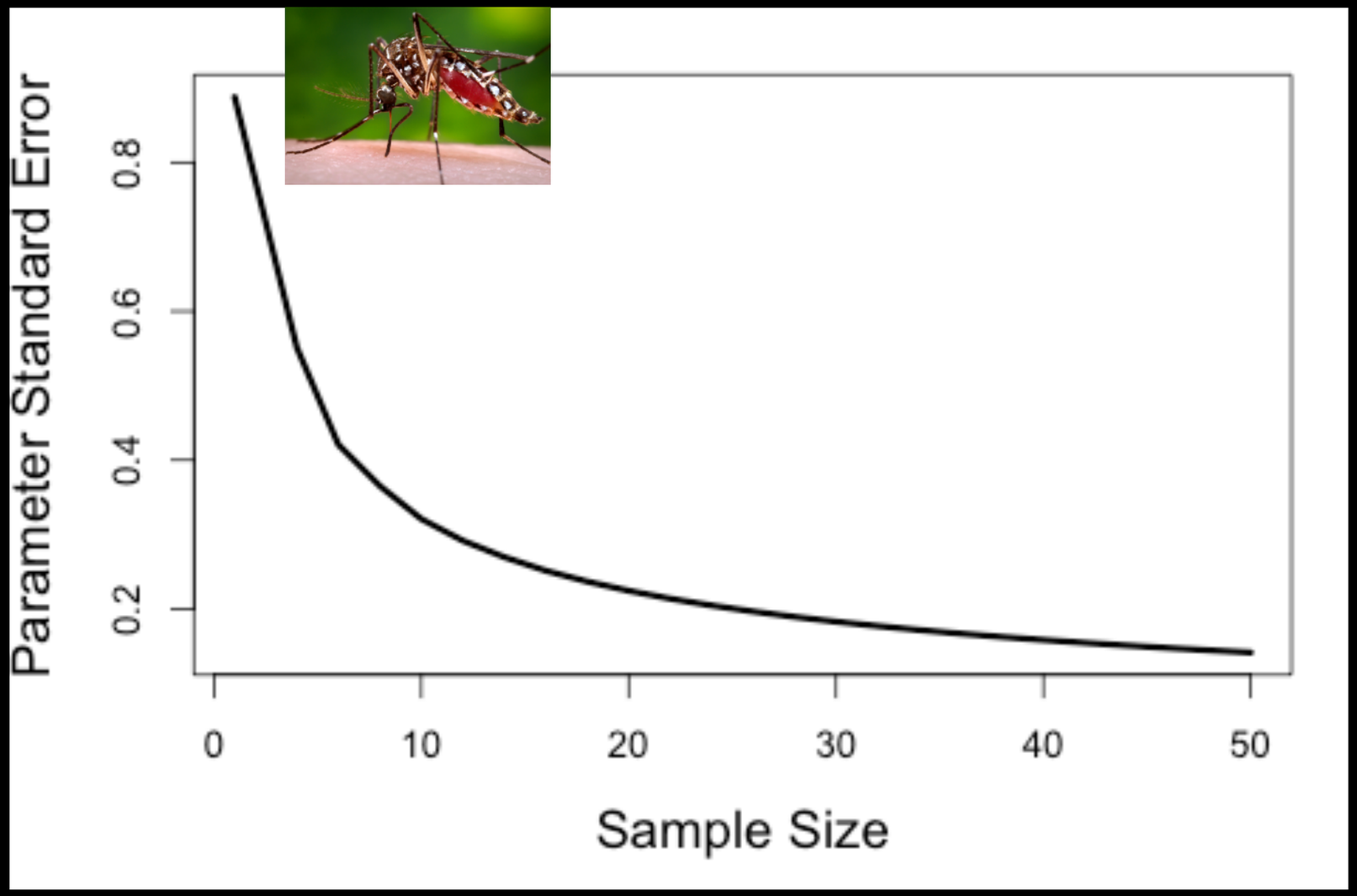
- Anova vs Regression design: *<u>How much</u> does X affect Y?*

- Endogenous (DD) vs Exogenous (DI) continuum

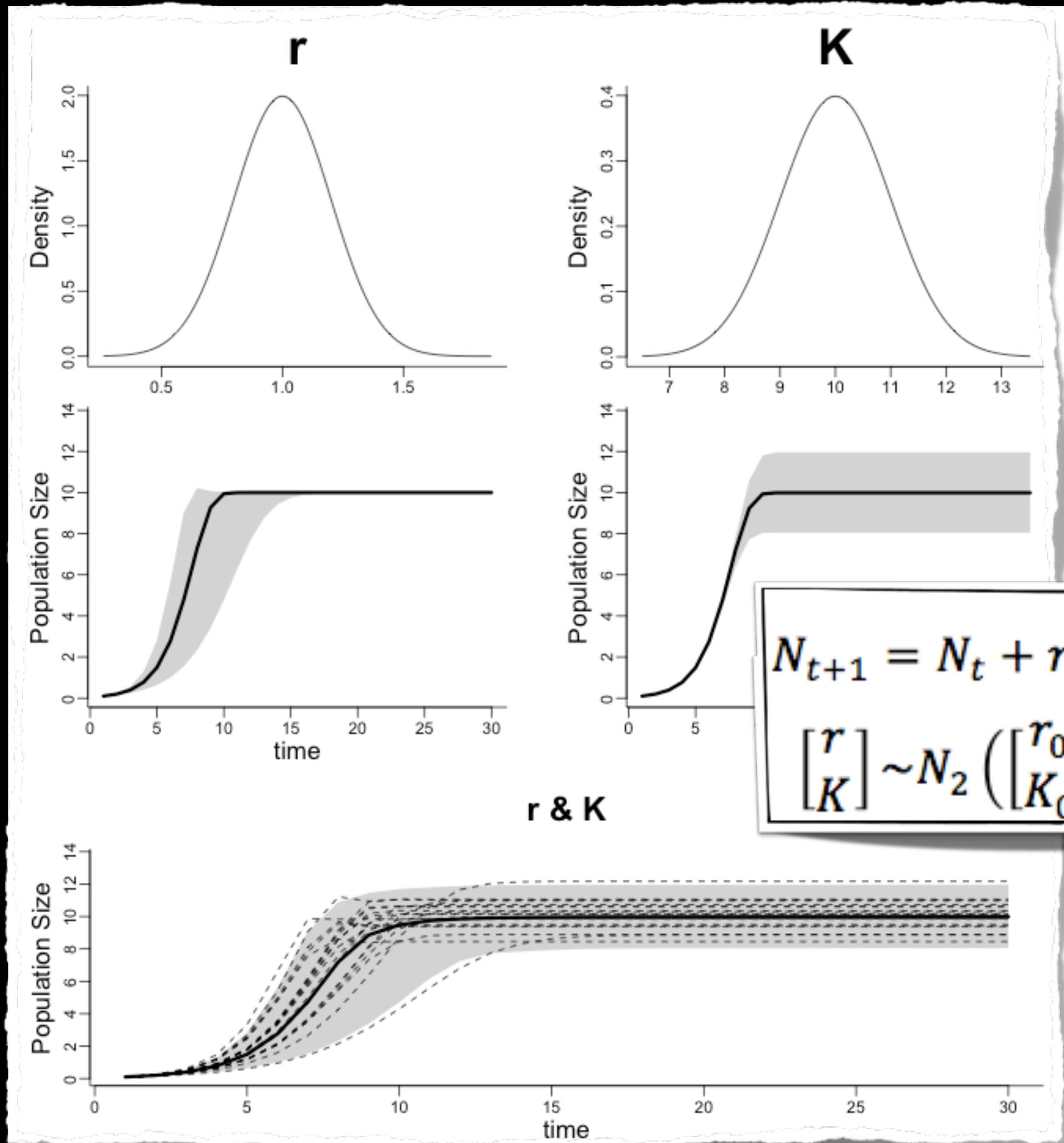
PARAMETER ERROR

$$+\underbrace{\left(\frac{\partial f}{\partial \theta}\right)^2}_{\substack{param\\sens}}\left(\underbrace{Var[\bar{\theta}]}_{\substack{param\\uncert}}+\underbrace{Var[\alpha]}_{\substack{param\\variability}}\right)+\underbrace{Var[\varepsilon]}_{\substack{process\\error}}$$





Parameter Error



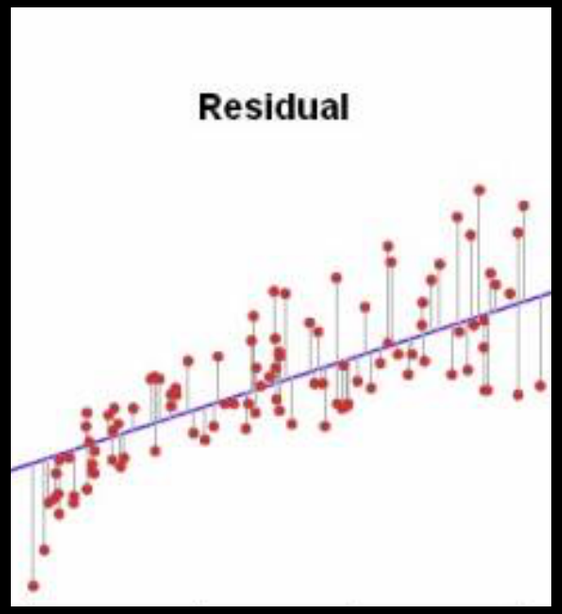
$$N_{t+1} = N_t + rN_t\left(1 - \frac{N_t}{K}\right)$$

$$\begin{bmatrix} r \\ K \end{bmatrix} \sim N_2\left(\begin{bmatrix} r_0 \\ K_0 \end{bmatrix}, \Sigma_{param}\right)$$

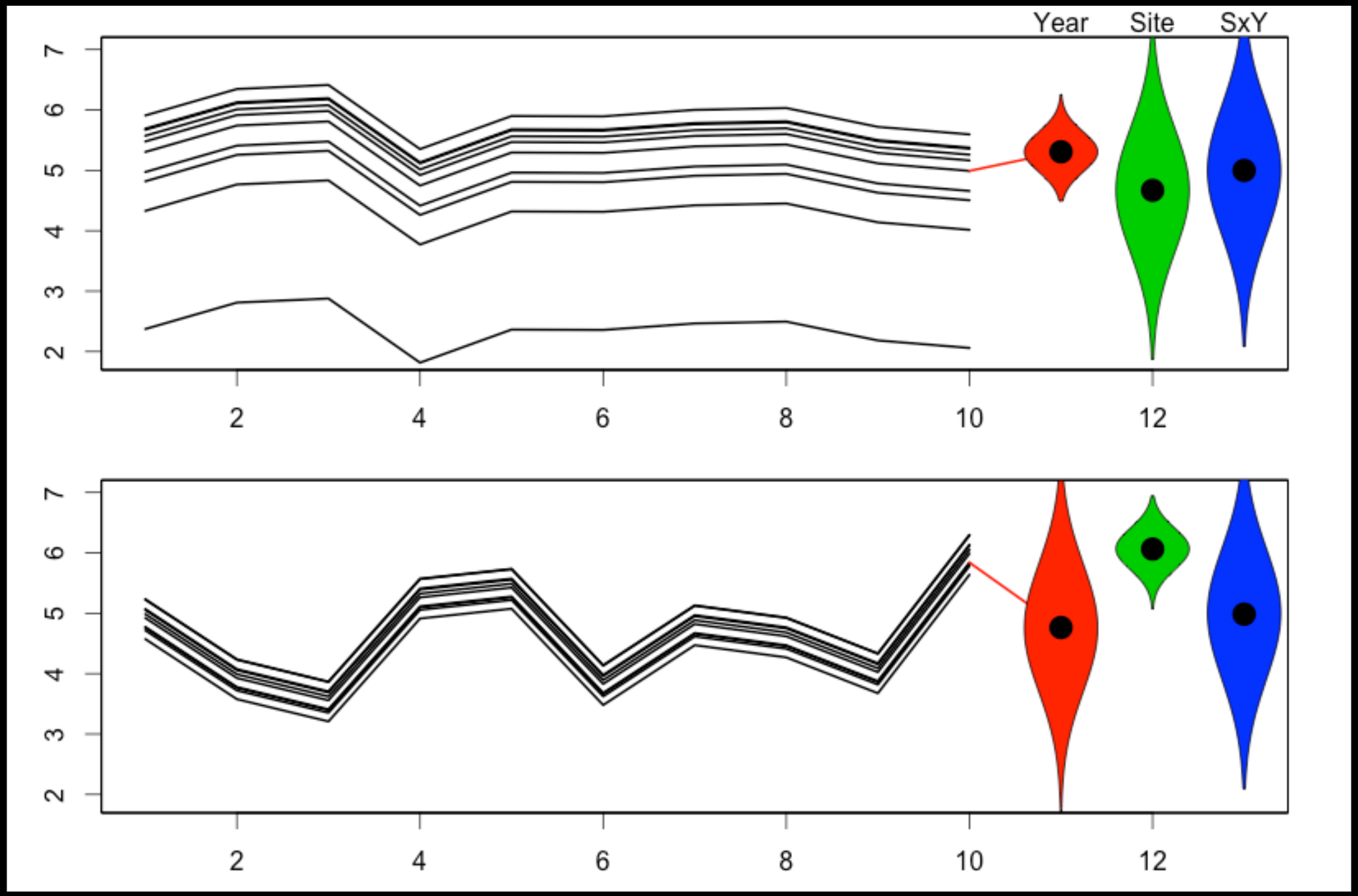
PROCESS ERROR

$$+\underbrace{\left(\frac{\partial f}{\partial \theta}\right)^2}_{\substack{param\\sens}}\left(\underbrace{Var[\bar{\theta}]}_{\substack{param\\uncert}}+\underbrace{Var[\alpha]}_{\substack{param\\variability}}\right)+\underbrace{Var[\varepsilon]}_{\substack{process\\error}}$$

