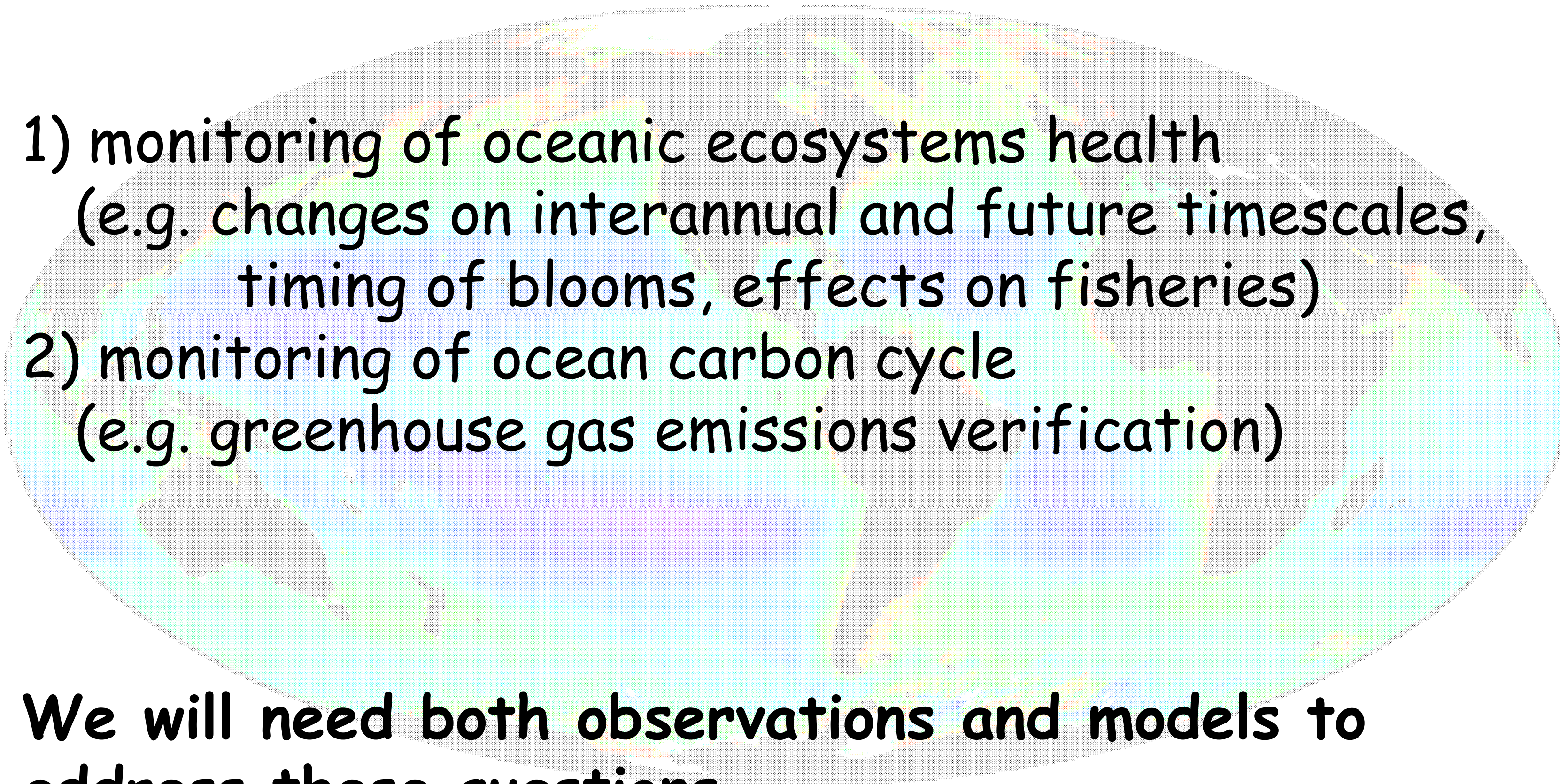
POTENTIAL BENEFITS OF SYNTHESIZING AUTONOMOUS NETWORK DATA WITH MODELS

Stephanie Dutkiewicz

Program of Atmospheres, Oceans and Climate
Massachusetts Institute of Technology

Societal relevant questions:



1) monitoring of oceanic ecosystems health
   (e.g. changes on interannual and future timescales,
   timing of blooms, effects on fisheries)
2) monitoring of ocean carbon cycle
   (e.g. greenhouse gas emissions verification)

We will need both observations and models to address these questions

MODELS $\leftrightarrow$ DATA

- end users: data used for initialization, verification, context

- synthesis: model and data used together

- feedback: model helps inform on observing system design

Observations: limited in time and space, and have errors

Models: deficient in representation of processes

Model-data synthesis can combine best of both

Massachusetts Institute of Technology

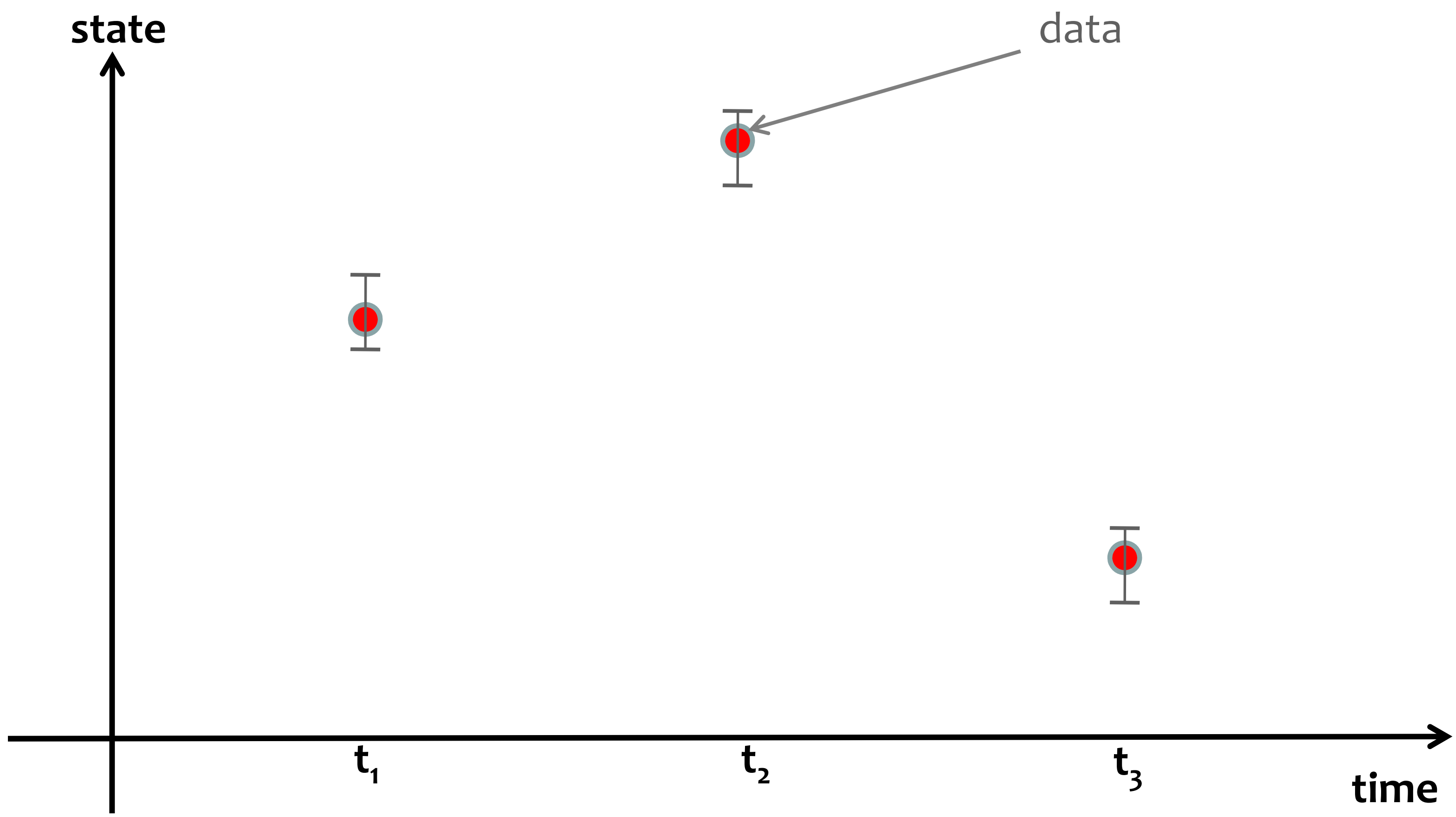
Movie credit: Oliver Jahn, MIT

MODEL-DATA SYNTHESIS

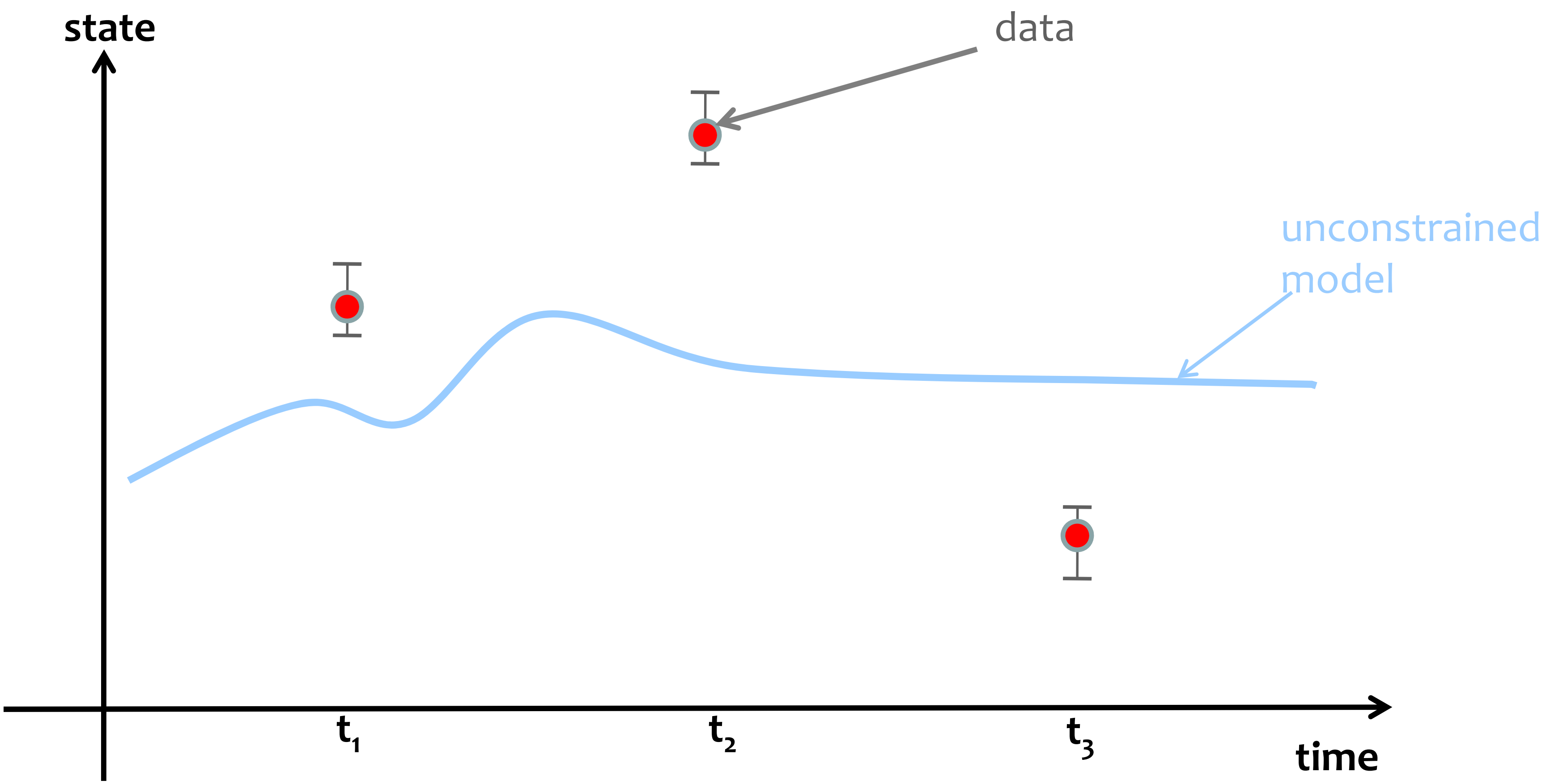
Assimilation methods:

