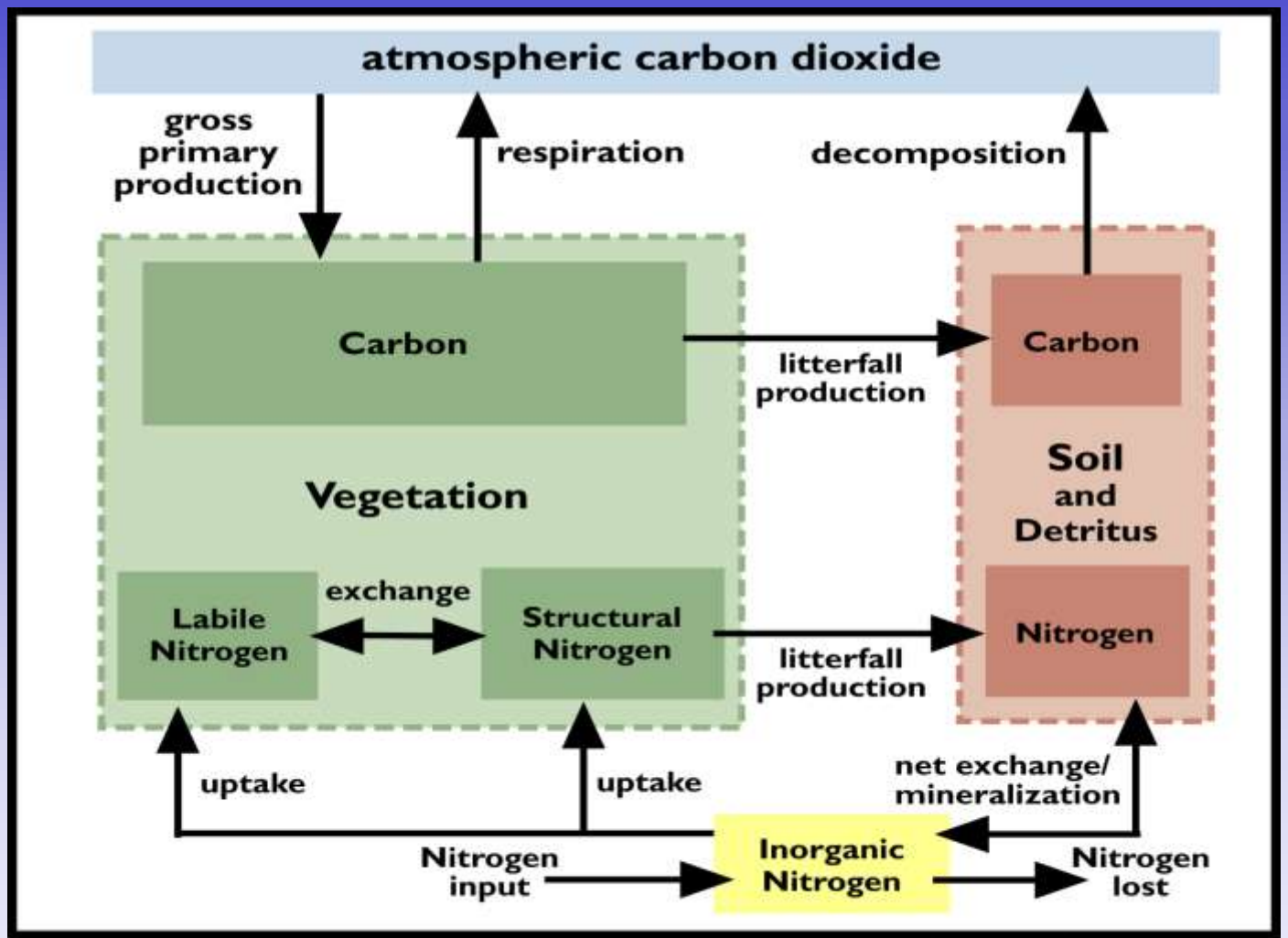
Quantification of Carbon and Water Dynamics in North America Using a Process-Based Ecosystem Model and a Bayesian Approach

Zhuang, Q., J. Tang, X. Lu, M. Chen, and Y. Jiang
Purdue University

With Contributions from Ameriflux Investigators

Improve Ecosystem Model Predictability

The Terrestrial Ecosystem Model



Terrestrial Ecosystem Model (TEM)

The Ecosystems Center, Marine Biological Laboratory (Woods Hole, Massachusetts)

Bayesian Inference Approach

- Key idea: Infer the probability of parameters (θ) conditioning on the prior probability of available data (V)

$$Pr(\theta|V) = \frac{Pr(\theta)\ Pr(V|\theta)}{Pr(V)}$$

Pr(θ|V) is the posterior

Pr(θ) is the prior of the ecosystem model parameters (uniform or Gaussian etc.)