- Inherent stochasticity (irreducible)

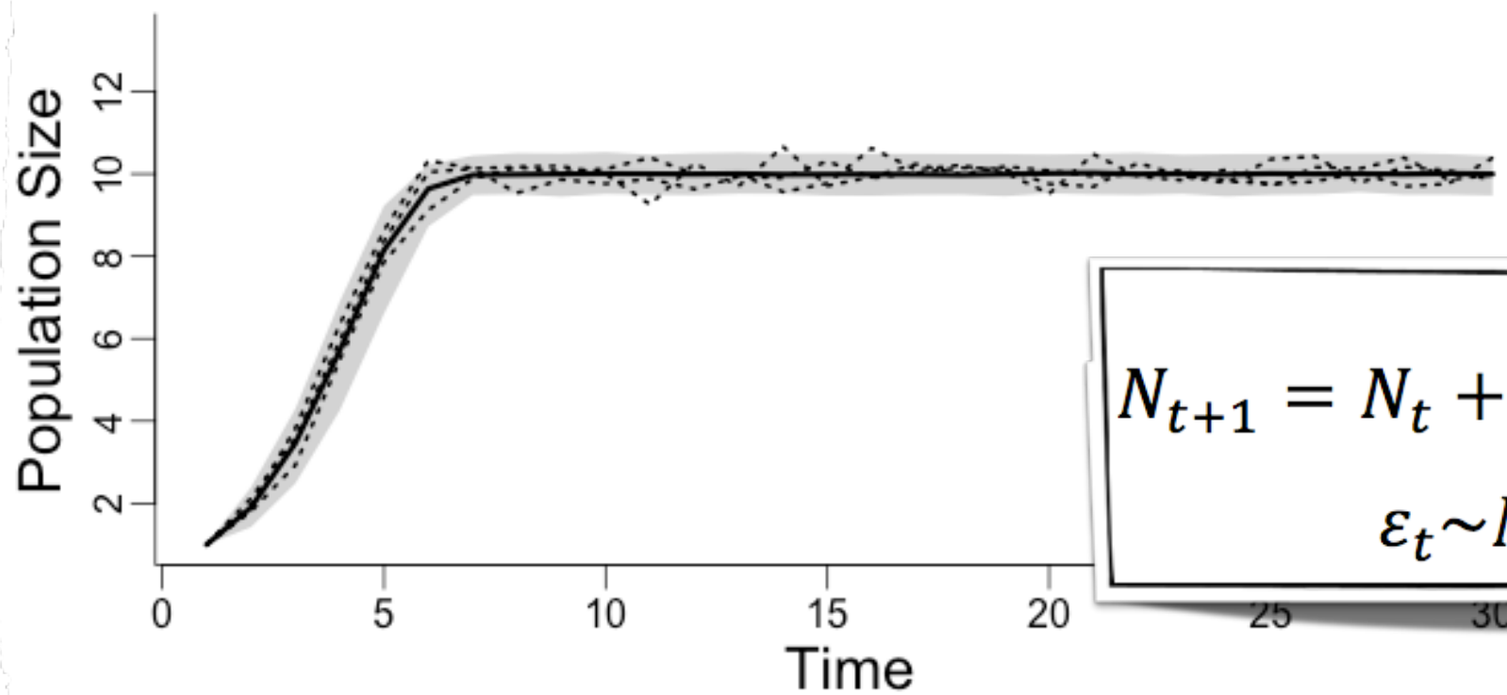


- Structural uncertainty
- Heterogeneity & variability
  - need to accommodate, even if can't explain


Year
Site
SxY
7
6
5
4
3
2
2
4
6
8
10
12
7
6
5
4
3
2
2
4
6
8
10
12

Process Error

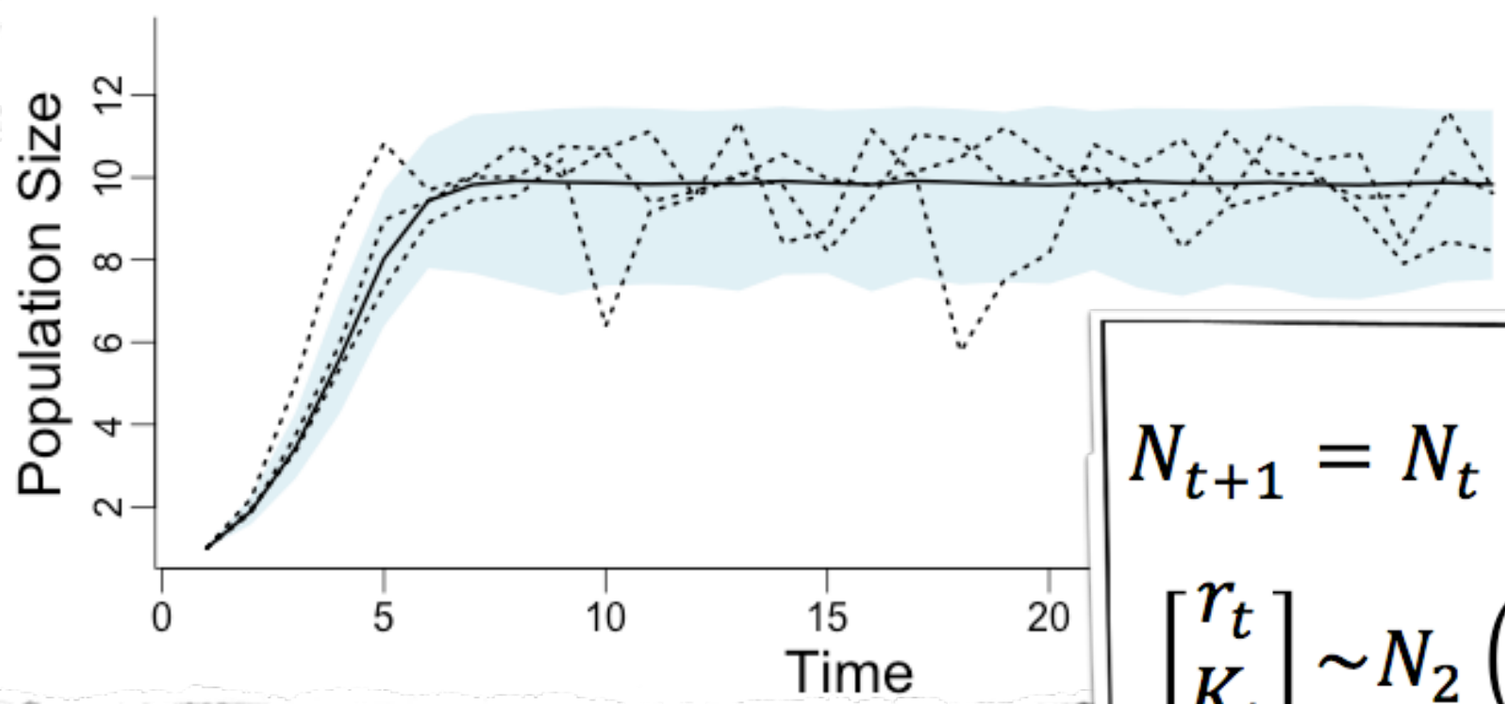
Additive Variability



$$N_{t+1} = N_t + rN_t\left(1 - \frac{N_t}{K}\right) + \varepsilon_t$$

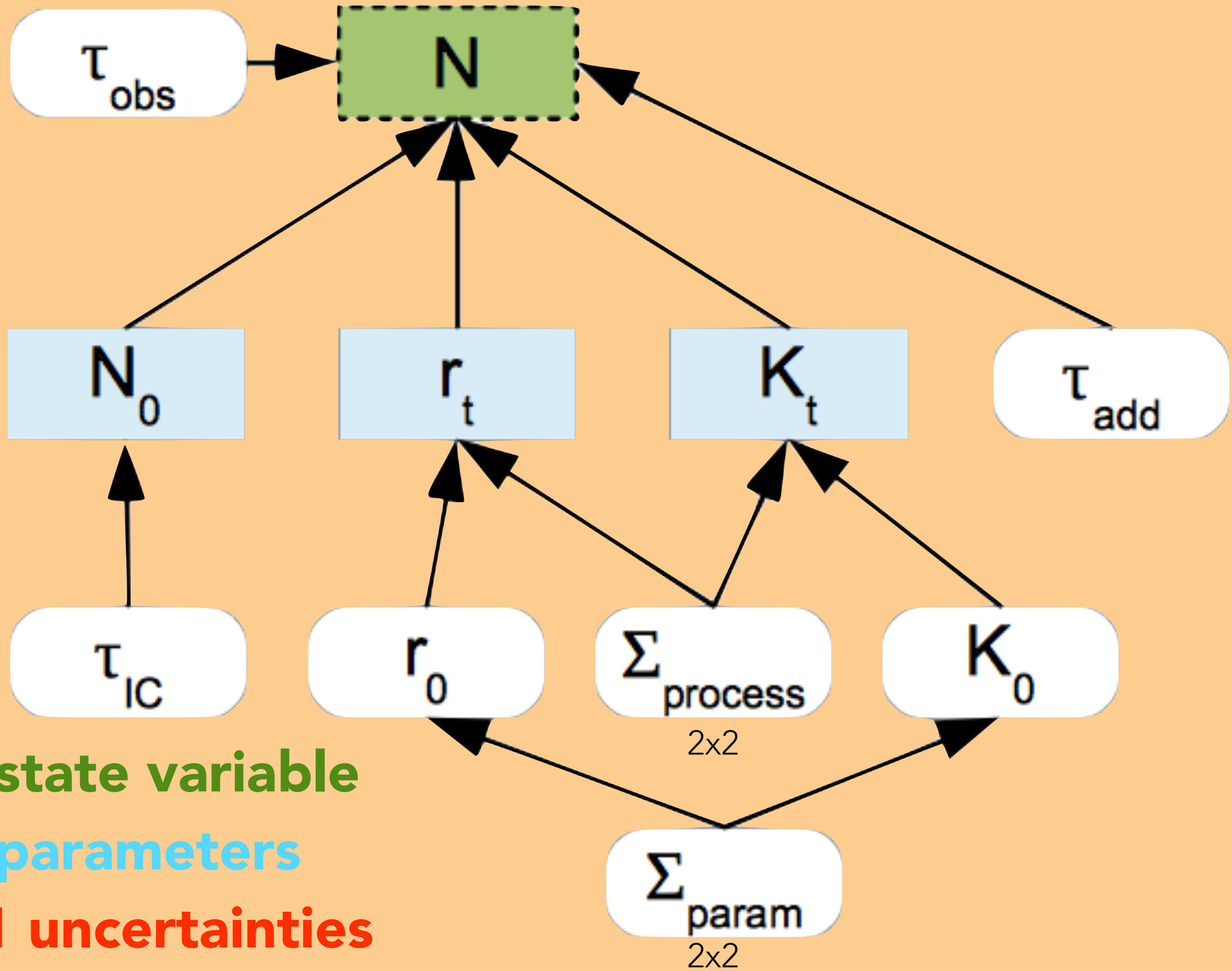
$$\varepsilon_t \sim N(0, \tau_{add})$$

Parameter Variability



$$N_{t+1} = N_t + r_t N_t\left(1 - \frac{N_t}{K_t}\right)$$

$$\begin{bmatrix} r_t \\ K_t \end{bmatrix} \sim N_2\left(\begin{bmatrix} r_0 \\ K_0 \end{bmatrix}, \Sigma_{process}\right)$$



