Data-model assimilation for manipulative experiments

Dr. Yiqi Luo
Botany and microbiology department
University of Oklahoma, USA
Manipulative experiments

More manipulative experiments under planning by NEON and DOE.
What is the nature of manipulative experiments?
How can we extrapolate results from the Duke FACE experiment to predict long-term, large-scale change
Luo and Reynolds
1999, Ecology, 80:1568-1583
Observed C and N dynamics in FACE experiments are not applicable to natural ecosystems in response to a gradual CO$_2$ increase.

Luo and Reynolds 1999
Results of manipulative experiments cannot be simply extrapolated to predict ecosystem responses to global change in the real world.
Luo and Reynolds (1999)

“Rigorous analysis of (results from) step experiments requires not only statistical but also other new approaches, such as deconvolution and inverse modeling”
Data-model assimilation at Duke FACE

**Tool development**

1. Deconvolution (Luo et al. 2001)
2. Adjoint function (White and Luo. 2002)
3. Stochastic inversion, Xu et al. 2006
4. Step-wise inversion, Wu et al. (in review)
5. Linear, nonlinear, ensemble Kalman Filter (Gao et al. see poster)

**Applications**

3. Uncertainty analysis, Xu et al. 2006
4. Forecasting of carbon sequestration
Framework for Uncertainty analysis

Measurement errors

Stochastic inversion

Variability in estimated parameter values

Forward model

Uncertainty in model predictions

Observation error (umol m⁻² s⁻¹)

Frequency

Error distribution
Norman Distribution
Double exponential distribution

GEE(µmol m⁻² s⁻¹)

Frequency

Rₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉᵉ>e

Re,ref(µmol m⁻¹ s⁻¹)

Frequency

Relative uncertainty (%)
1. Matrix to describe C flow

\[
\frac{dX}{dt} = \tau AX + Bu
\]

2. Mapping functions

\[Q'(A)(t) = q'(A)(t) \cdot X(A)(t)\]

3. Cost function

\[
J(A) = \sum_{j=1}^{m} \nu_j \left[ \sum_{i=1}^{n_j} (Q^j(A)(t_i) - Q_0^j(t_j))^2 \right]
\]

4. Search method

MCMC–Metropolis-Hastings algorithm
Criteria I: Data-model fitting

Luo et al. 2003, GBC
Criteria II: Probability Distribution

Prior knowledge

Posterior distribution

Inverse model

Well-constrained

Edge-hitting

No-information

Observed Data

Histogram of generated samples for $c_2$

Range of $c_2$

Sampling frequency

Histogram of generated samples for $c_3$

Range of $c_3$

Sampling frequency

Histogram of generated samples for $c_5$

Range of $c_5$

Sampling frequency
Uncertainty analysis

Variability in estimated parameter values

Multiple data sets

Stochastic inversion

Variability in estimated parameter values

Xu et al. 2006, GBC
\[ \ln(Y) = 0.966 \ln(X) + 0.215 \]

\[ R^2 = 0.969 \]

Luo et al. 2003, GBC
Model predictions

- Multiple data sets
- Stochastic inversion
- Variability in estimated parameter values
- Forward model
- Uncertainty in model predictions

Xu et al. 2006, GBC
Estimated initial values of pools and residence times to partition C sink to two components caused by climate change and forest regrowth.
Uncertainty Analysis

1. Magnitudes (50%, 100%, and 200%) of measurement errors (Weng et al. poster)

2. Distributions (Normal vs. double exponential) of measurement errors of eddy-flux data (Liu et al. in review)

3. Different assimilation algorithms (least squares, maximal likelihood, and Kalman filter (Gao et al. poster)

4. Continental analysis on residence times (Tao Zhou et al. in review), $Q_{10}$ values (Tao Zhou et al. in review, and their uncertainties (Xuhui Zhou et al. poster)

Applications

Measurement errors

\[ \text{Stochastic inversion} \]

Variability in estimated parameter values

\[ \text{Forward model} \]

Uncertainty in model predictions

1. Magnitudes (50%, 100%, and 200%) of measurement errors (Weng et al. poster)

2. Distributions (Normal vs. double exponential) of measurement errors of eddy-flux data (Liu et al. in review)

3. Different assimilation algorithms (least squares, maximal likelihood, and Kalman filter (Gao et al. poster)

4. Continental analysis on residence times (Tao Zhou et al. in review), $Q_{10}$ values (Tao Zhou et al. in review, and their uncertainties (Xuhui Zhou et al. poster)
Summary

1. Data from manipulative experiments cannot be directly extrapolated to the real world. We have to extract information from data on fundamental processes.

2. Model assimilation of multiple data sets is one of the best approaches to synthesis of experimental results with processing thinking and can better balance evidence from different lines.
Acknowledgement

**Idea and math development**
- Luther White
- James Reynolds
- S. Lakshmivarahan
- Dafeng Hui
- Tao Xu
- Tao Zhou
- Ensheng Weng
- Chao Gao
- Ensheng Weng
- Li Zhang
- Min Liu

**Data from the Duke FACE**
- Schlesinger
- Ellsworth
- Finzi
- DeLucia
- Katul
- Oren

**Financial support**
- DOE TCP
- NSF

http://bomi.ou.edu/luo