Pr(V|θ) is the likelihood function that could be calculated based on model simulation and observation data of fluxes and states

Pr(V) does not depend on θ, and with fixed V, can thus be considered as a constant

Calculation of the Likelihood Function Pr(V|θ)

$$p(\mathbf{V}|\boldsymbol{\beta},\boldsymbol{\theta}) = \frac{1}{2^N}\prod_{i=1}^{N}\Gamma\left[(1+\beta_i)\left(T-\frac{1}{2}\right)\right]\left[\omega(\beta_i)\right]^T \cdot \left[c(\beta_i)\sum_{t=0}^{T}|v_{it}|^{2/(1+\beta_i)}\right]^{(1/2-T)(1+\beta_i)}$$

β are coefficients associated with observational data error distribution
T is model simulation temporal domain
N is number of state and flux variables
V_{ti} are the difference between model simulations and observations at time t for variable i

(Tang and Zhuang, 2009)

The Posterior Generation Using Sampling Importance Resampling Technique

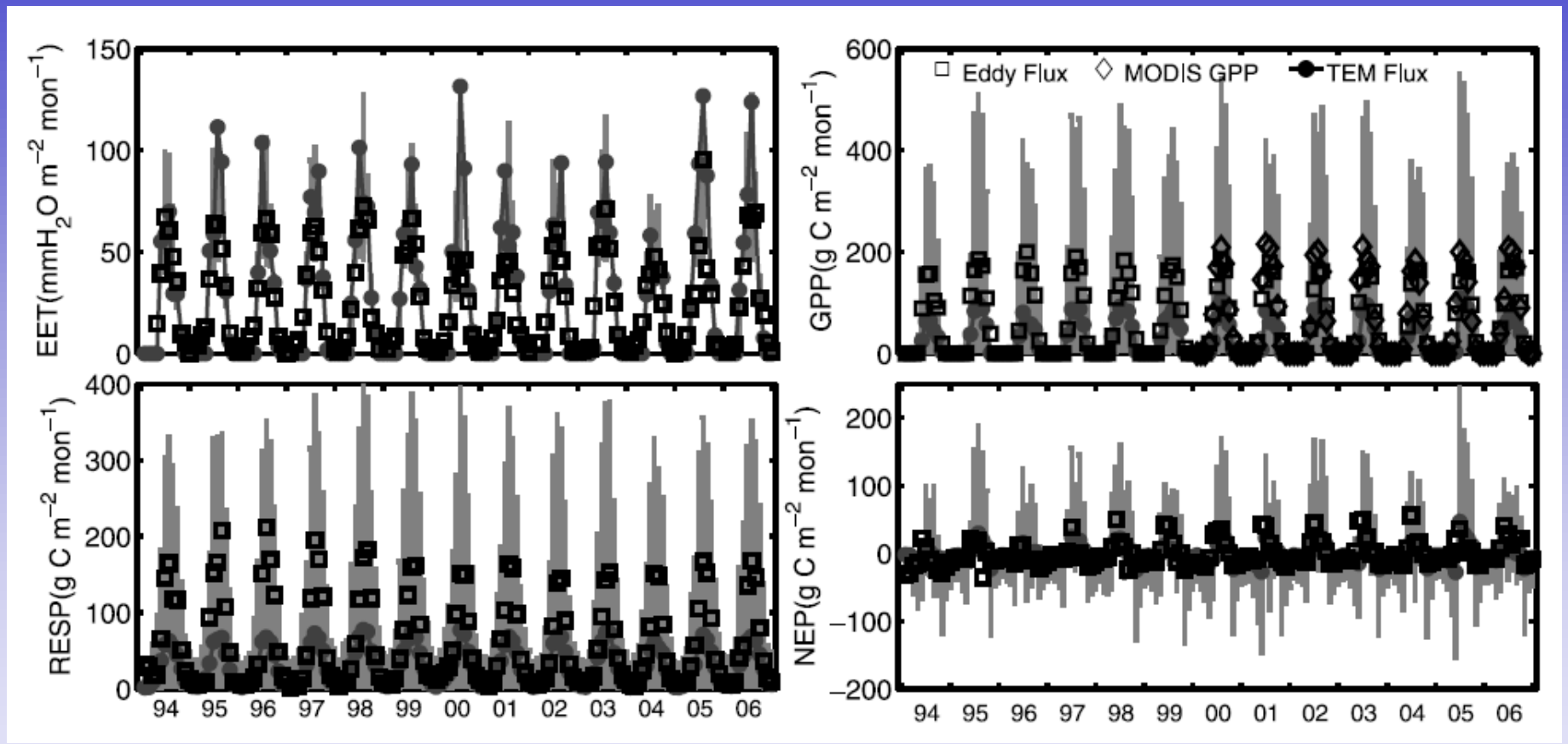
[26] (1) Computing the importance ratio $h(\boldsymbol{\theta}_j) = p(\boldsymbol{\theta}_j|\mathbf{V})/p(\boldsymbol{\theta}_j)$, using equation (6) for each sample, so that a vector of length 500,000 is formed. In this study, with the uniform prior, the importance ratio is just the value of likelihood function.