1 state variable

3 parameters

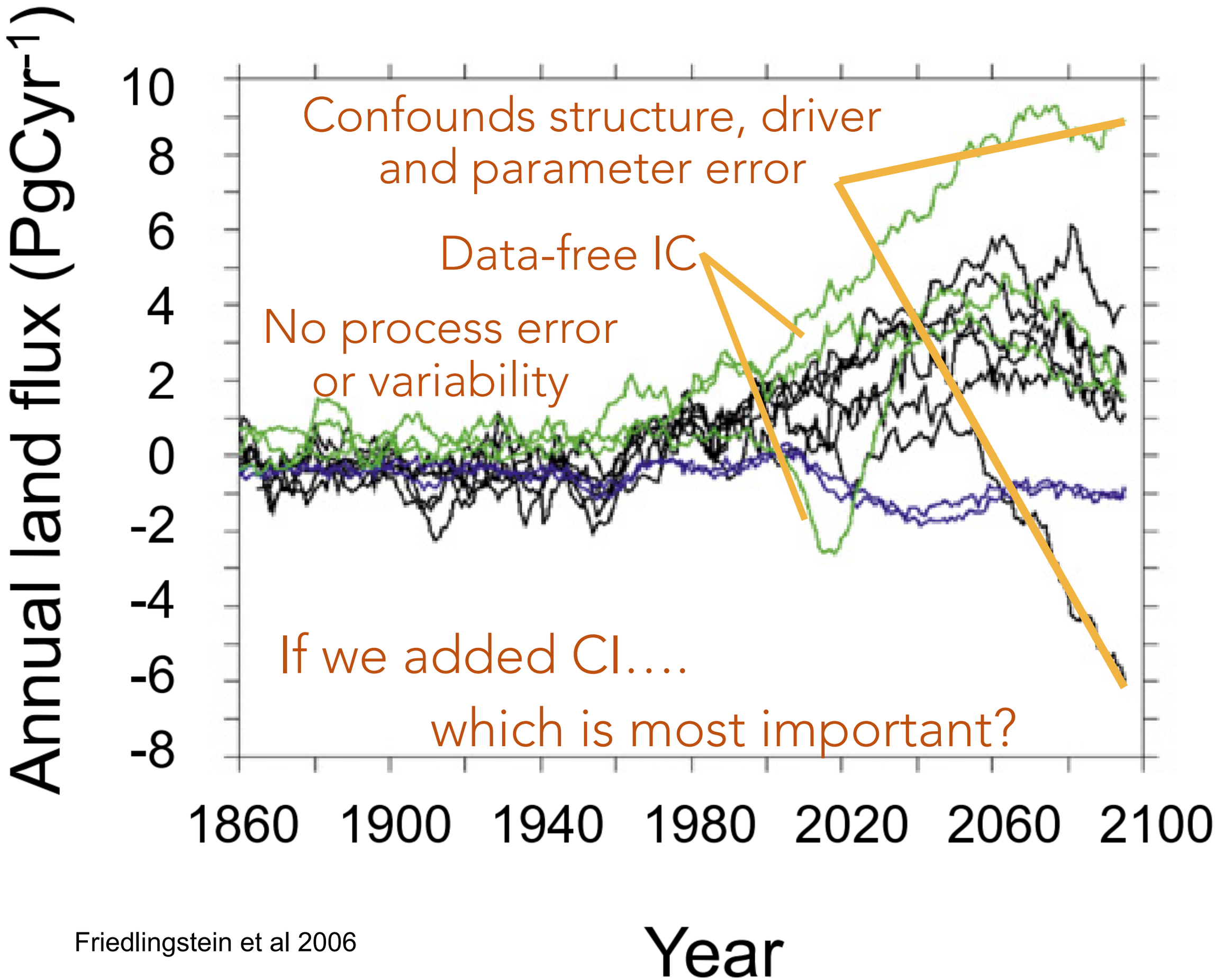
11 uncertainties

THINK DISTRIBUTIONS !!

COV & SCALING

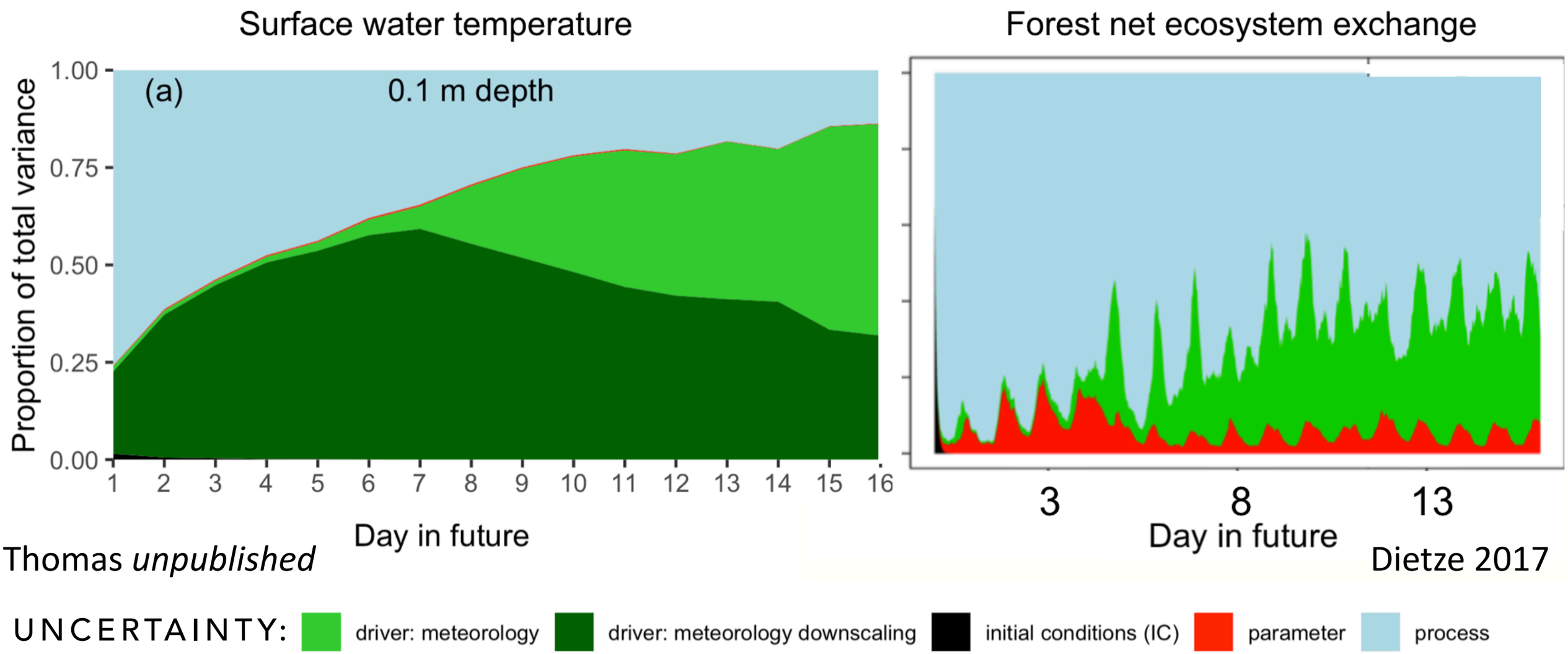
- At large scales, average over drivers (X), heterogeneity (α), & variability (ε)
  - Internal stability (Y) increases in importance
  - Scaling very dependent on spatial and temporal auto- & cross-correlation

$$\sum\sum \frac{\partial f}{\partial X_i}\frac{\partial f}{\partial X_j} COV[X_i, X_j]$$



Friedlingstein et al 2006

How do the drivers of forecast uncertainty vary across ecological system?



Thomas *unpublished*

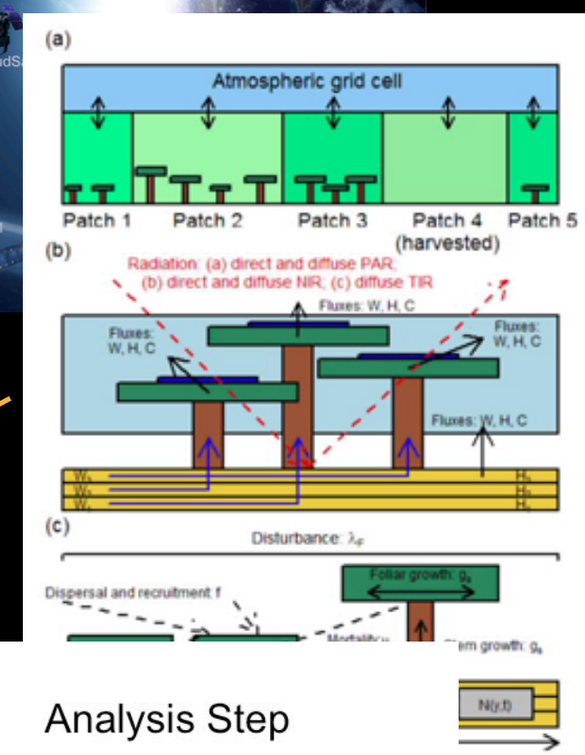
Dietze 2017

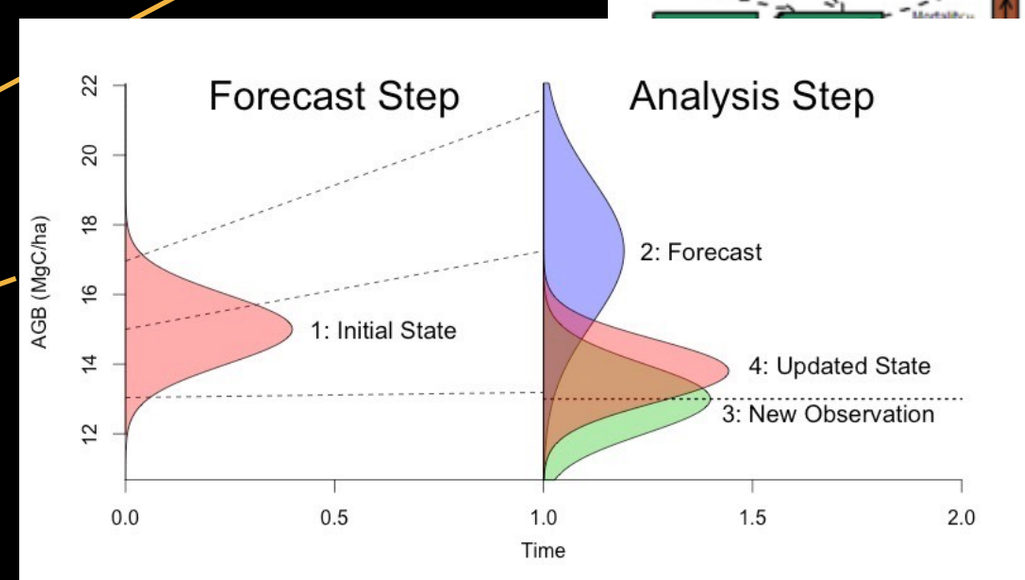
NATURE OF THE PREDICTION PROBLEM…

$$Var[Y_{t+1}] \approx \underbrace{\left(\frac{\partial f}{\partial Y}\right)^2}_{stability} \underbrace{Var[Y_t]}_{\substack{IC\\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial X}\right)^2}_{\substack{driver\\ sens}} \underbrace{Var[X]}_{\substack{driver\\ uncert}} + \underbrace{\left(\frac{\partial f}{\partial \theta}\right)^2}_{param} \left(\underbrace{Var[\bar{\theta}]}_{\substack{param\\ uncert}} + \underbrace{Var[\alpha]}_{\substack{param\\ variability}} \right) + \underbrace{Var[\varepsilon]}_{\substack{process\\ error}}$$

$$= INTERNAL + EXTERNAL + I$$

- Theory
  - What drives dynamics?
  - Generality across processes and locations
- Practice
  - What can we predict?
  - How to tackle new systems
- Methods
  - What to measure
  - How we build models
  - How we assimilate data