Some lessons learned from biogeochemical data assimilation:

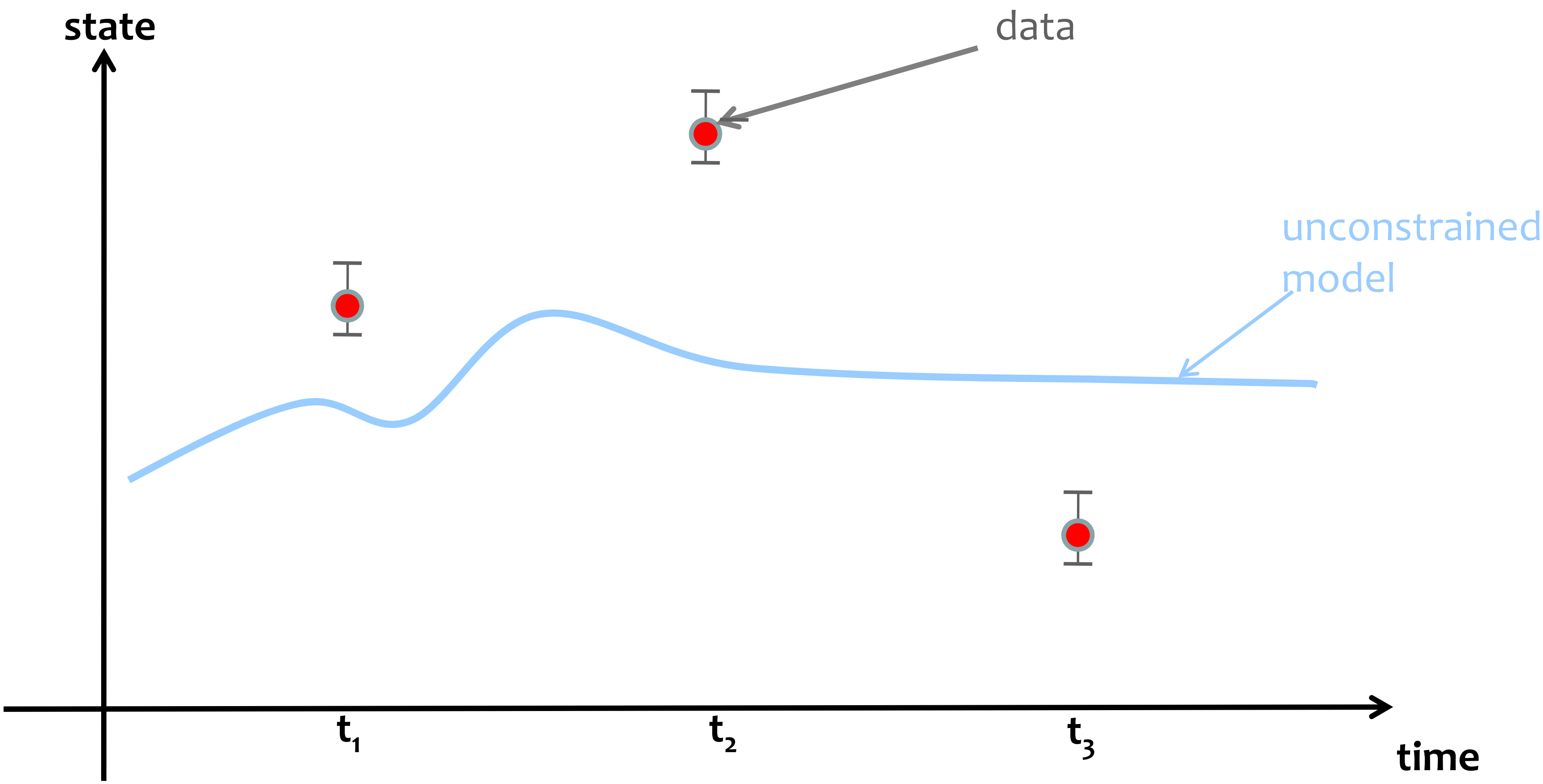
Some lessons learned from physical data assimilation:



Massachusetts Institute of Technology



Massachusetts Institute of Technology



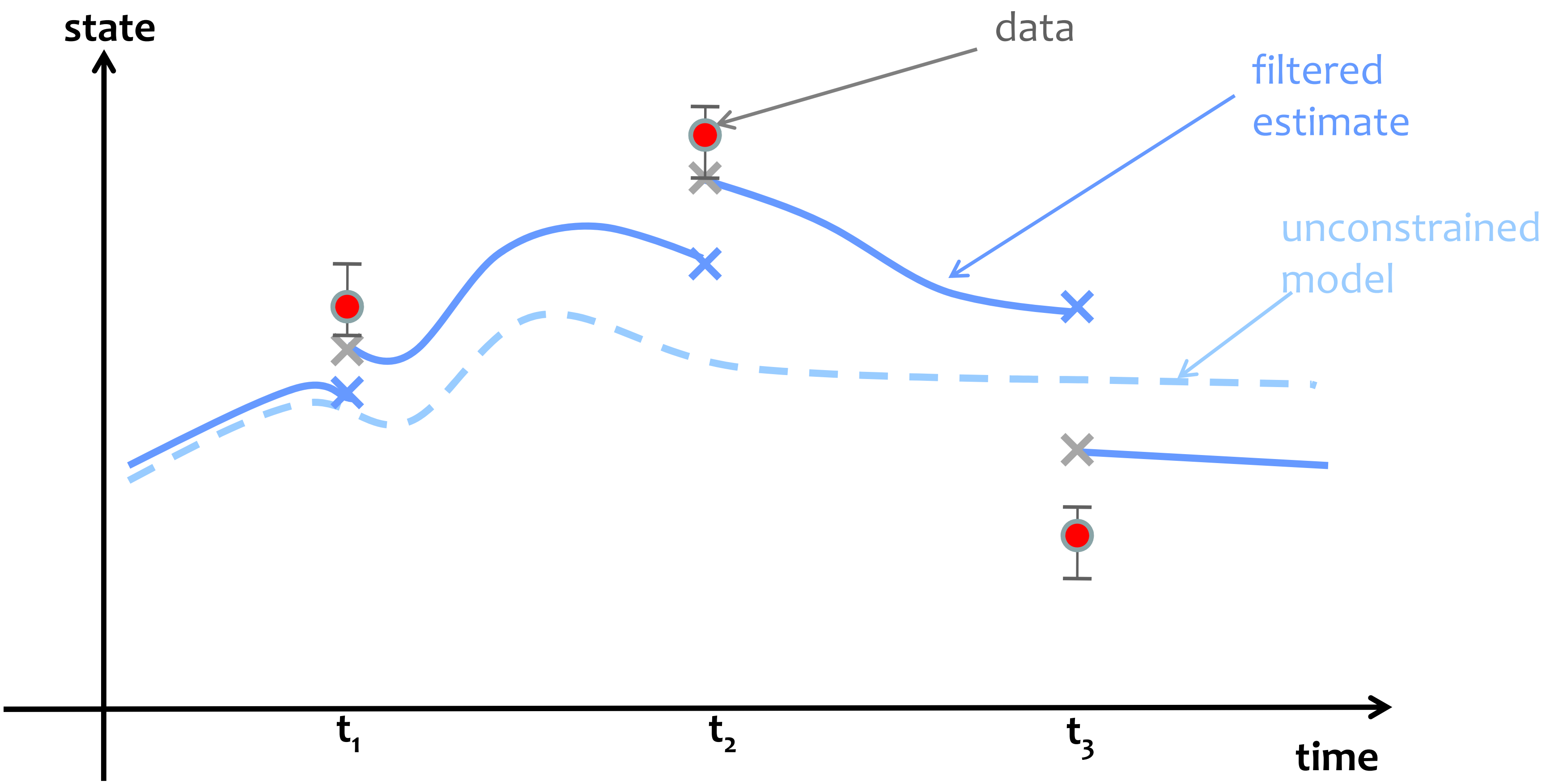
least square fit of model to observations:

$$J=\Sigma_i W_i^2(\text{model}_i-\text{data}_i)^2$$

J - cost function

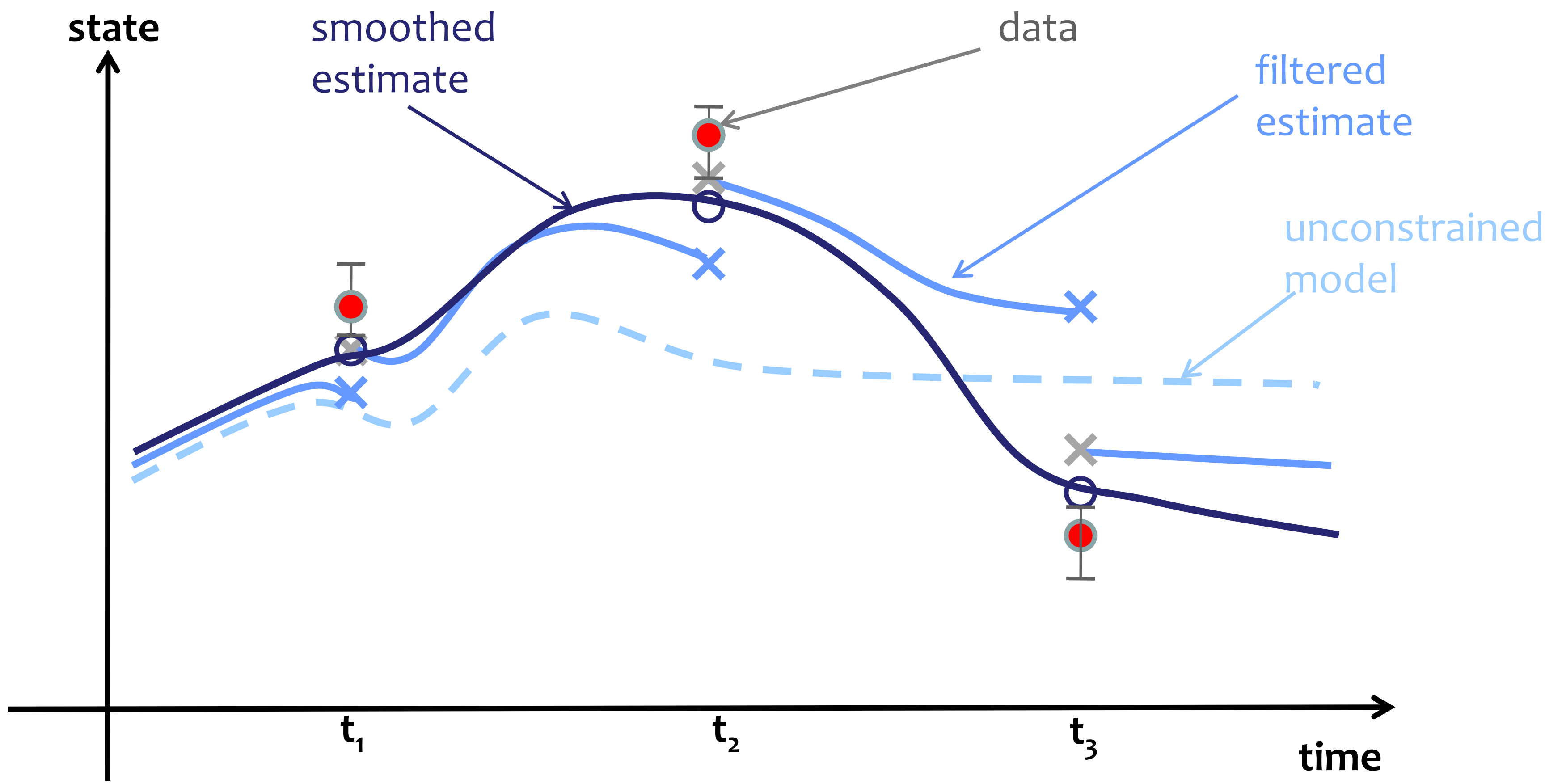
W_i - weighting function (error estimates)