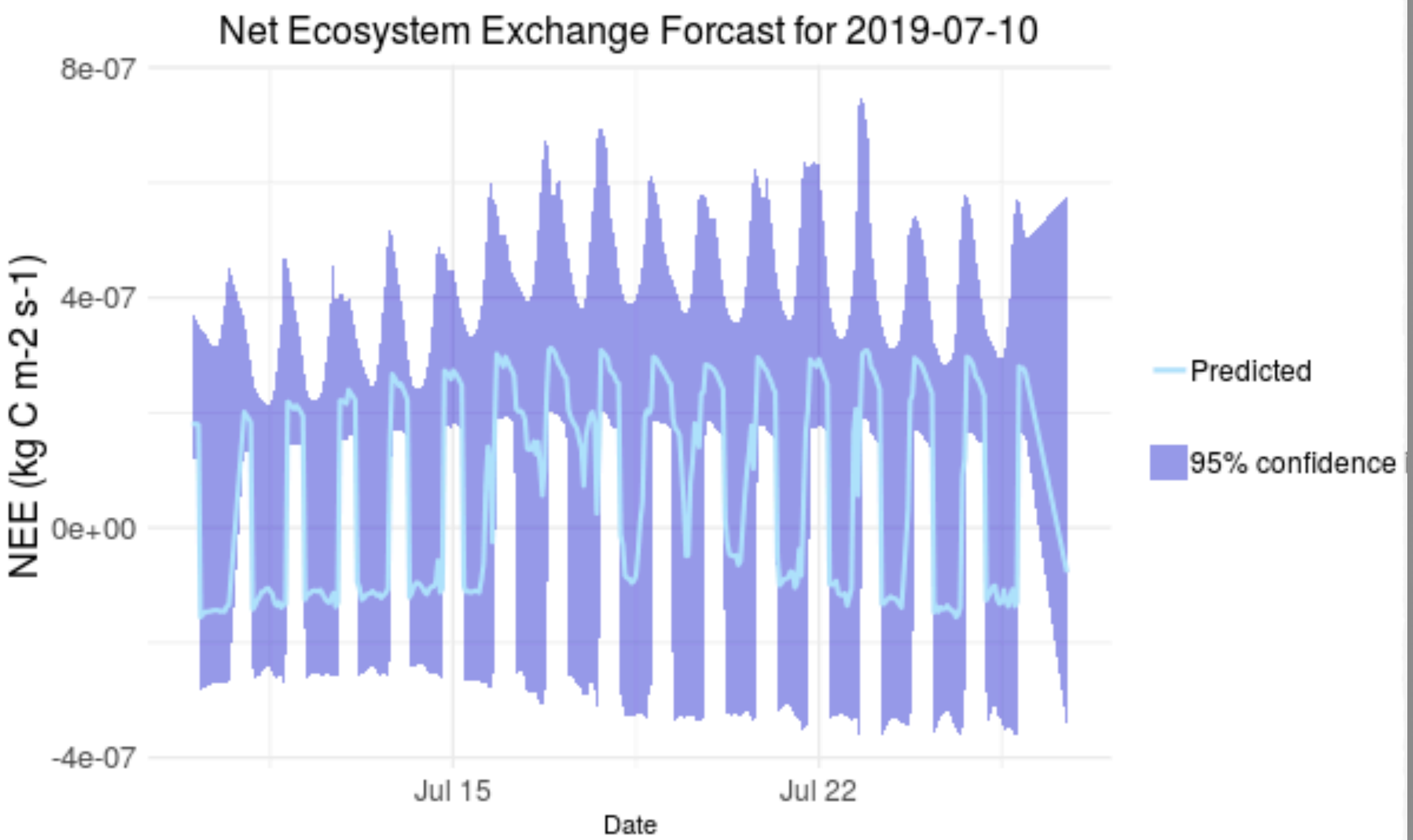
DISCOVER WHETHER NATURE IS PREDICTABLE

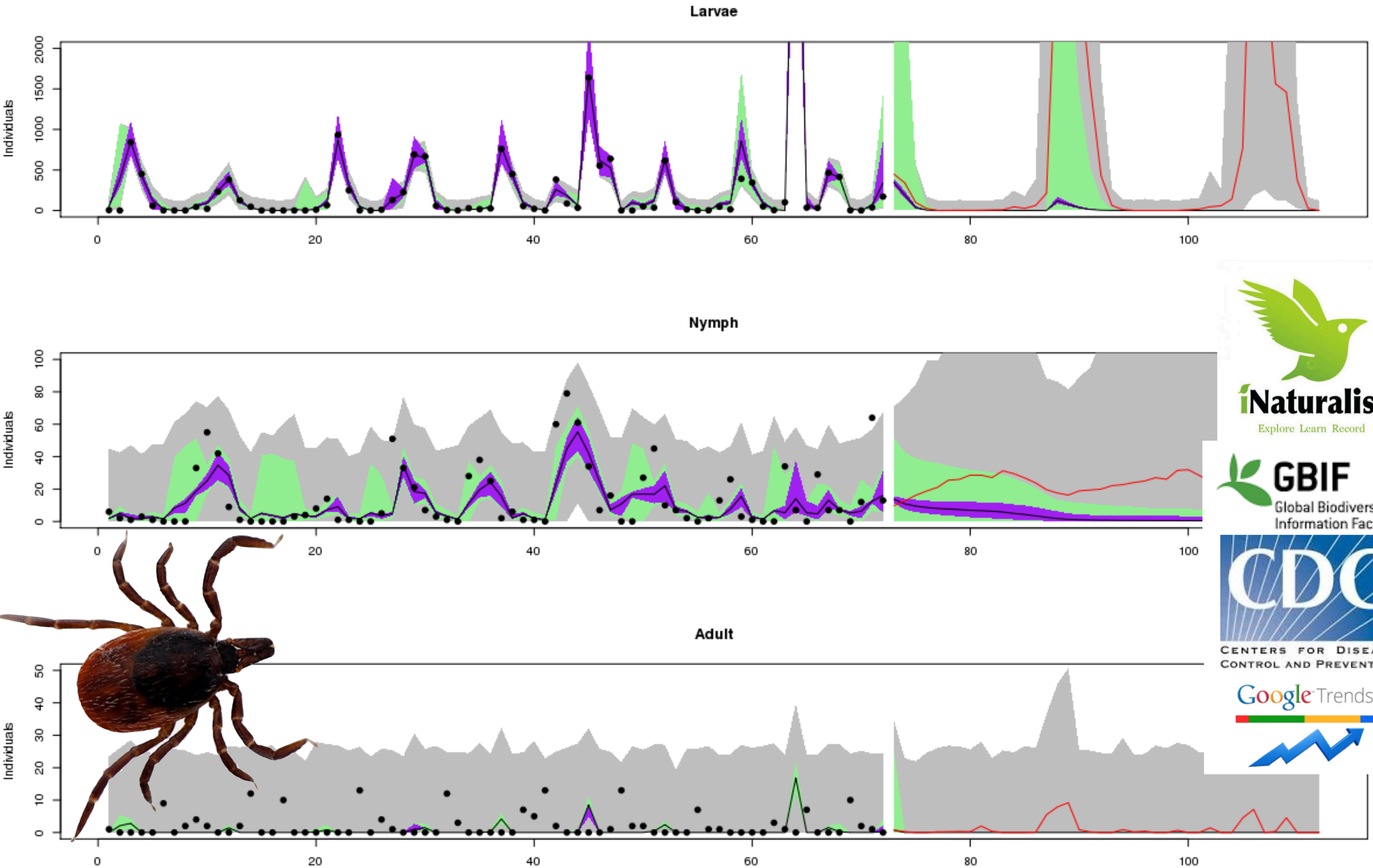

Dietze & Wheeler:
Fluxes & Phenology
Weathers: Aquatic Productivity
Near-term Ecological Forecasting Initiative
LaDeau & Foster:
Ticks & Small Mammals
Talbot & Averill:
Microbial diversity
neon
National Ecological Observatory Network

Willow Creek Forecast for 2019-07-10

To: Michael Dietze

Below is today's 16 day forecast for NEE, LE, and Soil Moisture Fraction at Willow Creek, Wisconsin.



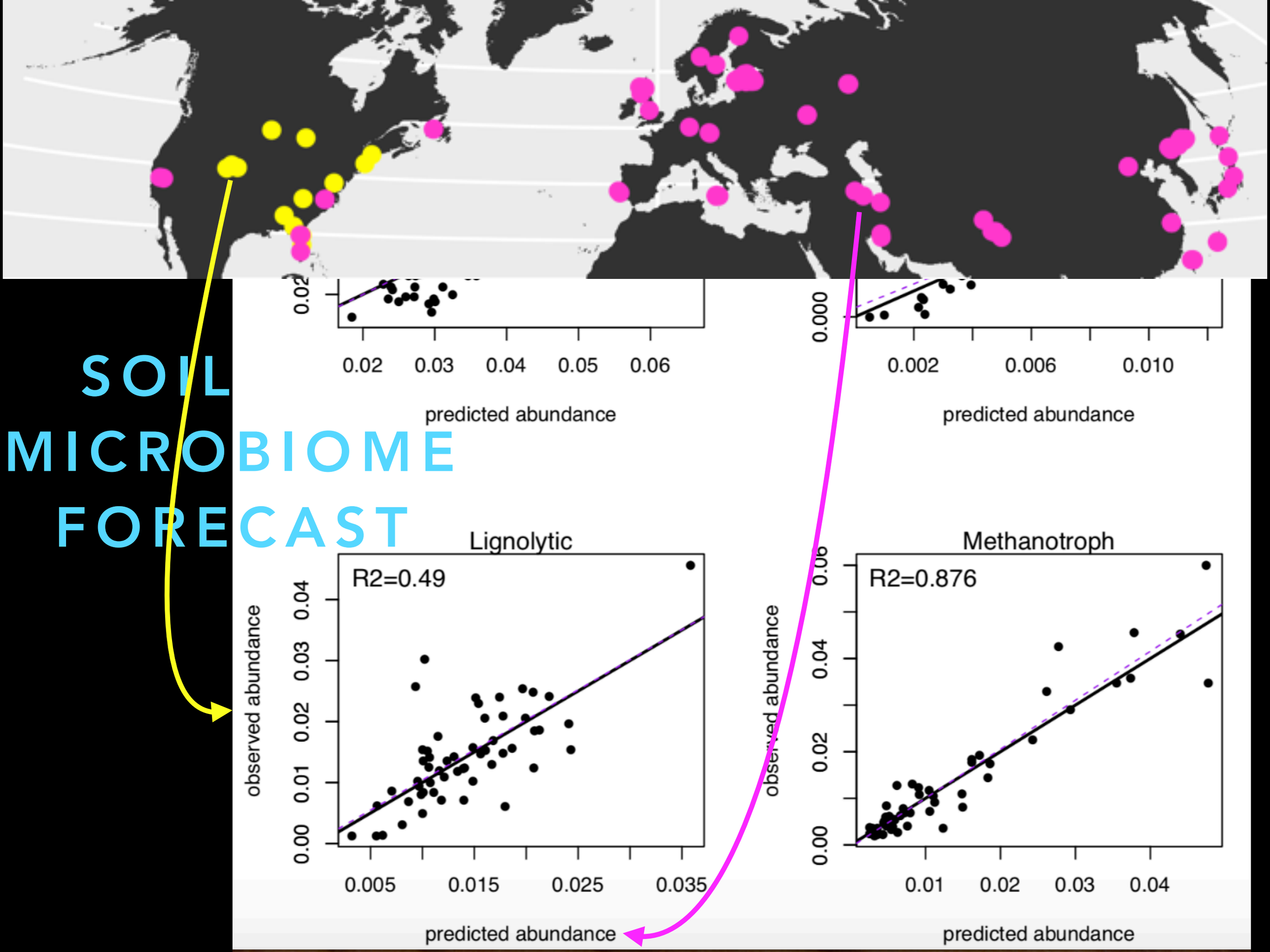


iNaturalist
Explore Learn Record

GBIF
Global Biodiversity Information Facility

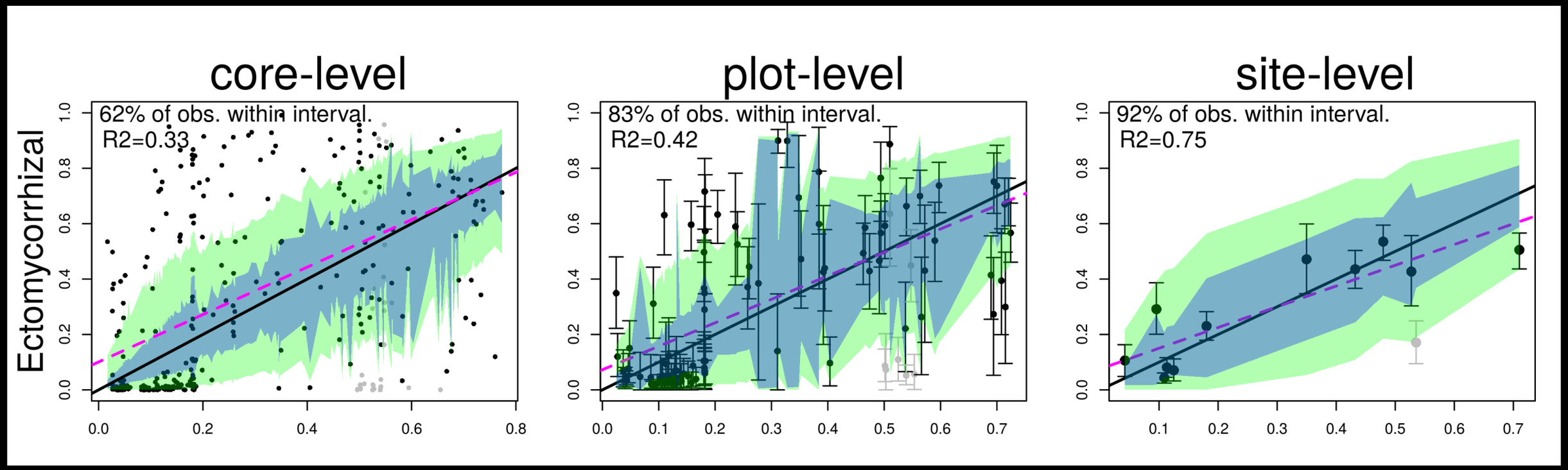
CDC
Centers for Disease Control and Prevention

Google Trends

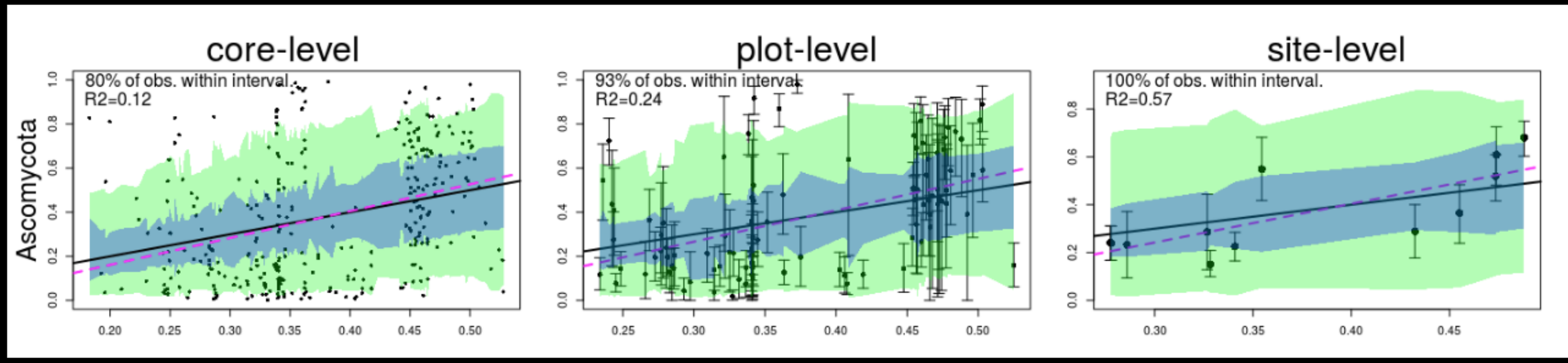

SOIL MICROBIOME FORECAST
0.02
0.03
0.04
0.05
0.06
predicted abundance
0.000
0.002
0.006
0.010
predicted abundance
Lignolytic
R2=0.49
observed abundance
0.00
0.01
0.02
0.03
0.04
0.005
0.015
0.025
0.035
predicted abundance
Methanotroph
R2=0.876
observed abundance
0.00
0.02
0.04
0.01
0.02
0.03
0.04
predicted abundance