Assimilation: Sequential Method



(e.g. Kalman Filter)

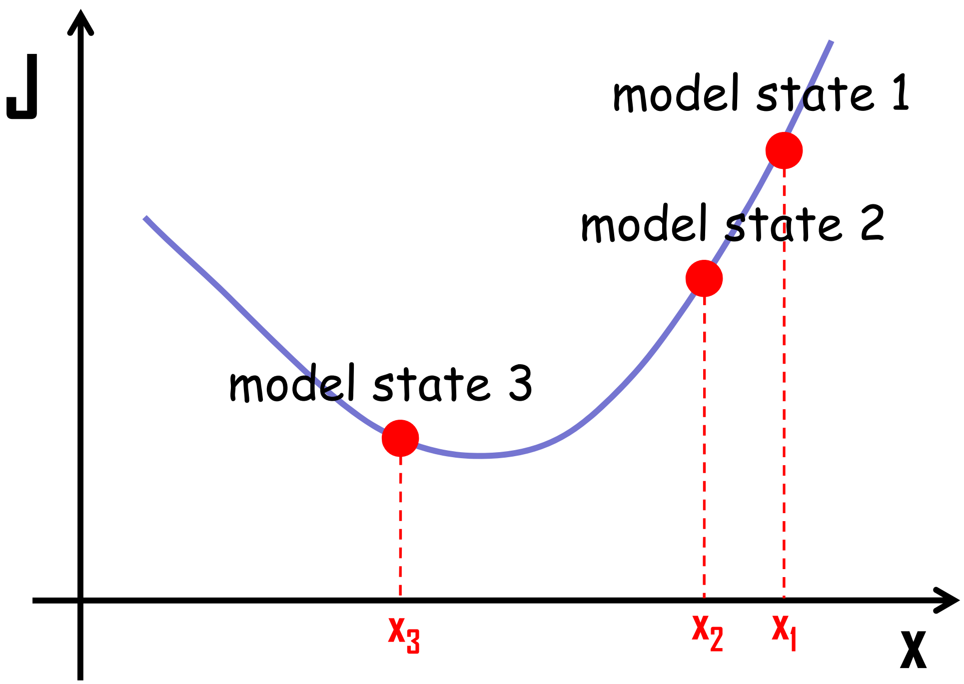
Assimilation: Smoother Method



(e.g. Adjoint)

State Estimation: Adjoint Method

- essentially the "backward" model
- efficiently computes sensitivity of model (cost function) to perturbations in parameters/initial conditions/forcing fields



$$J=\Sigma_i W_i^2(\text{model}_i-\text{data}_i)^2$$

Use adjoint to find gradient **dJ/dx** to iteratively minimize **J**

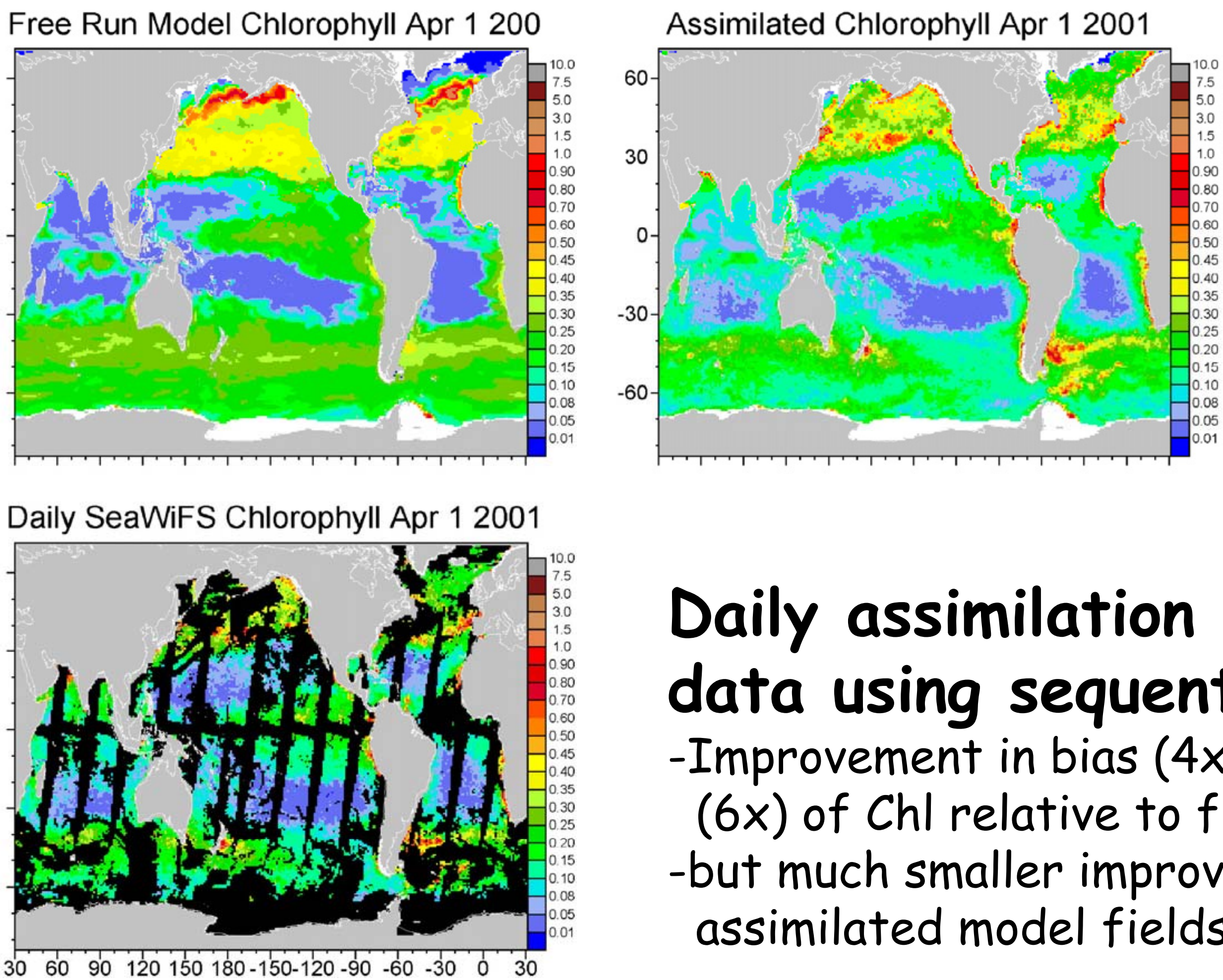
x can be:
- initial conditions
- forcing fields
- model parameters

MODEL-DATA SYNTHESIS

Assimilation methods:

Some lessons learned from biogeochemical data assimilation:

Some lessons learned from physical data assimilation:

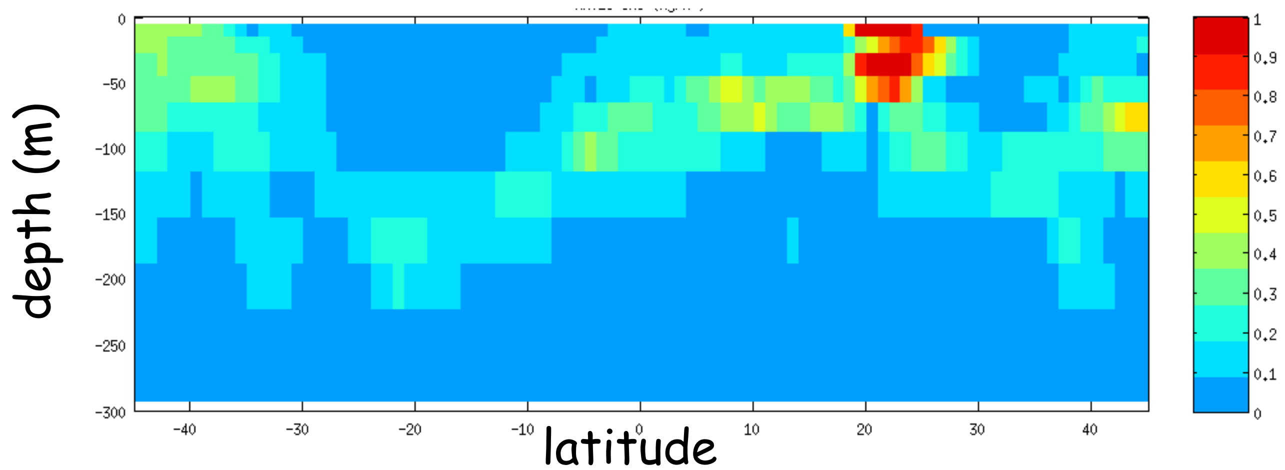


Daily assimilation of SeaWiFS data using sequential method:

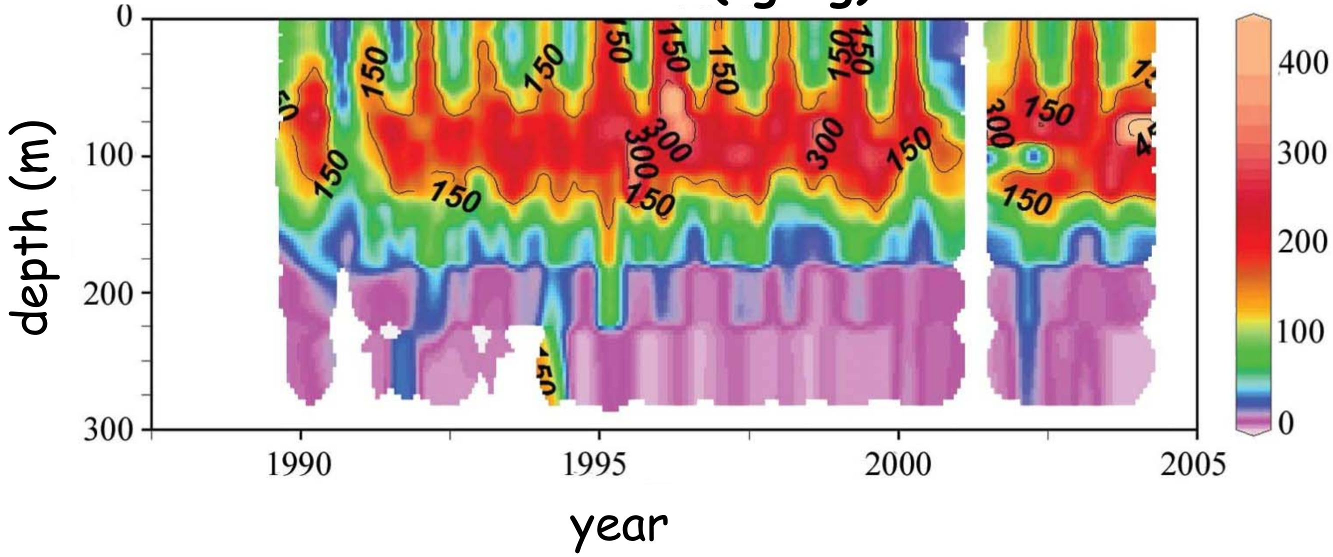
- -Improvement in bias (4x) and uncertainty (6x) of Chl relative to free run,
- -but much smaller improvement in non assimilated model fields

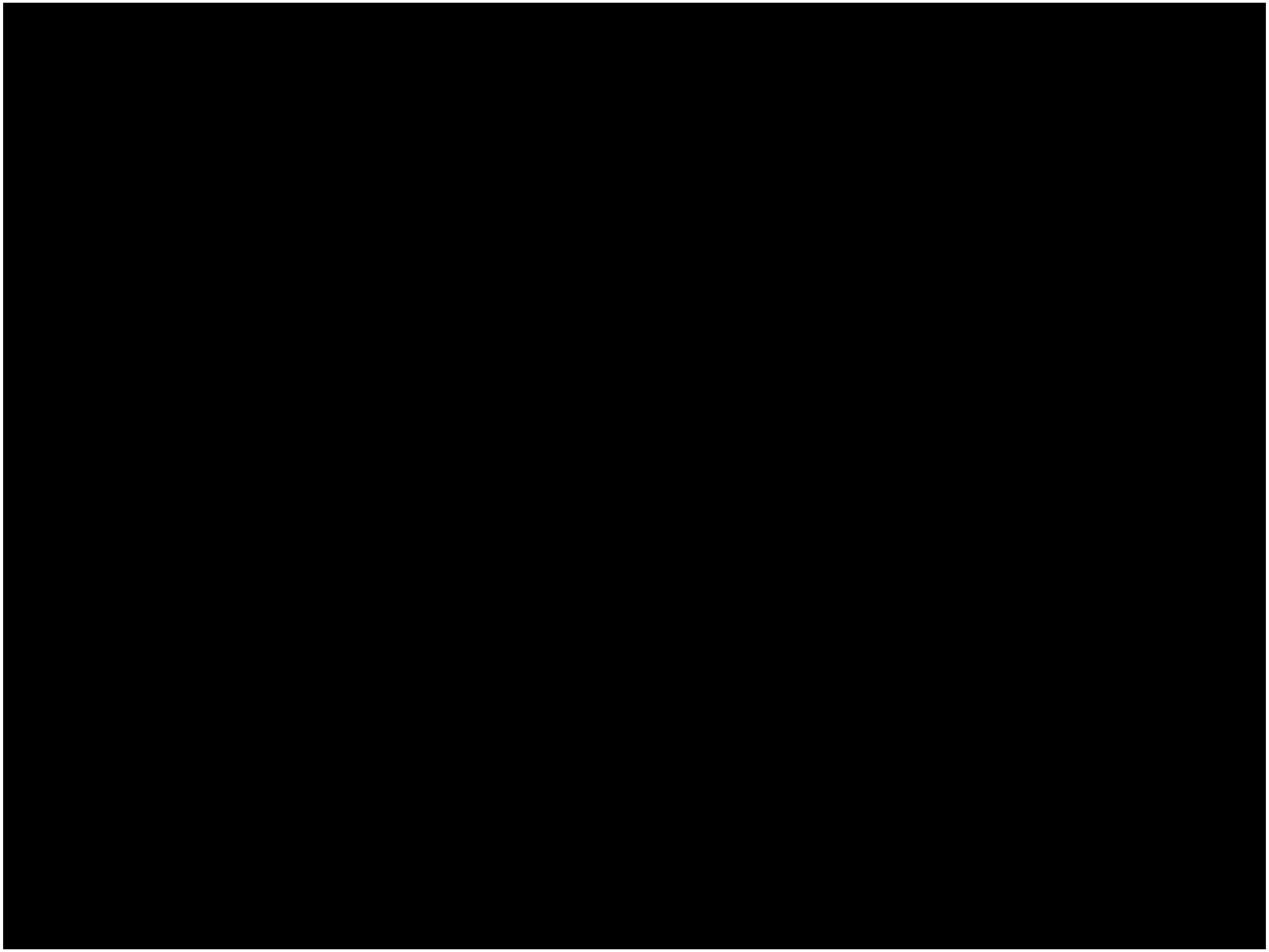
Gregg, JMS, 2008

AMT15 Chl (mg/m^3)



BATS Chl (ng/kg)





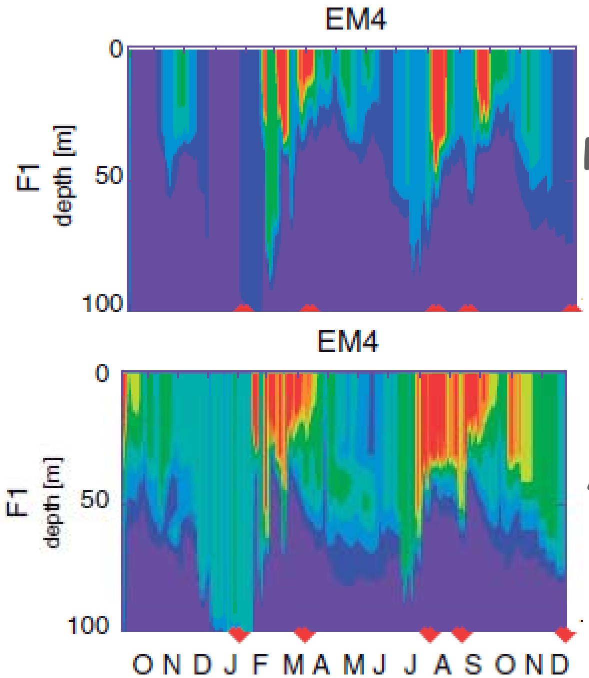
Massachusetts Institute of Technology

Movie credit: Oliver Jahn, MIT

Adjoint assimilation of Chl, DIN, PP, export, zooplankton biomass at Arabian Sea Site:

• Only subset of uncorrelated parameters could be optimized.

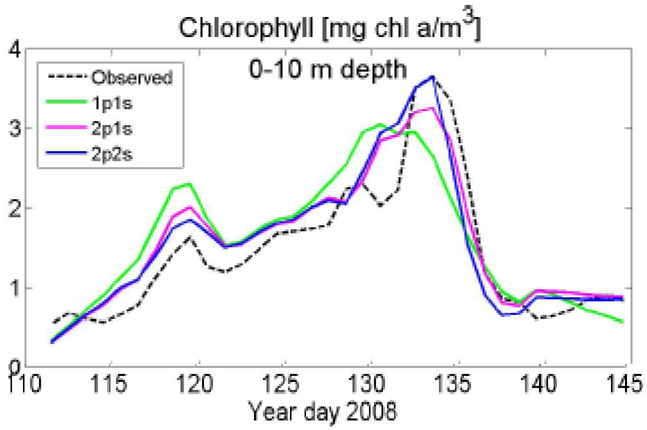
• Change in physics had greater affect than changes in model complexity: good physics essential before optimizing

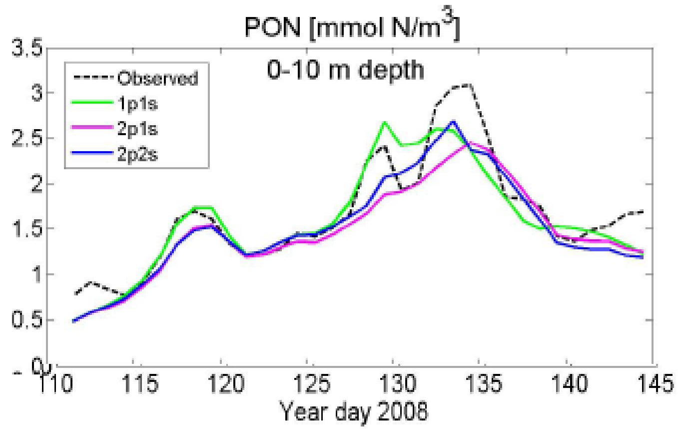


Pre-assimilated Chl

Assimilated Chl

Friedrichs et al, DSRII, 2006





Assimilation of Chl, PON, DIN, O_2 from Lagrangian float using variational technique

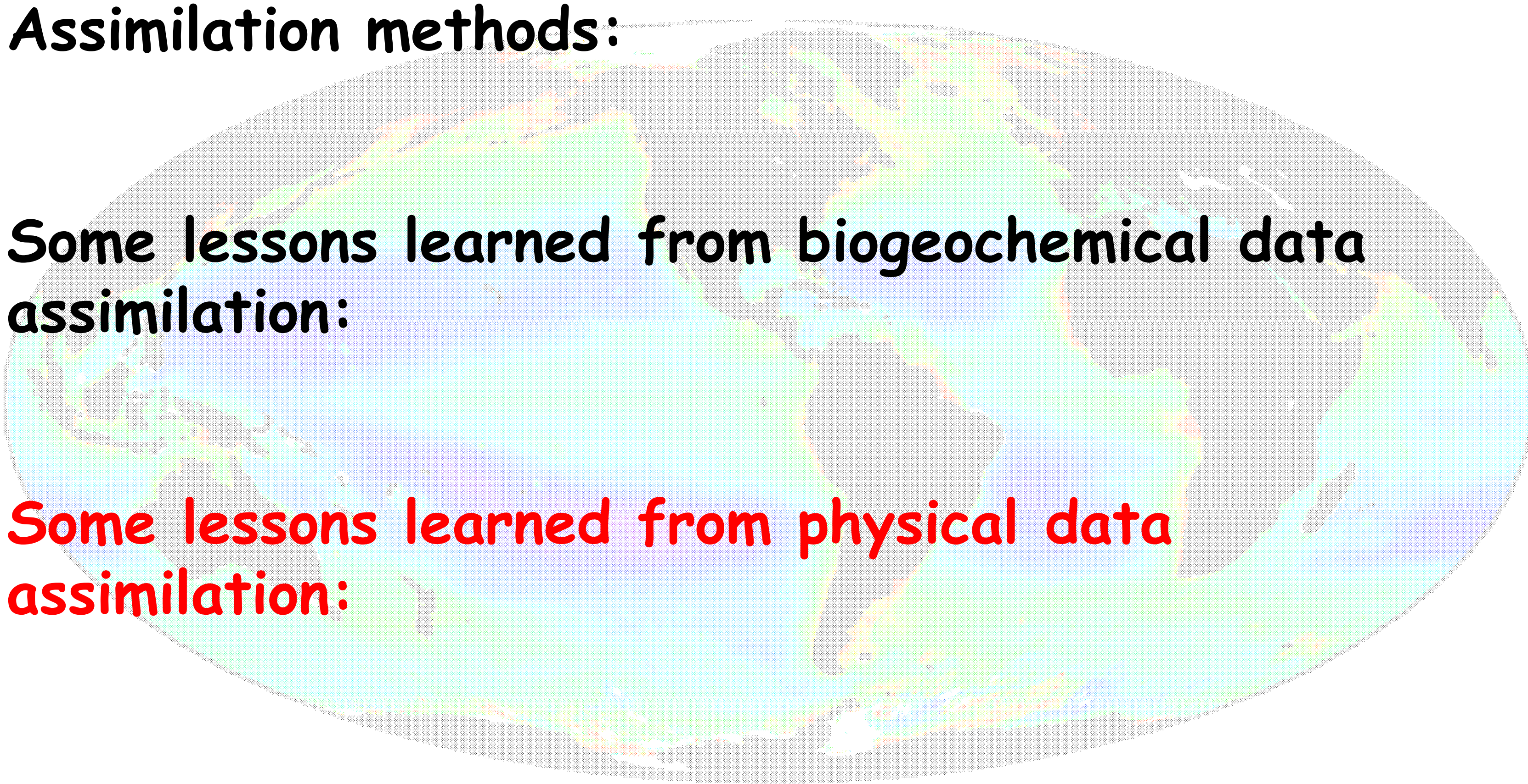
- phytoplankton-related parameters constrained better than previous optimization models

- Autonomous daily profiles provide better constraint to models, but even this data can only constrain a subset of parameters (12 of 25)

- Data not able to constrain deep carbon export

Bagniewski et al, BG, 2011

MODEL-DATA SYNTHESIS

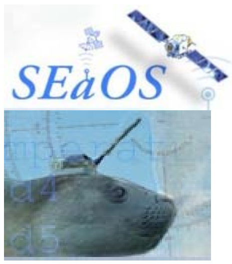


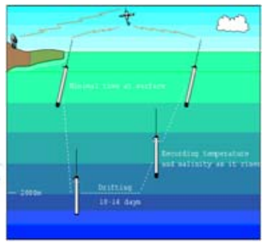
Assimilation methods:

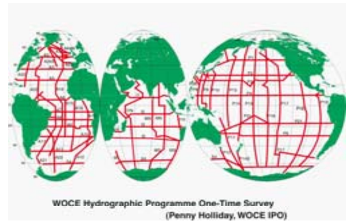
Some lessons learned from biogeochemical data assimilation:

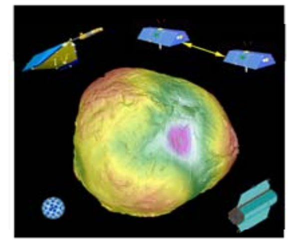
Some lessons learned from physical data assimilation:

ECCO-GODAE: global (multi-)decadal adjoint produced dynamically consistent state estimates for climate research



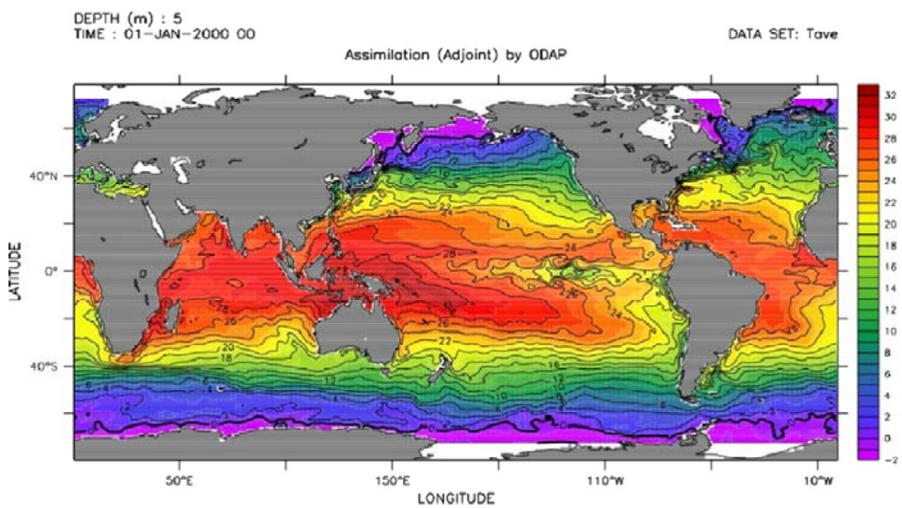






MITgcm
1° x 1° 23 levels
adjoint.

1992-present



Temperature (Deg C)

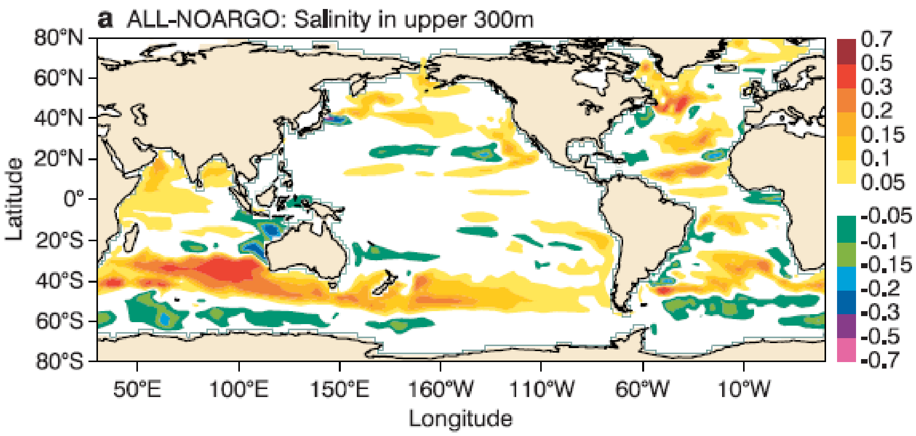
Wunsch +Heimbach, Physica D, 2007;
Wunsch et al., Oceanography, 2009

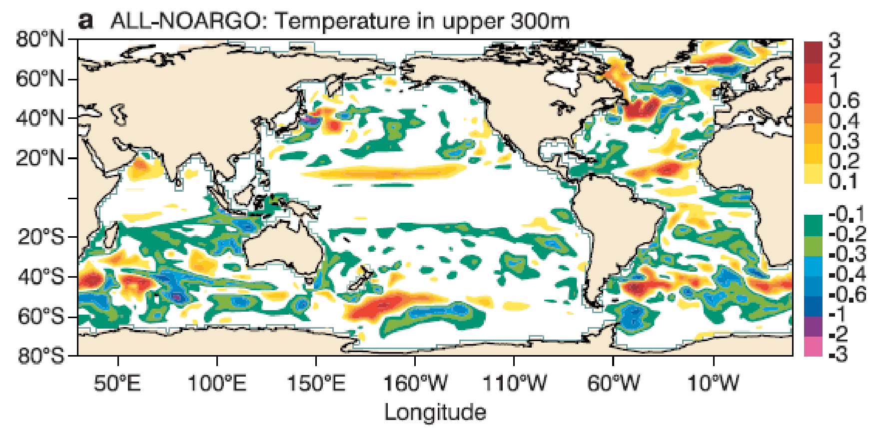
Massachusetts Institute of Technology

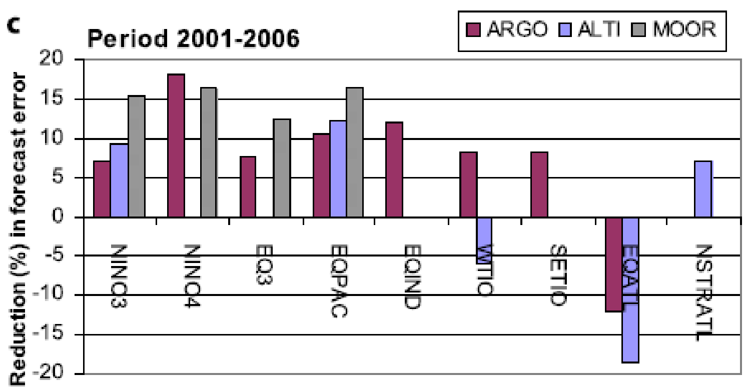
Slide courtesy of Patrick Heimbach

Data withholding experiments from ECMWF operational forecast analysis (sequential assimilation):

- ARGO essential correcting basin scale salinity
- ARGO less important for Eq. Pacific, due to TAO/TRITON array





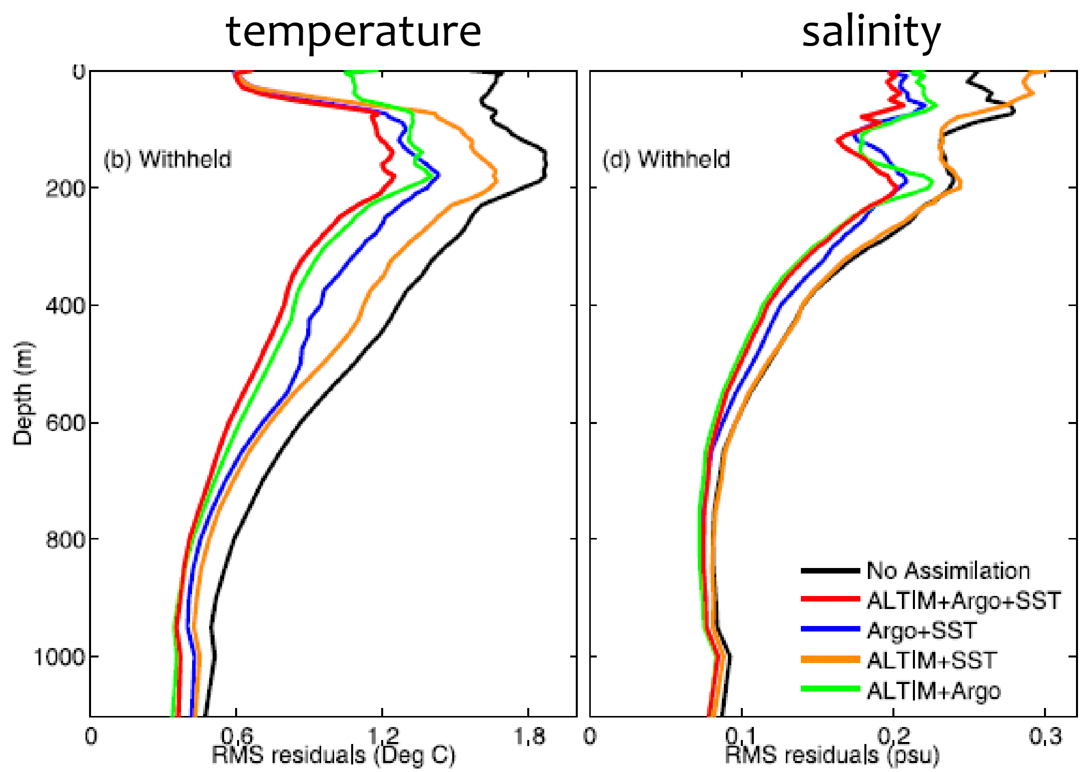


- impact of initialization in forecast skill: ARGO essential in many regions

Balmaseda et al, GRL, 2007; 2009

Data withholding experiments from BlueLink Australian operational Forecast model (sequential assimilation):

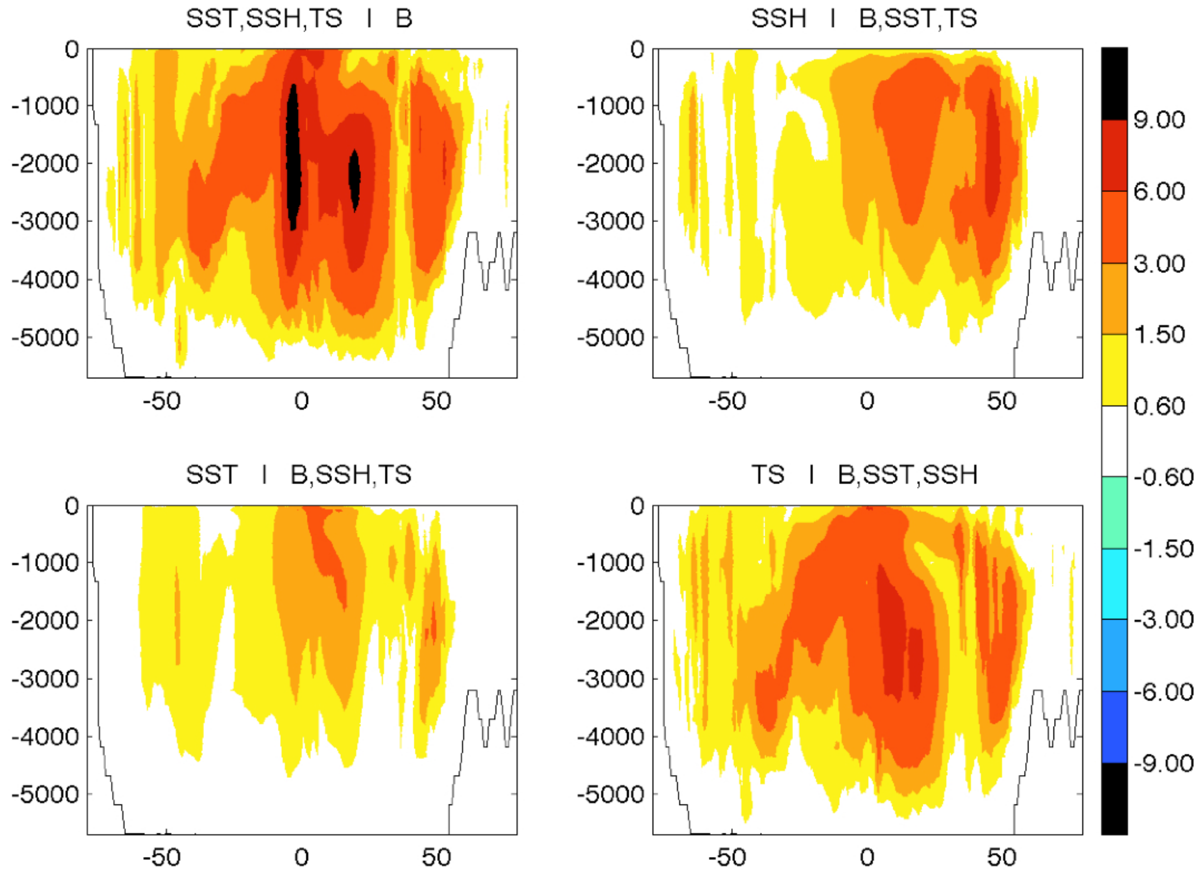
•Different data sets constrain different variables
•ARGO essential for sub-surface T, full profile S