PREDICTABILITY INCREASES WITH SCALE

…both for functional groups



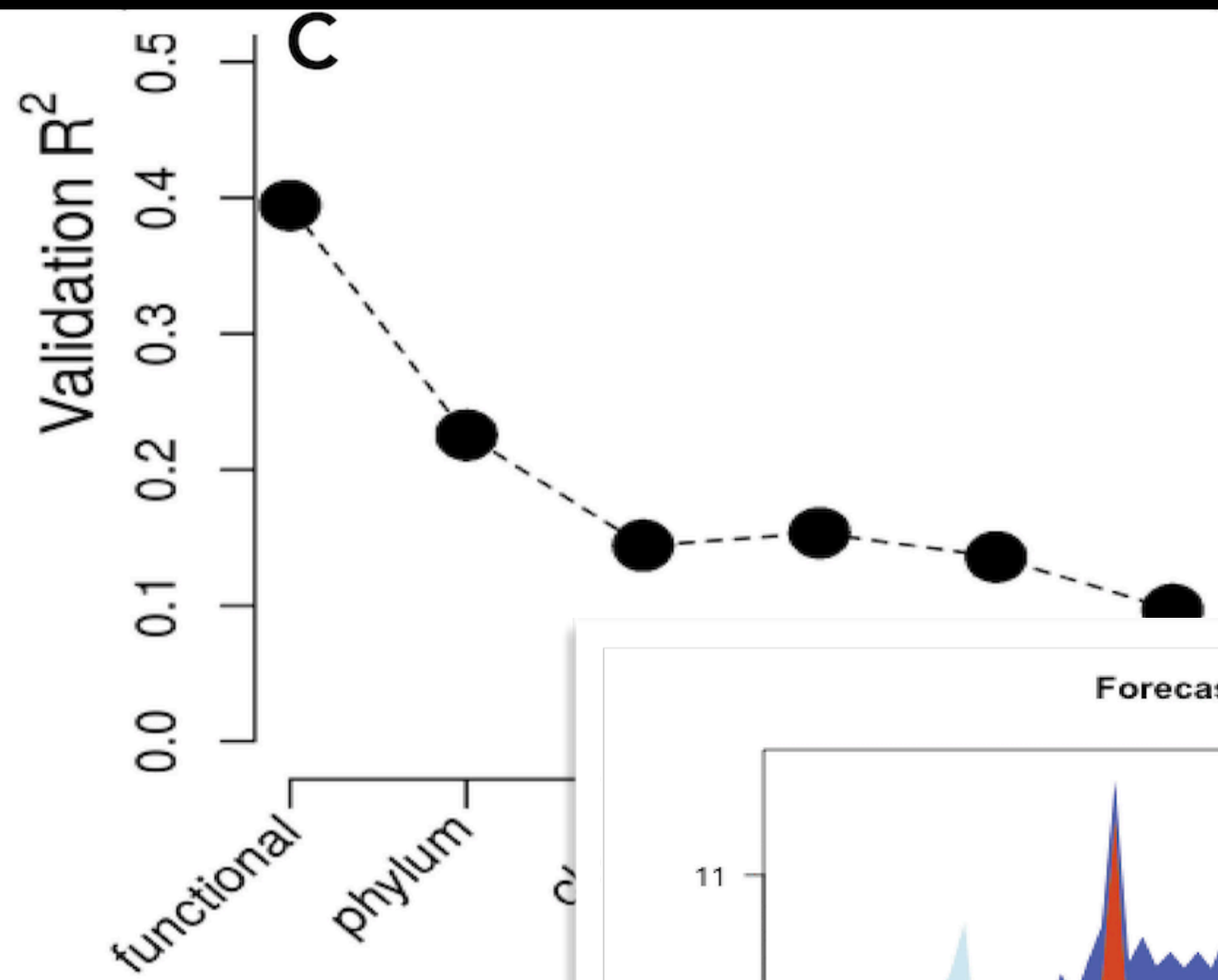
…and for higher-level taxonomic groups

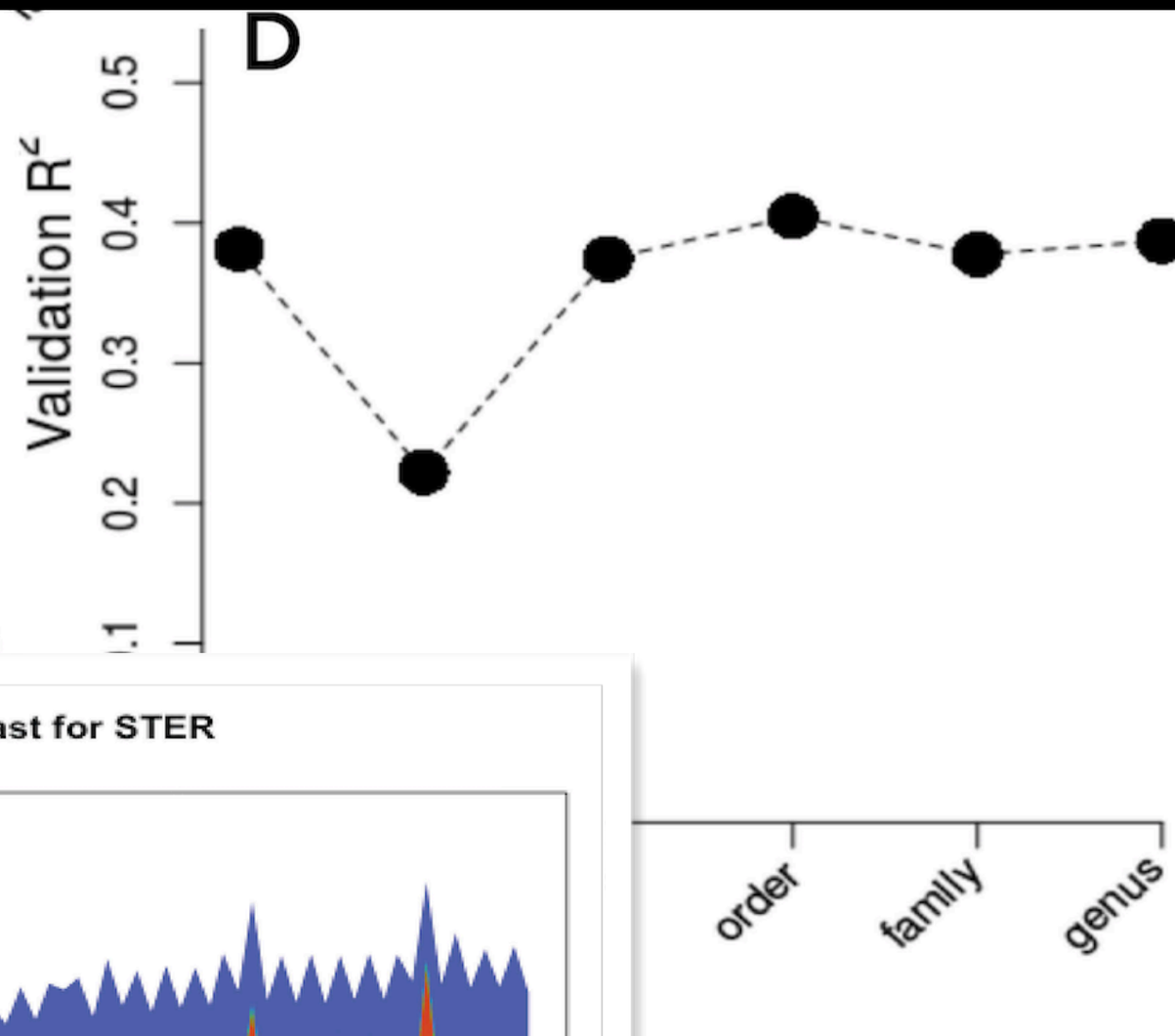


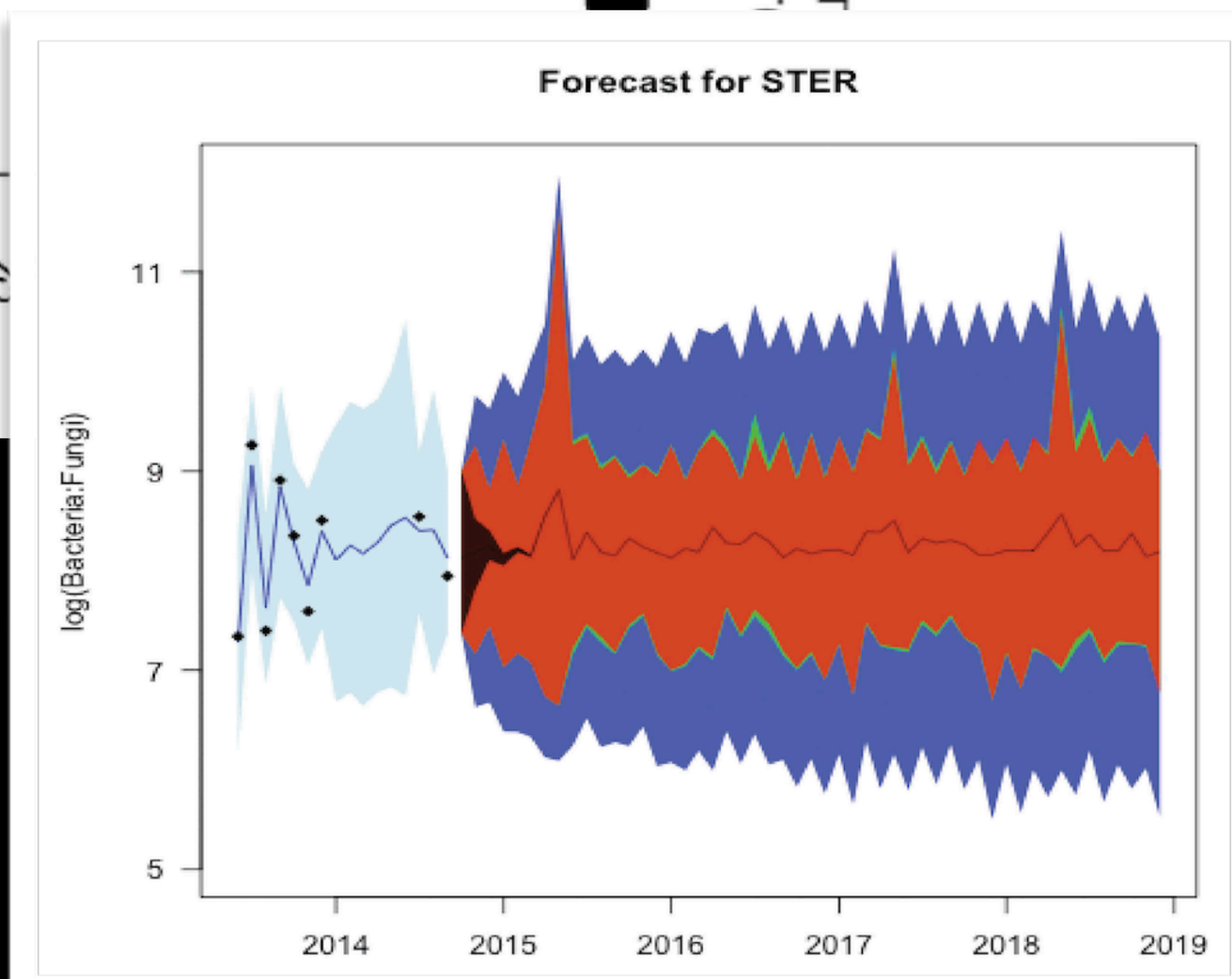
Out-of-sample validation at NEON sites

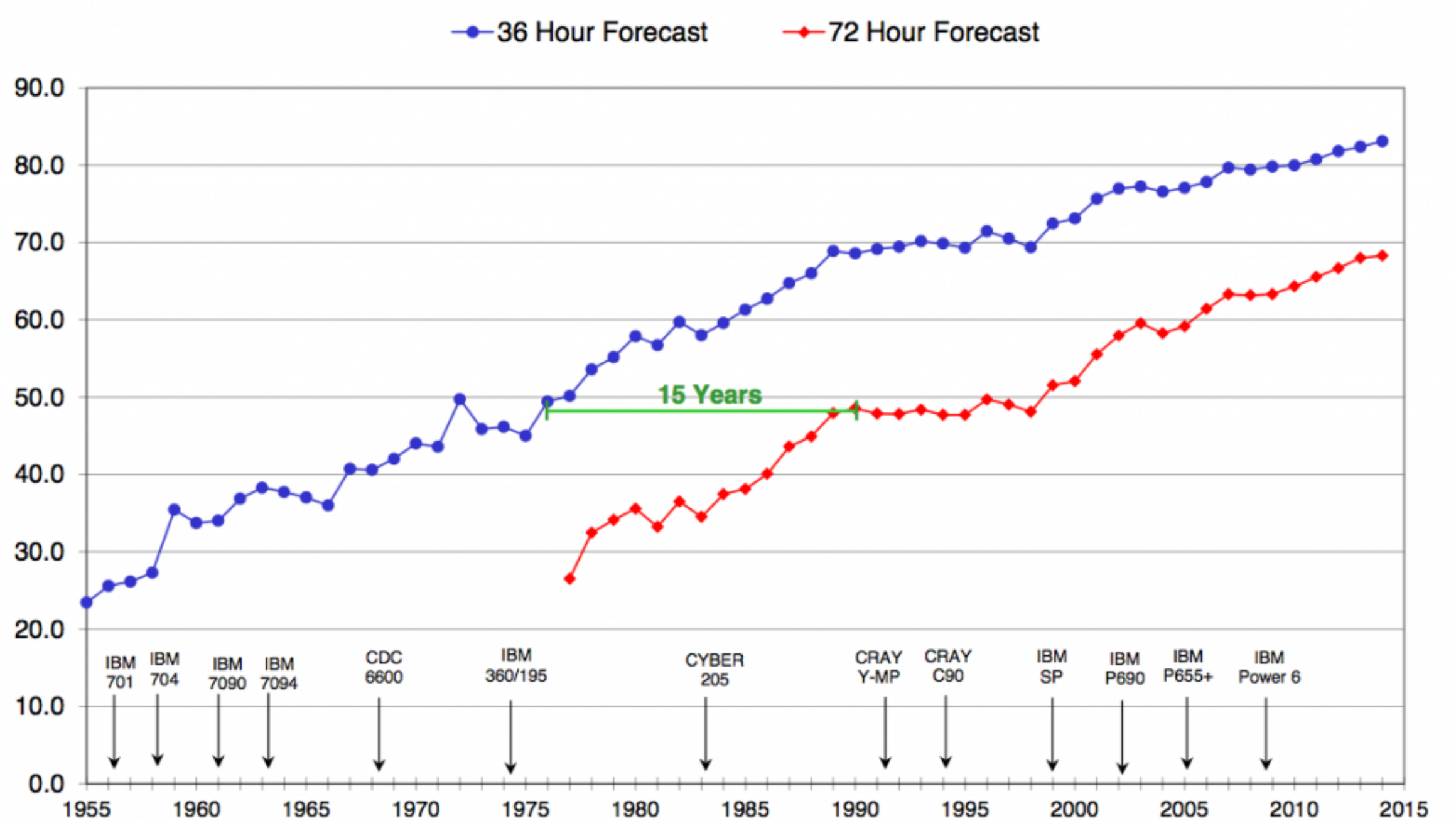
FUNGI

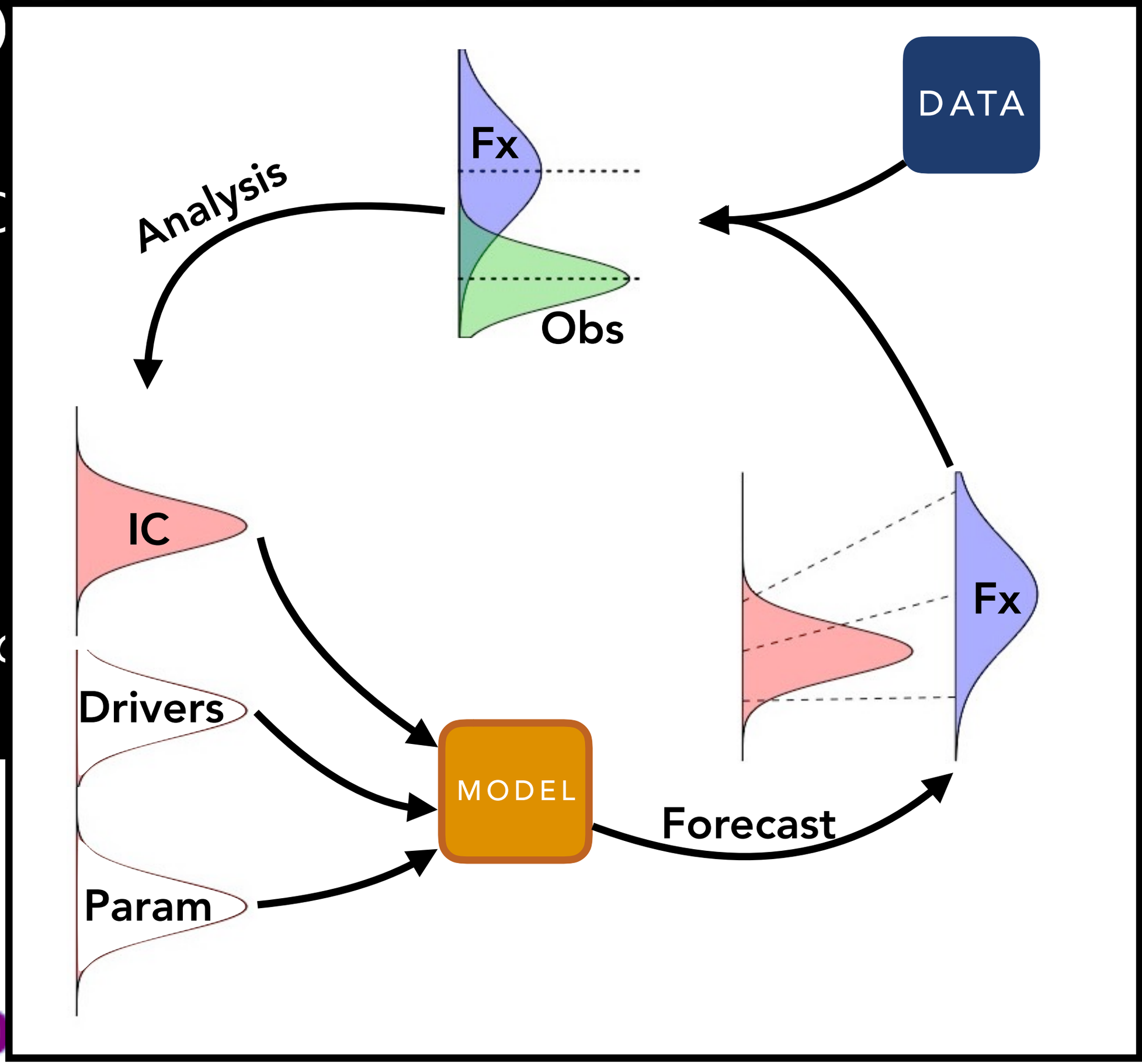
BACTERIA

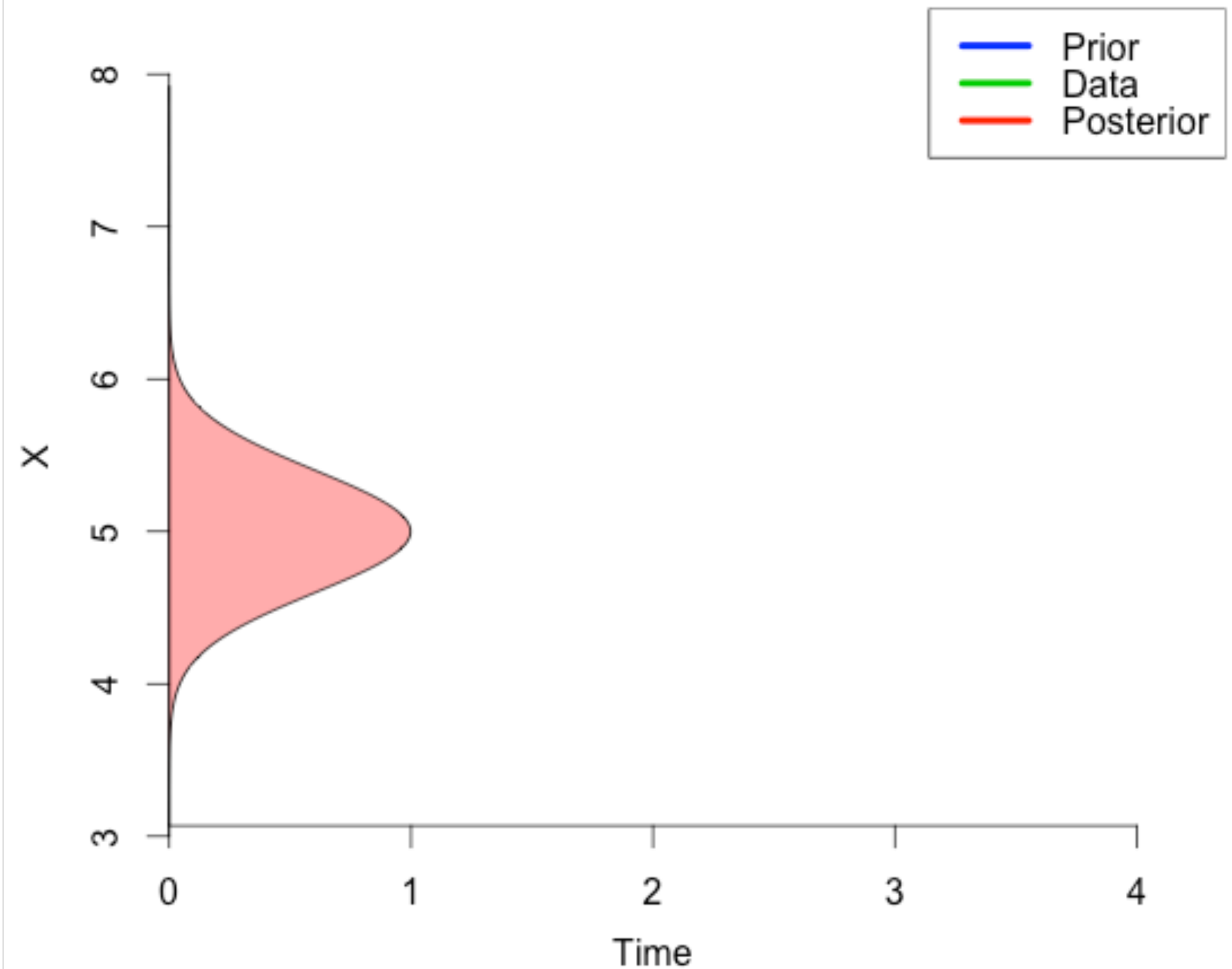


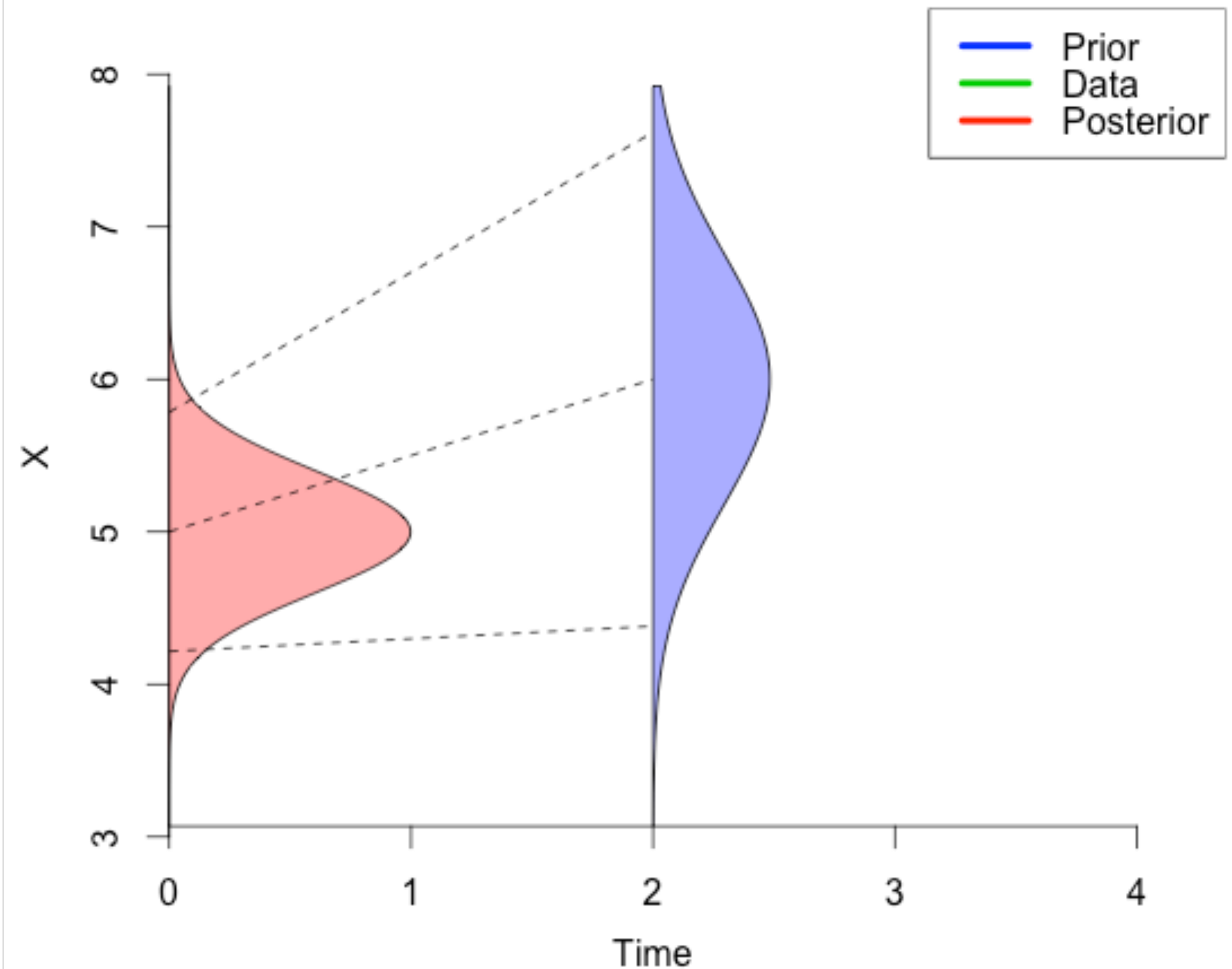


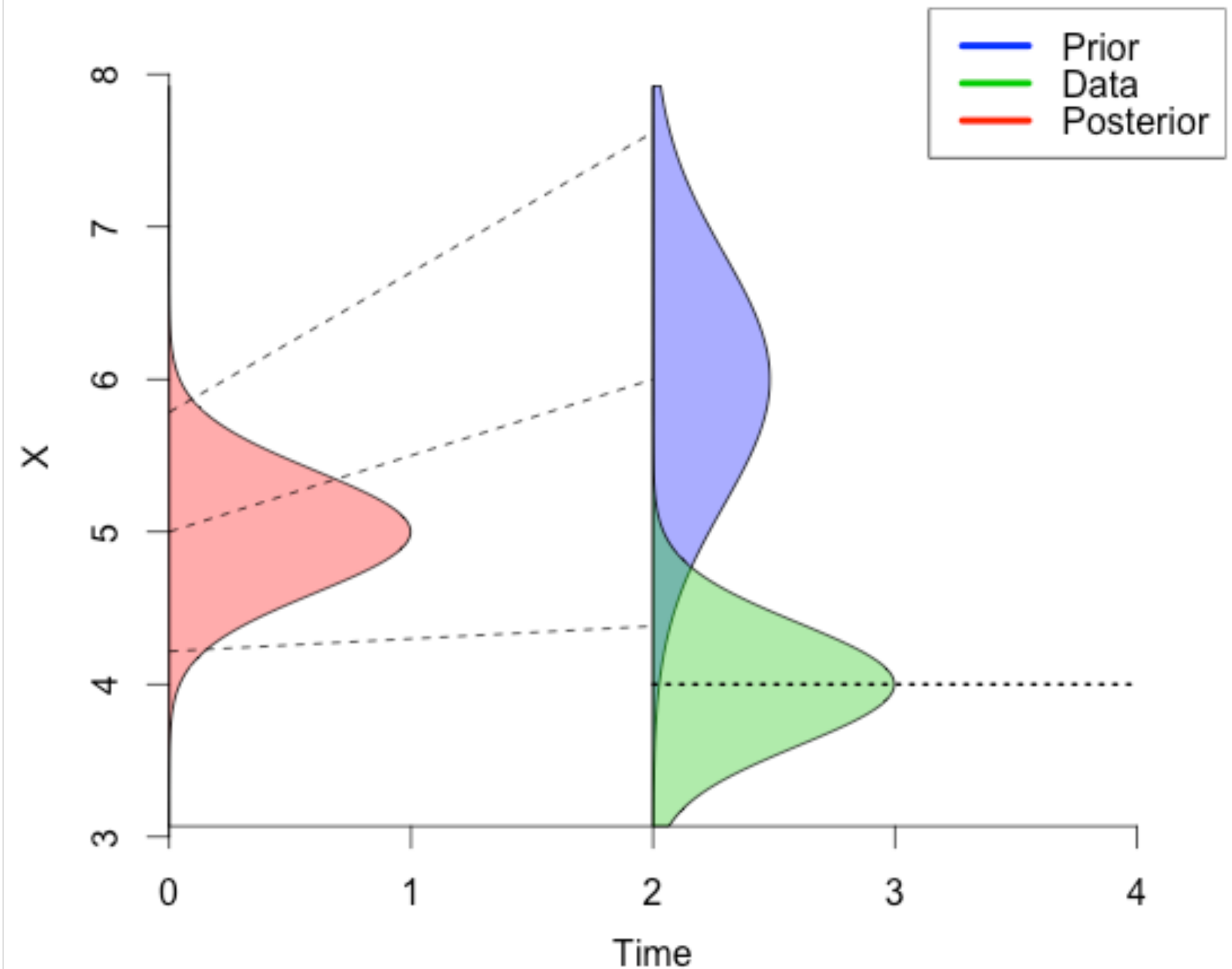