Oke and Schiller, GRL, 2007

OCCA (ECCO) adjoint assimilation product:

Difference in variability of MOC relative to a base experiment with only hydrography as "data input"



B: hydrography only
SSH: altimetry
SST: satellite SST
TS: Argo

Gael Forget, in prep

MODELS ↔ DATA

- end users: data used for initialization, verification, context

- synthesis: model and data used together

- feedback: model helps inform on observing system design

Feedback from model to observing systems:

Interrogate model in manner consistent with observational network

Some examples:

Henson et al., BG, 2010

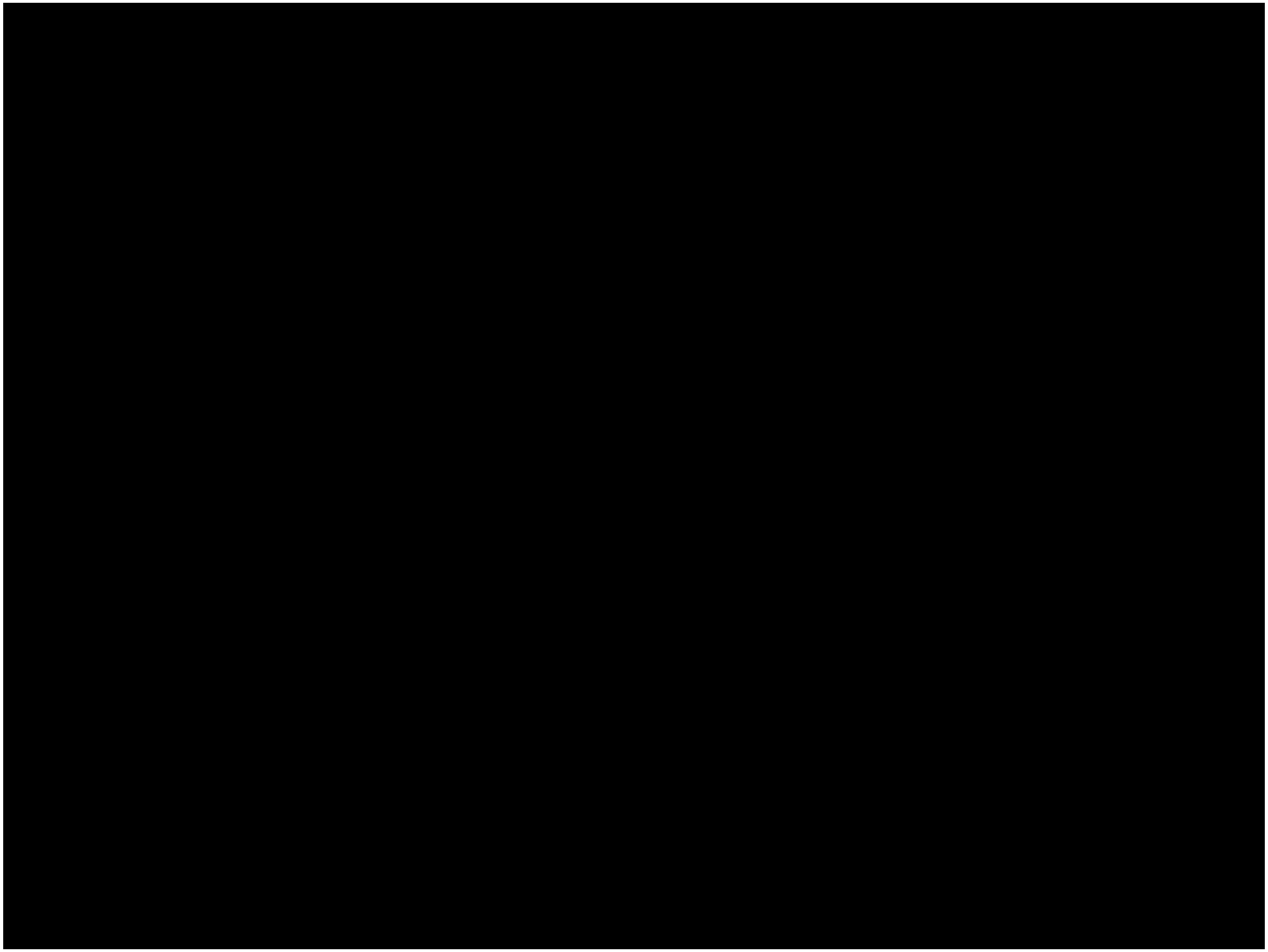
~40 years satellite Chl data needed to capture trend from natural variability

McKinley et al. (poster here, *Nature Geoscience* article in press)

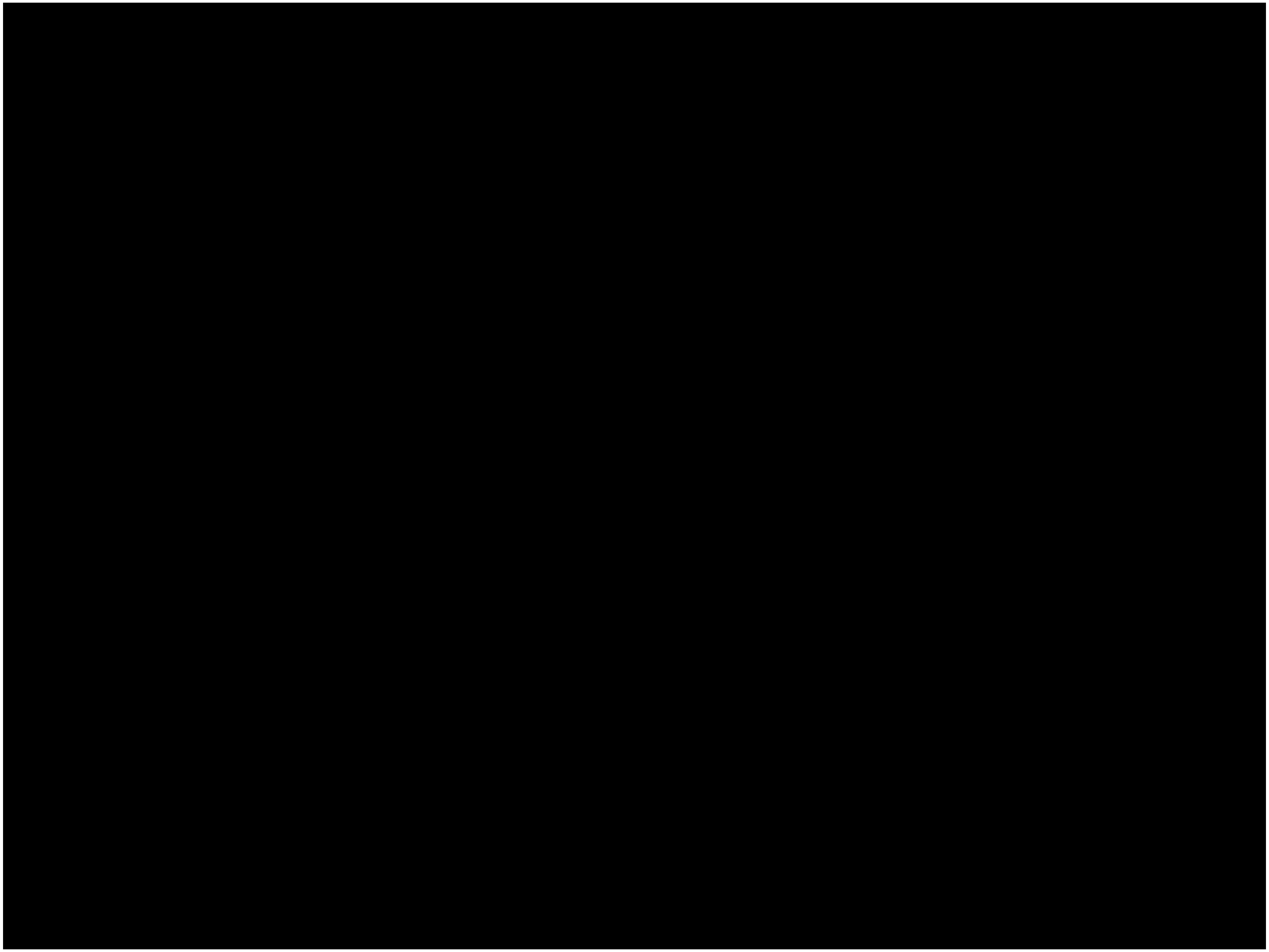
25 years for pCO2 measurements to show trend in ocean carbon sink?

Bennington et al, GBC, 2009

What does Chl variability tell us (or not) about pCO2, export?



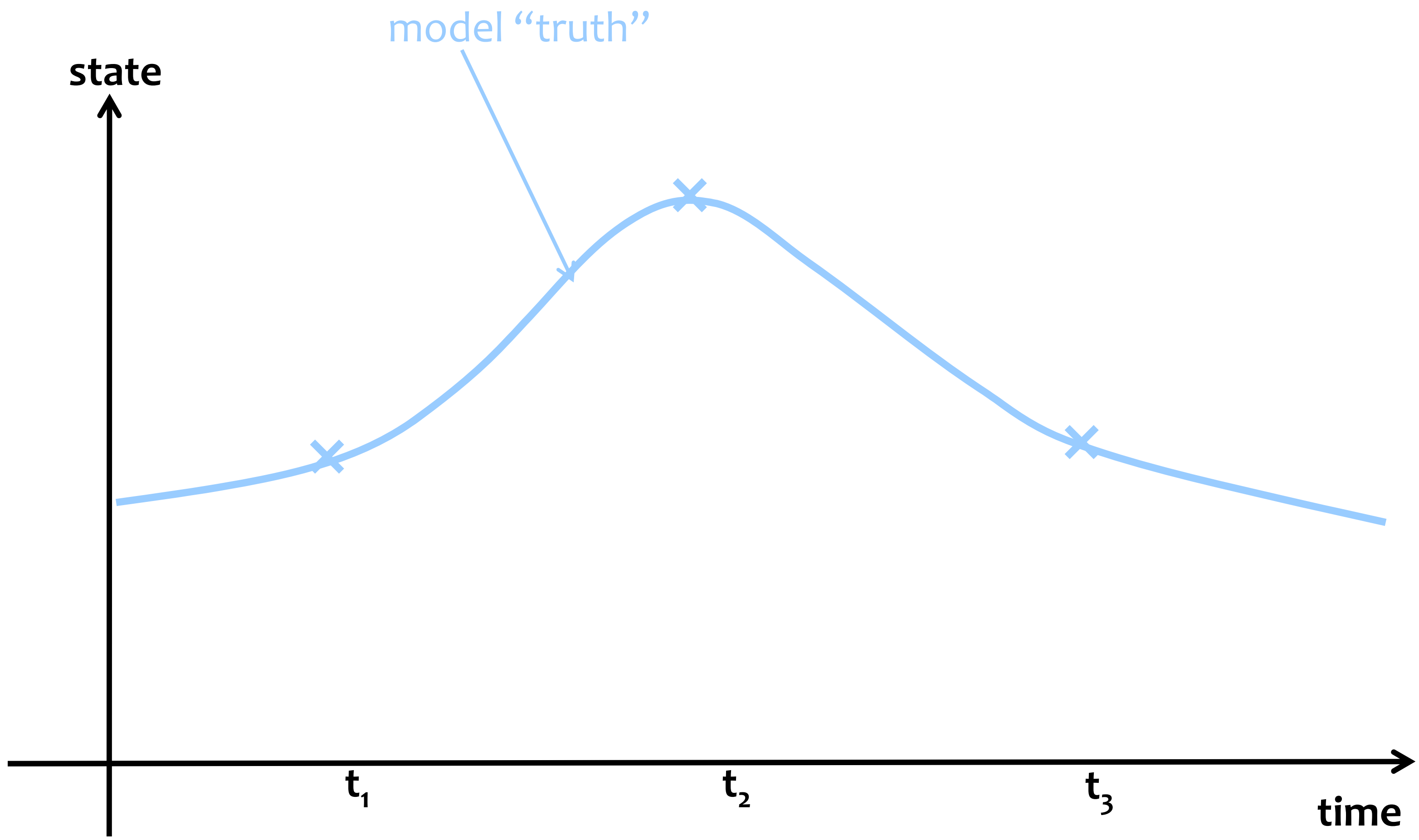
Movie credit: Oliver Jahn, MIT



Movie credit: Oliver Jahn, MIT

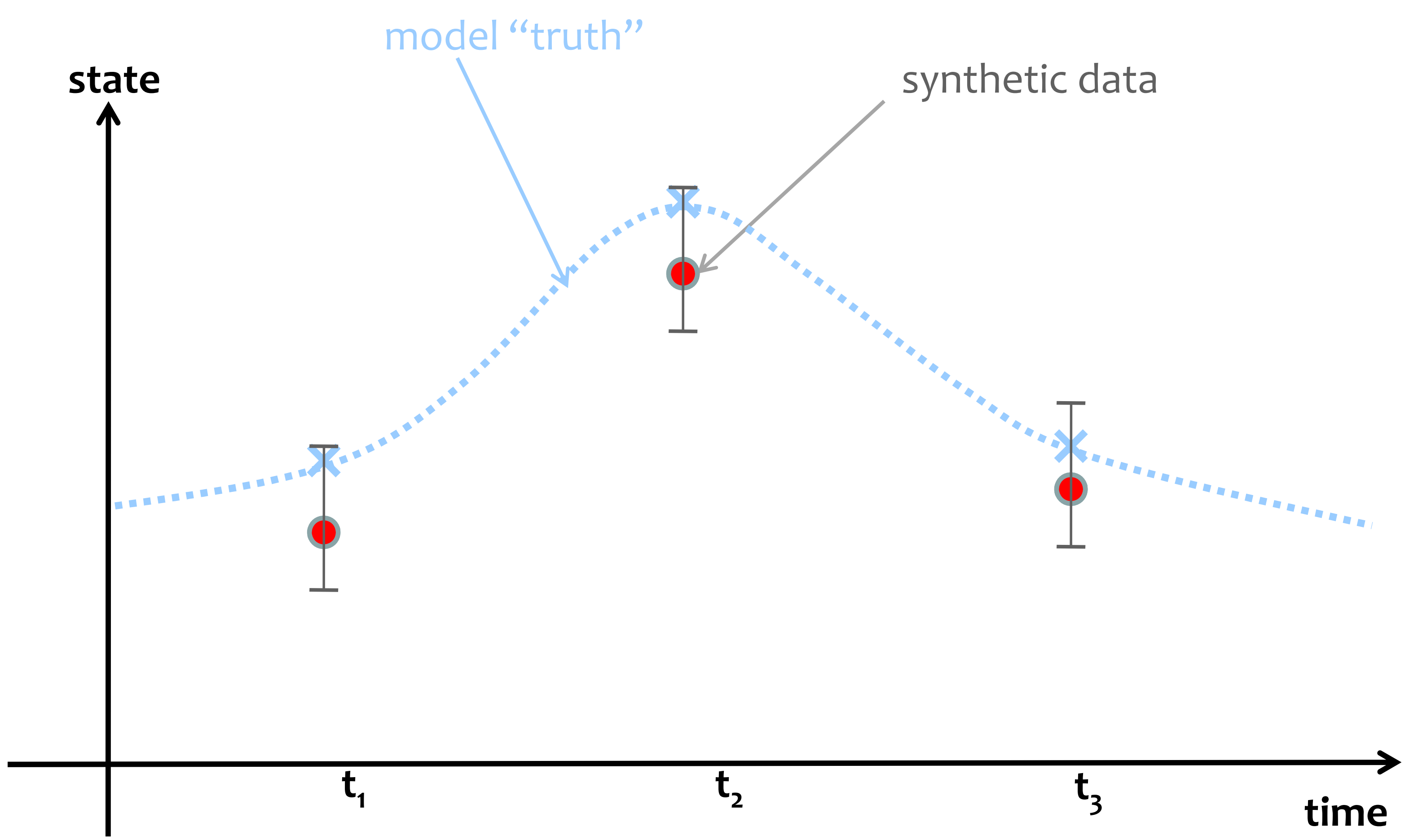
Massachusetts Institute of Technology

Observing System Simulation Experiment (OSSE)

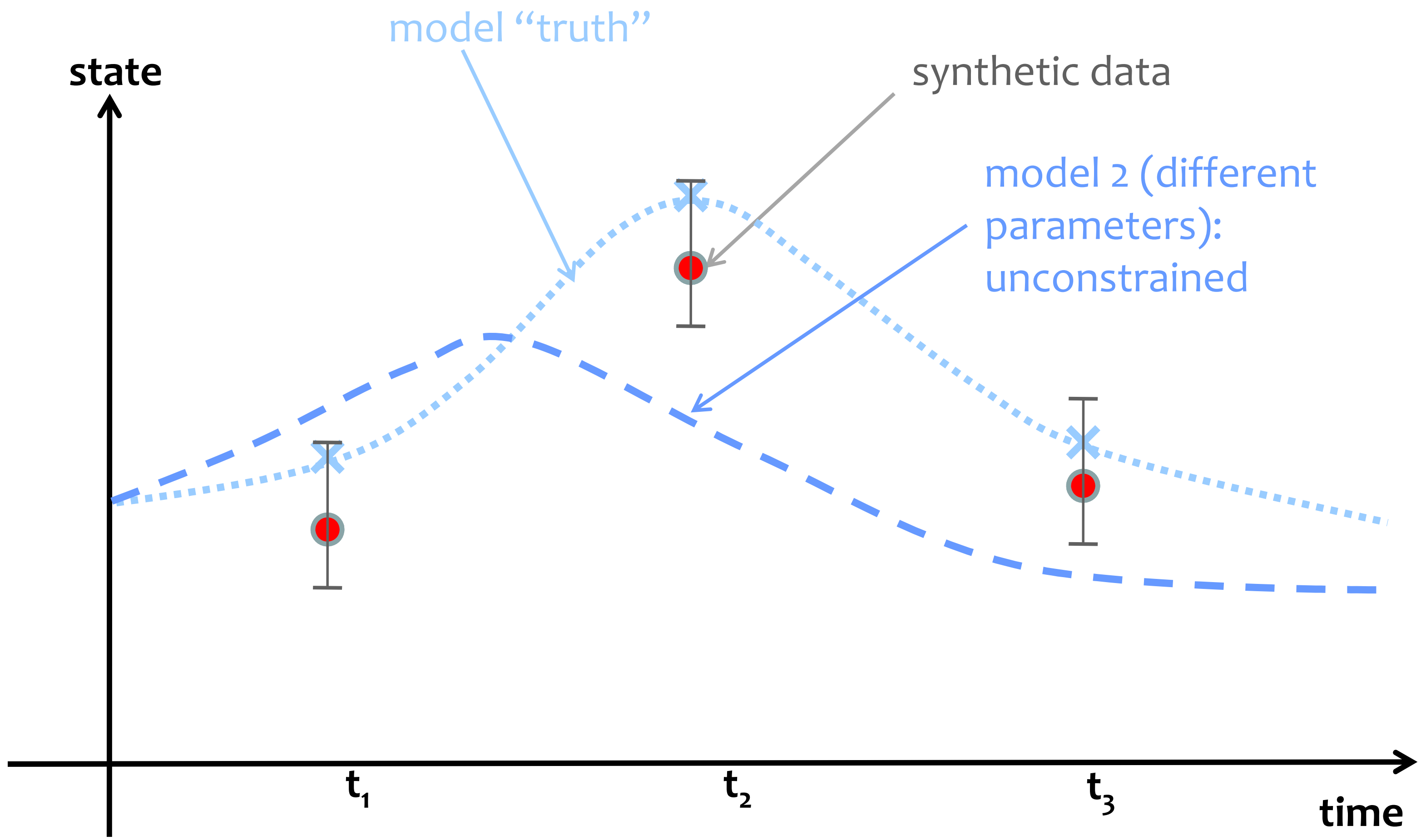


Massachusetts Institute of Technology

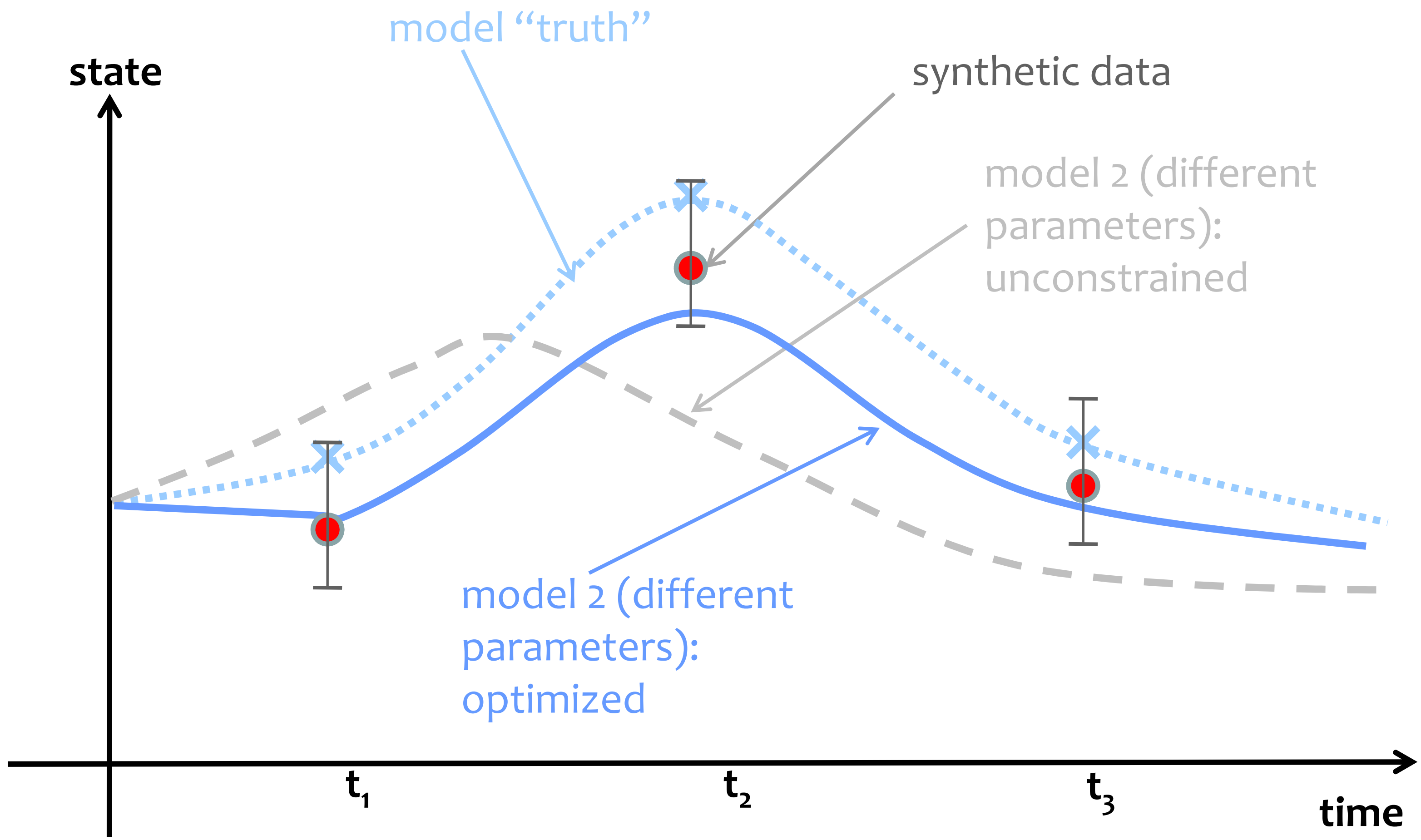
Observing System Simulation Experiment (OSSE)