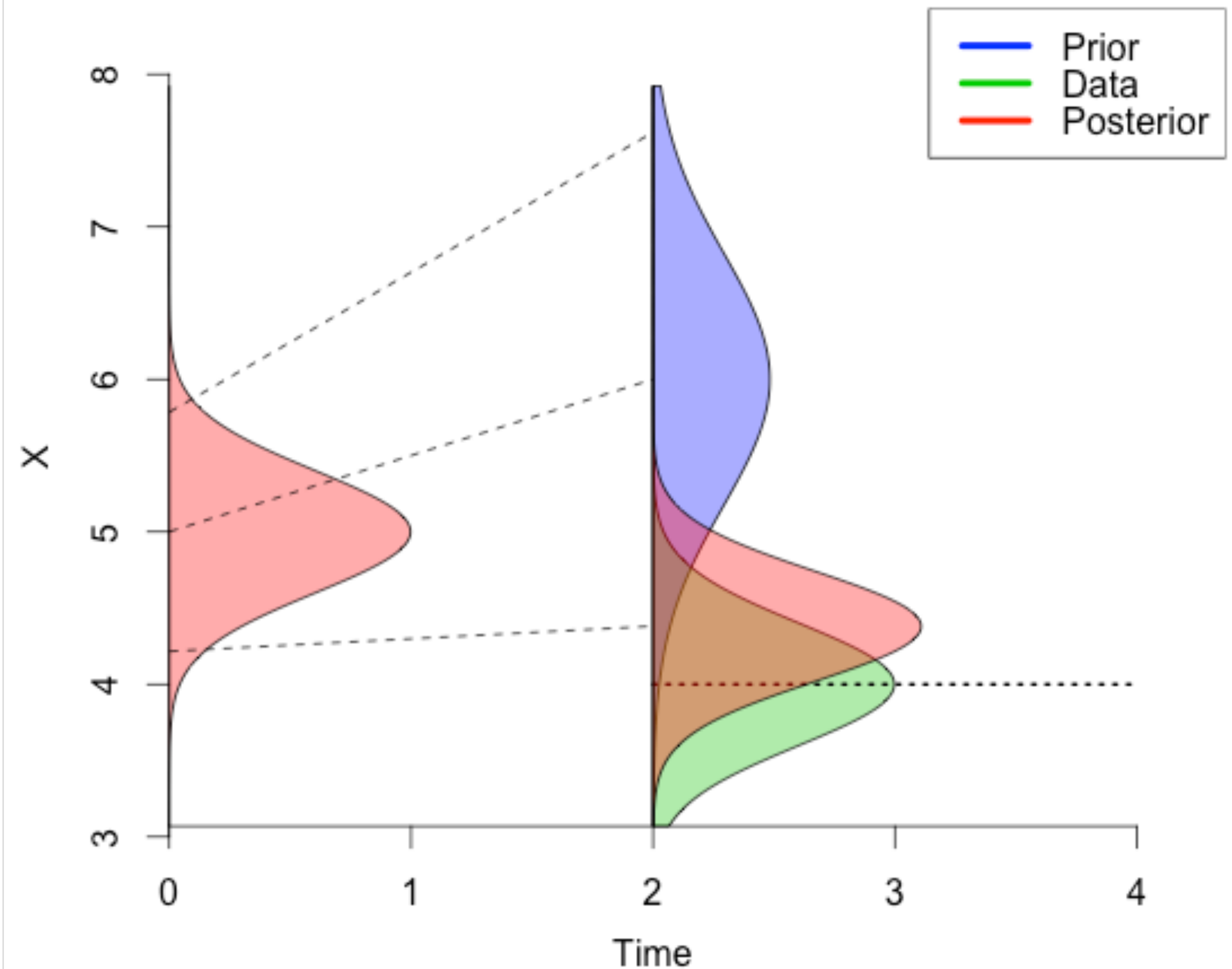

NCEP Operational Forecast Skill
36 and 72 Hour Forecasts @ 500 MB over North America
[100 * (1-S1/70) Method]
36 Hour Forecast
72 Hour Forecast
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
15 Years
IBM 701
IBM 704
IBM 7090
IBM 7094
CDC 6600
IBM 360/195
CYBER 205
CRAY Y-MP
CRAY C90
IBM SP
IBM P690
IBM P655+
IBM Power 6
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
NCEP Central Operations January 2015