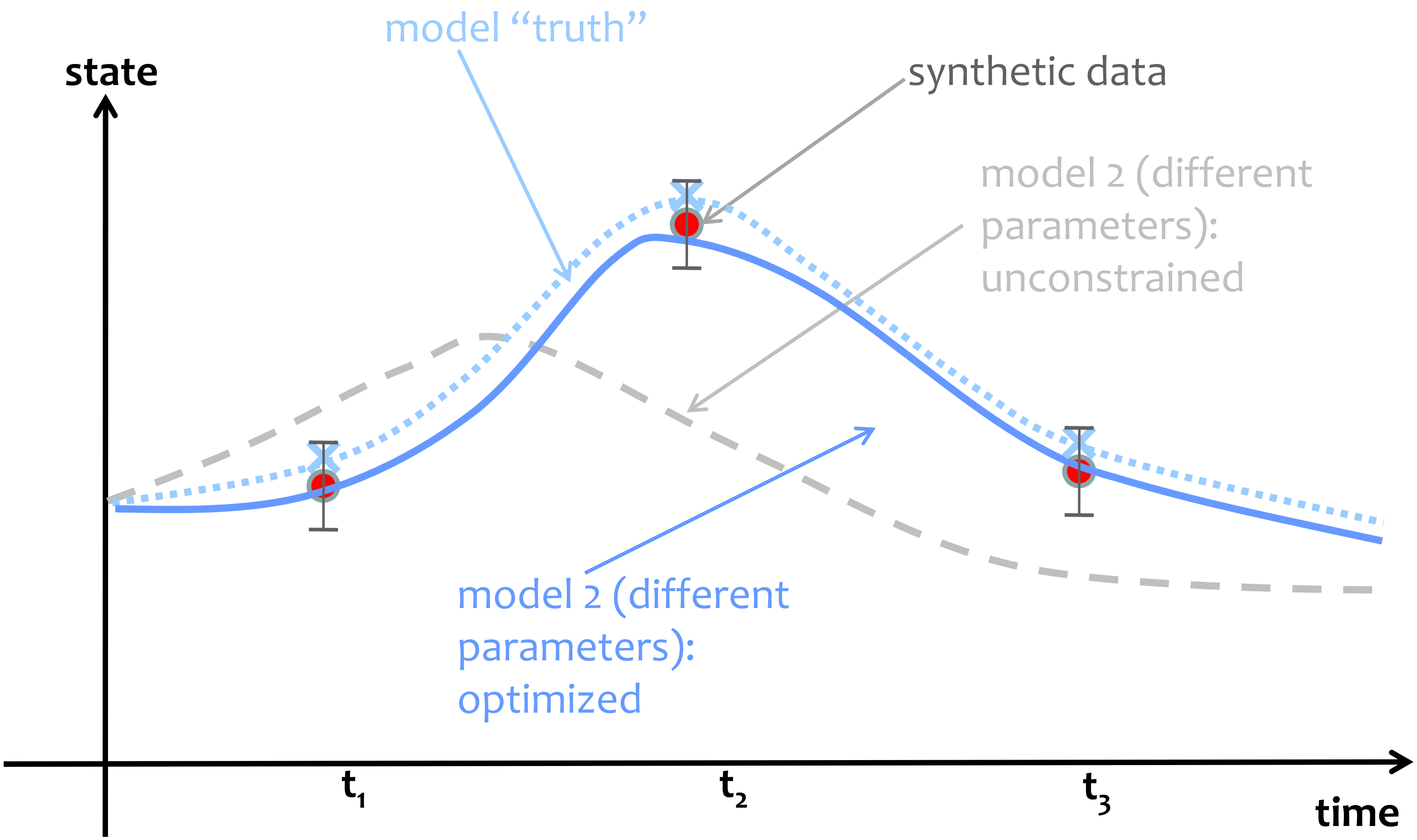
Observing System Simulation Experiment (OSSE)



Observing System Simulation Experiment (OSSE)

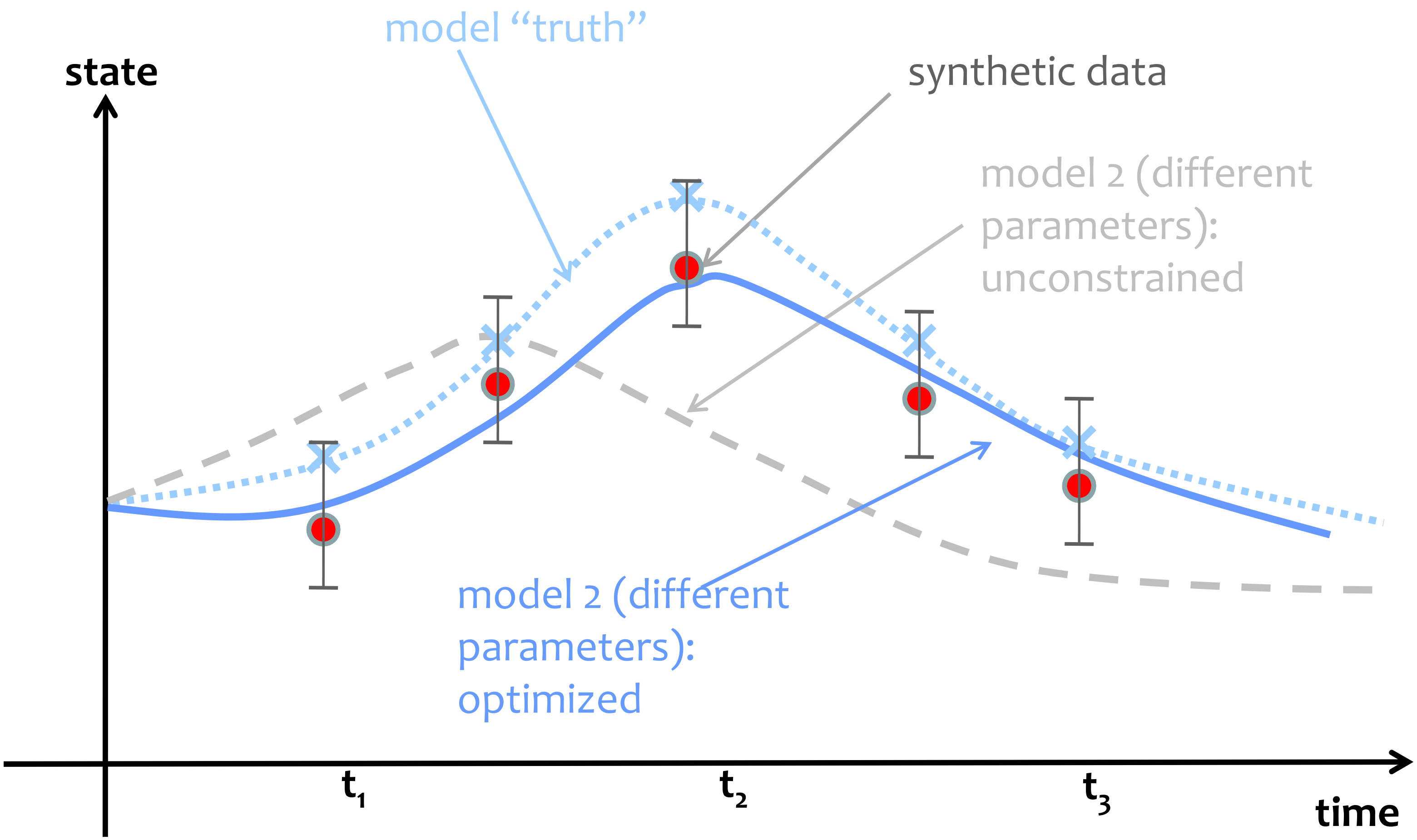


Observing System Simulation Experiment (OSSE)



Improvement with less noise in "synthetic data"

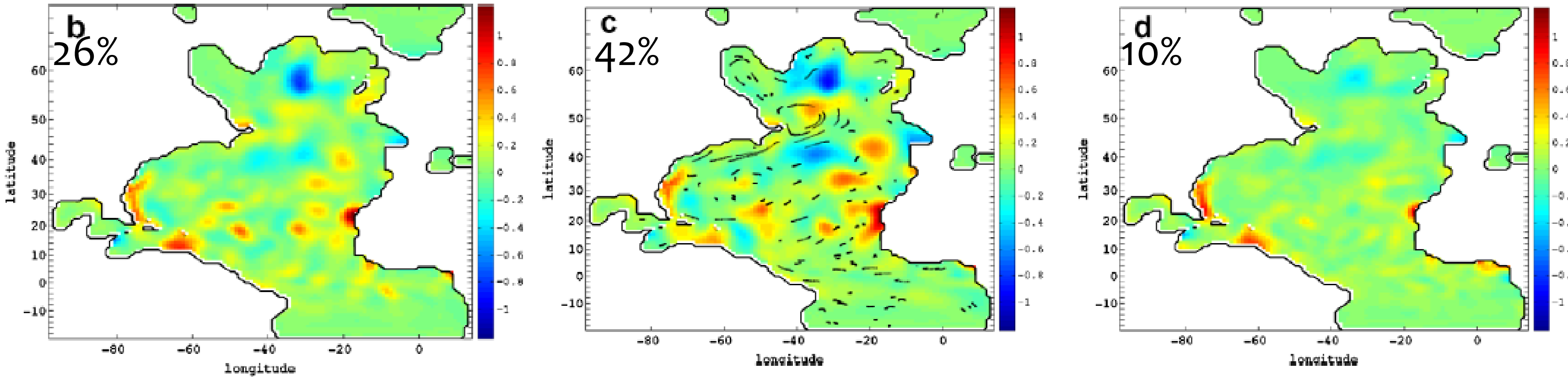
Observing System Simulation Experiment (OSSE)



Improvement with more "synthetic data"

OSSE experiment to explore ARGO deployment objectives:
Model sub-sampled as for ARGO deployment, noise added

Residual at 847.5m for T for different amounts of "ARGO floats"



ARGO-like (300) less floats (100) more floats (900)

- ARGO adequate as a large scale, low frequency observing system

- non-assimilated variables also benefit:
  velocities improved relative to free model by 2 between 300 and 900 floats

Forget et al, OM, 2008

Models feedback to observing strategies:

Some potential questions for OSSE for biogeochemical autonomous network:

- What is optimal vertical sampling strategy?
- What are optimal combinations of mission lifespan versus number sensors?
- Which regions are most crucial to sample?

MODEL-DATA SYNTHESIS

- Data assimilation methods:
  - brings together observations with the 4-D capabilities of models to provide best "state estimates" of system
  - powerful method to bring together diverse data sets
  - ARGO essential for many metrics in the physical oceanography system (but multiple observing systems needed)
  - can be used to refute, compare and improve models

- Global biogeochemical data assimilation requires:
  - better coverage of sub surface ocean
  - multiple diverse datasets
  - good physical state estimates

MODEL FEEDBACK TO OBSERVING SYSTEMS

- Models can be used to help us understand what observing systems are capable of telling us
- Data assimilation frameworks can be used for OSSEs to help design optimal observational network

Societal relevant questions:

1) monitoring of oceanic ecosystems health
(e.g. changes on interannual and future timescales, effects on fisheries)
2) monitoring of ocean carbon cycle
(e.g. greenhouse gas emissions verification)

We will need data assimilated models to address these questions;

But need autonomous network for essential measurements with which to constrain the model