DATA
Fx
Obs
Analysis
IC
Drivers
Param
MODEL
Forecast
Fx

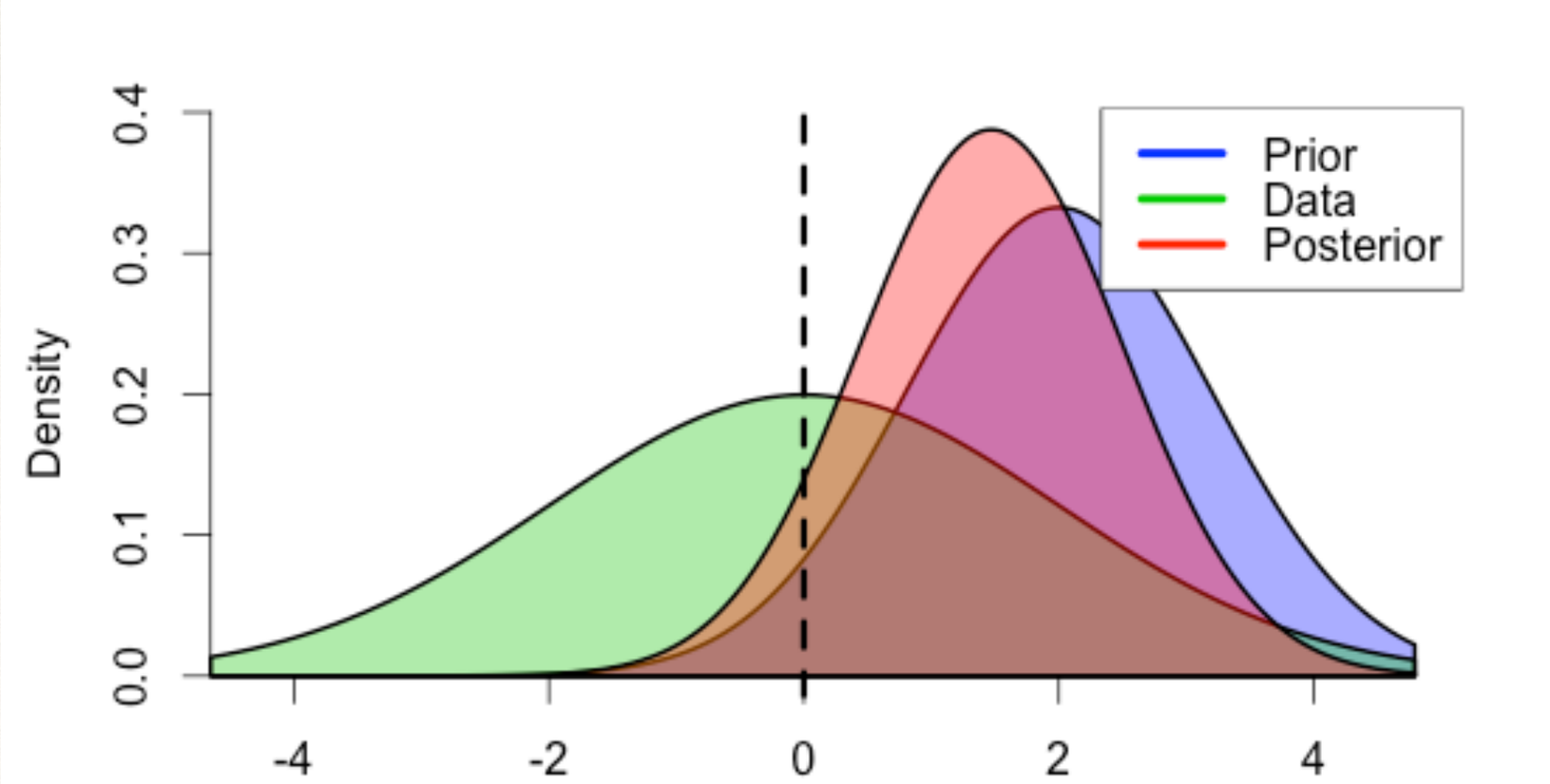

Prior
Data
Posterior
X
3
4
5
6
7
8
0
1
2
3
4
Time


Prior
Data
Posterior
X
3
4
5
6
7
8
0
1
2
3
4
Time

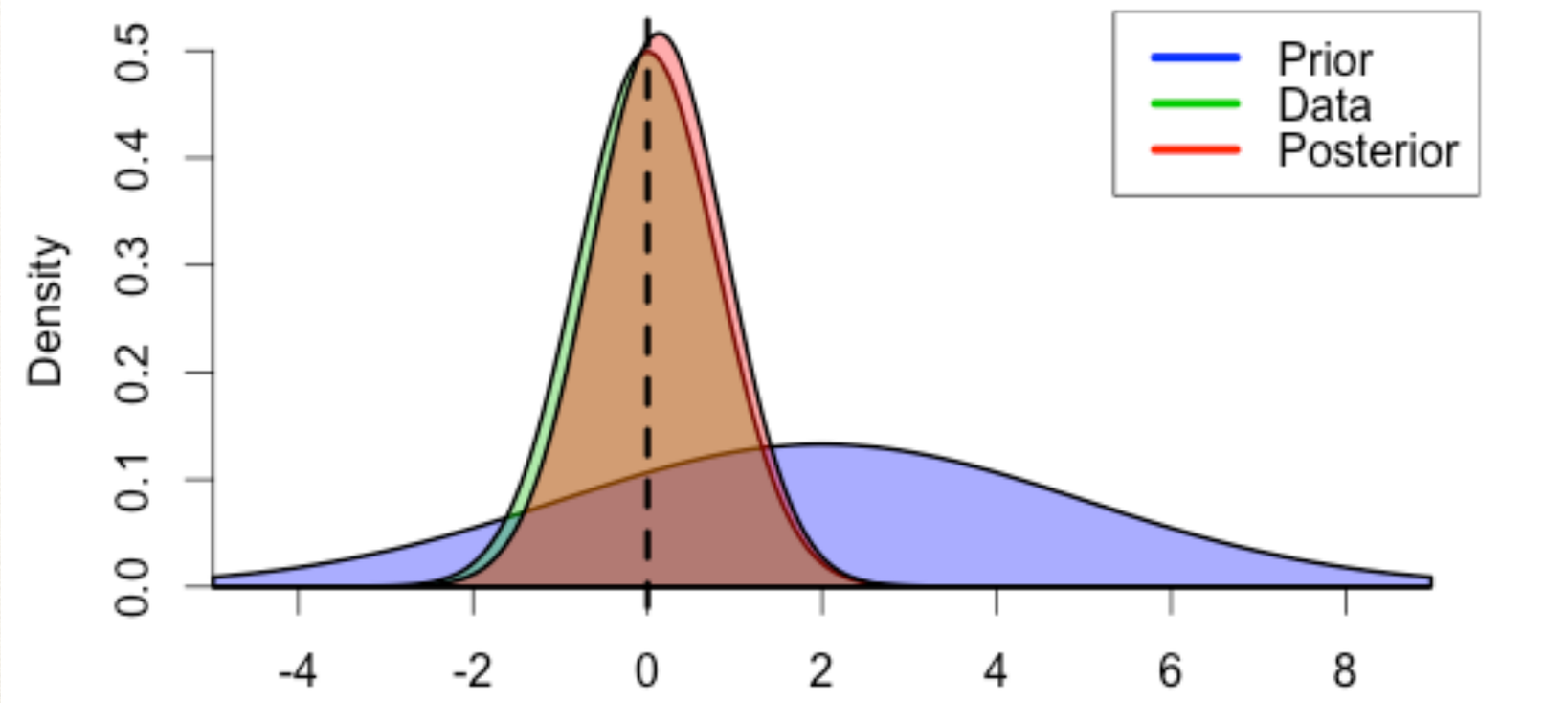

Prior
Data
Posterior
X
3
4
5
6
7
8
0
1
2
3
4
Time


Prior
Data
Posterior
X
3
4
5
6
7
8
0
1
2
3
4
Time

Less Precise Data

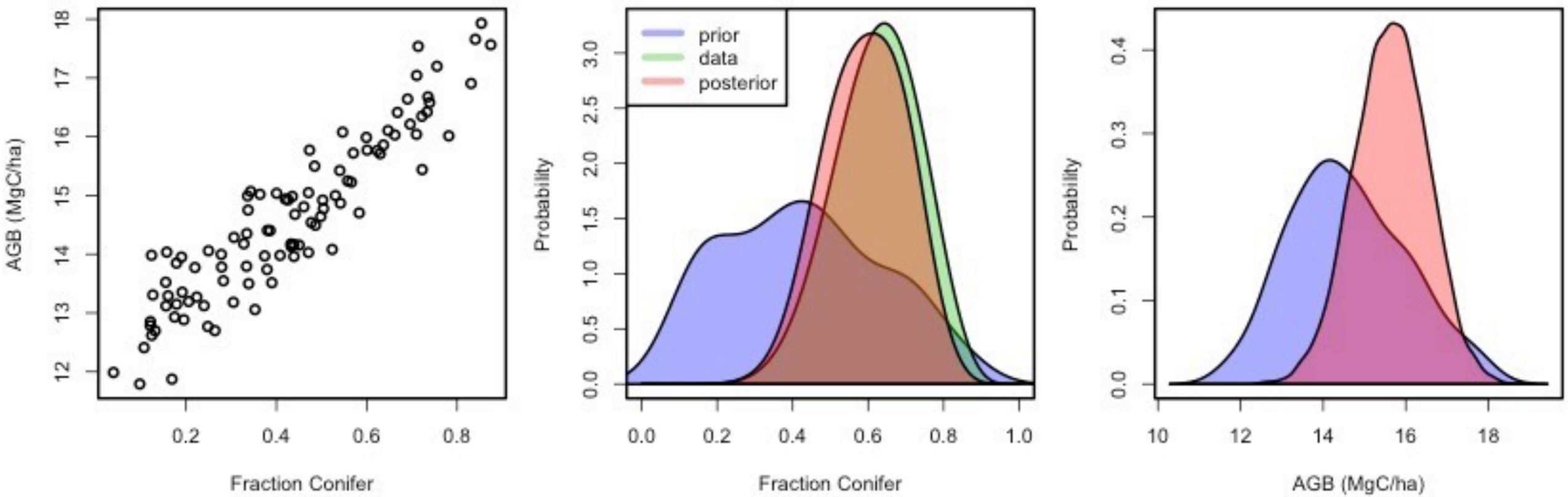


Less Precise Model



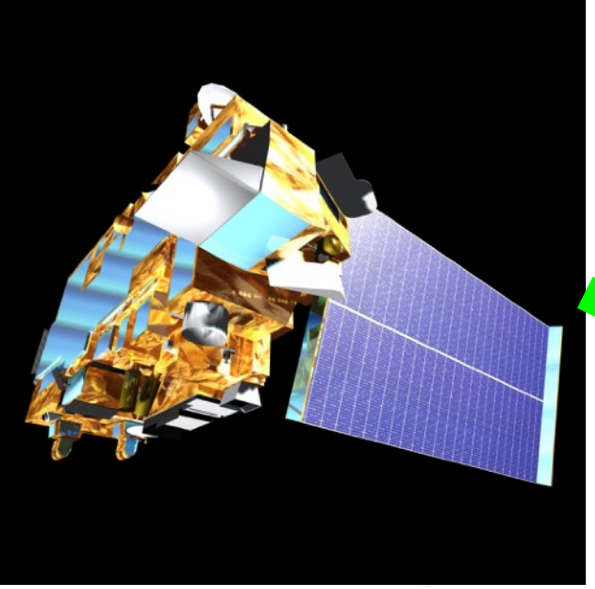
Precision controls influence

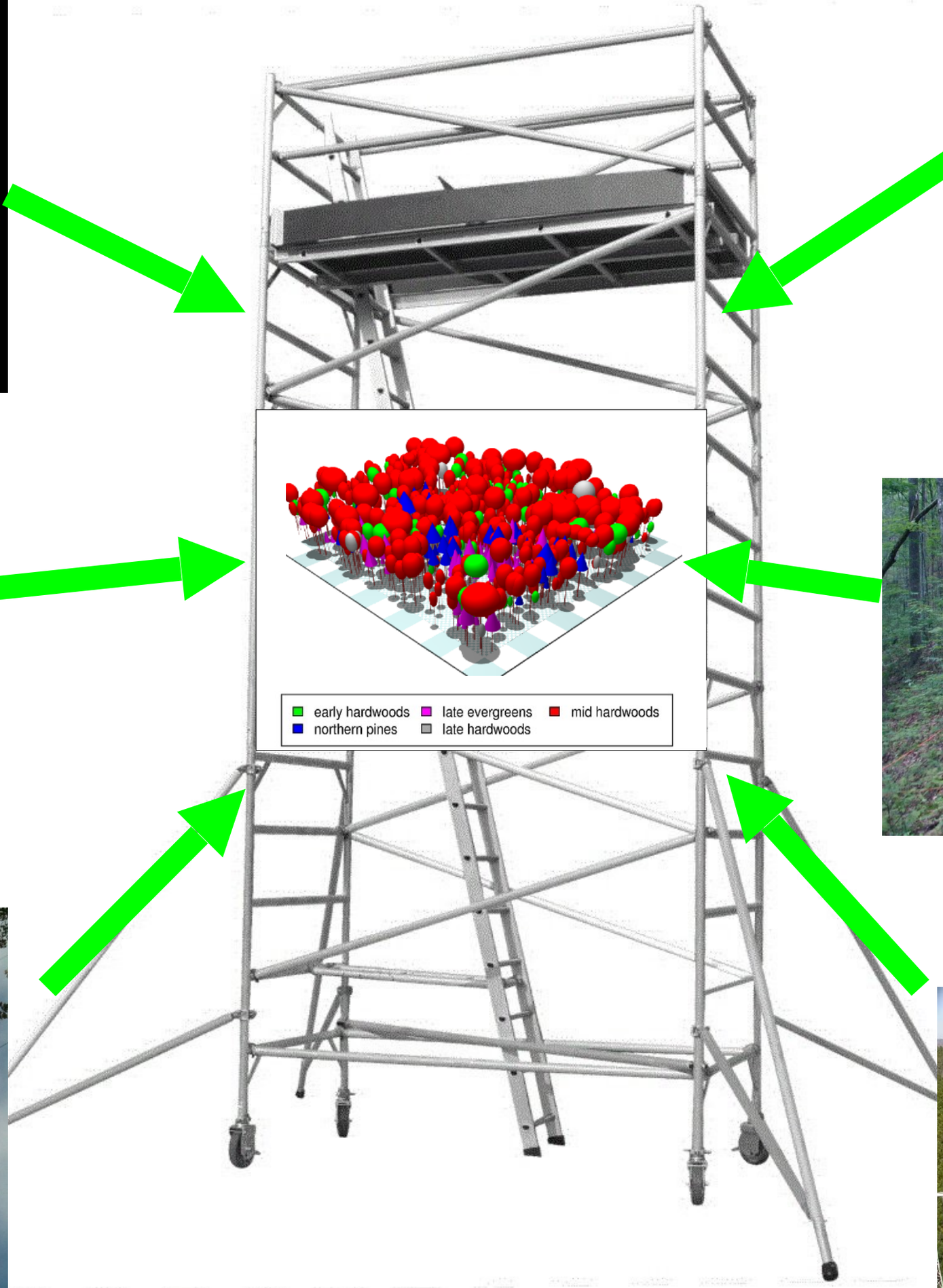
State-Variable Data Assimilation



$$P(\theta|y) \propto P(y|\theta)\,P(\theta)$$

Updated State **Data** **Model**




early hardwoods
late evergreens
mid hardwoods
northern pines
late hardwoods


©2007 Jeffrey Pippen

ECOLOGICAL FORECASTING

- Is more than forward simulation
- Requires a fusion of models and data
- Must address multiple sources of uncertainty and variability
- Think Probabilistically!!