[27] (2) Let $S_{-j} = \sum_{i \neq j} h(\boldsymbol{\theta}_i)$ be the sum of all importance ratios excluding $h(\boldsymbol{\theta}_j)$.

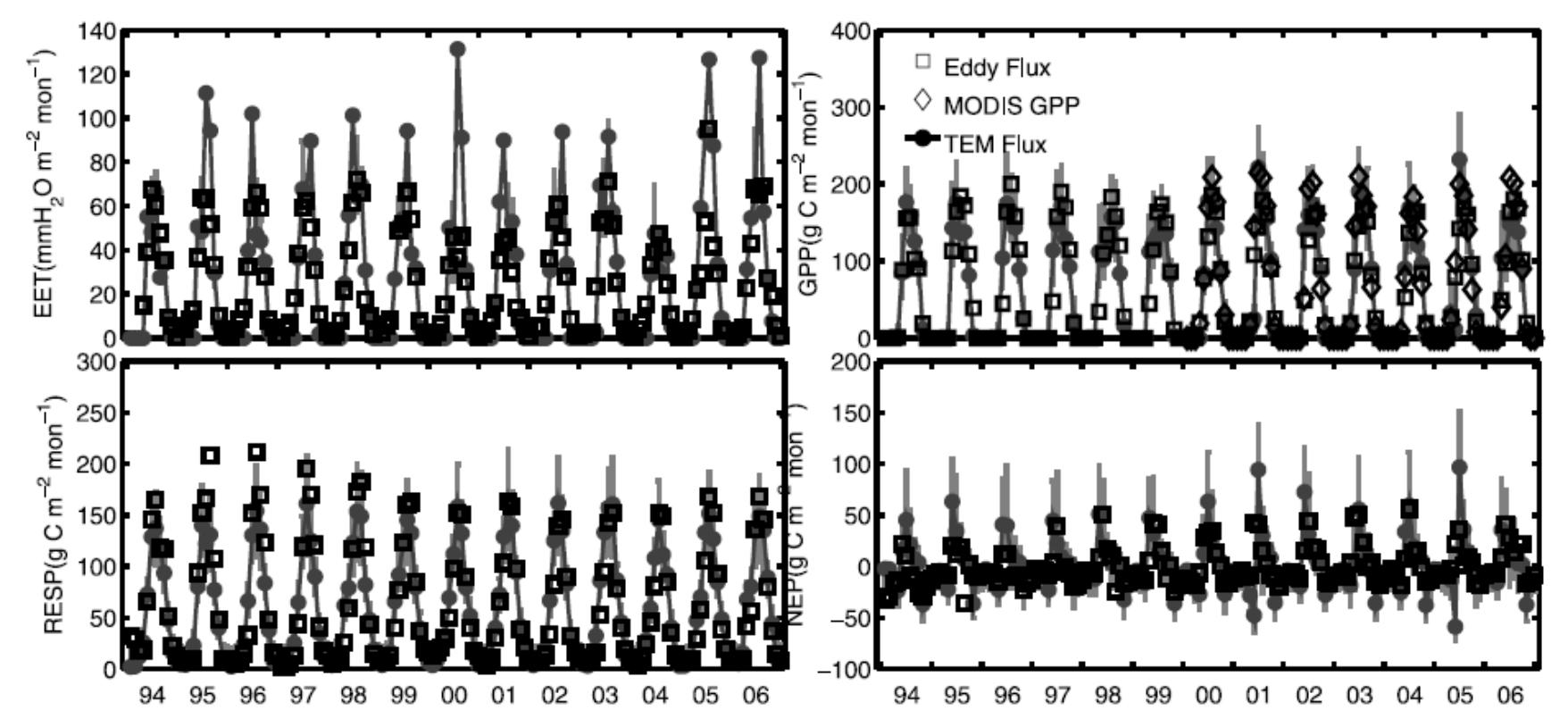
[28] (3) For $k = 1, \ldots, m$, draw parameter samples j_k with probability q_{j_k} from the 500,000 samples, where $q_j \propto h(\boldsymbol{\theta}_j)/S_{-j}$. We used the Probability Proportional to Size (PPS) sampling technique [*Hong et al.*, 2005] to draw the samples. The resampling was done with replacement.

[29] (4) For the picked j_k, get the corresponding $\boldsymbol{\theta}_{j_k}$, $k = 1, \ldots, m$. In our analysis, we set $m = 50{,}000$. This number was chosen according to the suggestion that, to produce stable results, setting the ratio of prior sample size to posterior sample size as 10 is appropriate [*Green et al.*, 1999; *Bates et al.*, 2000].

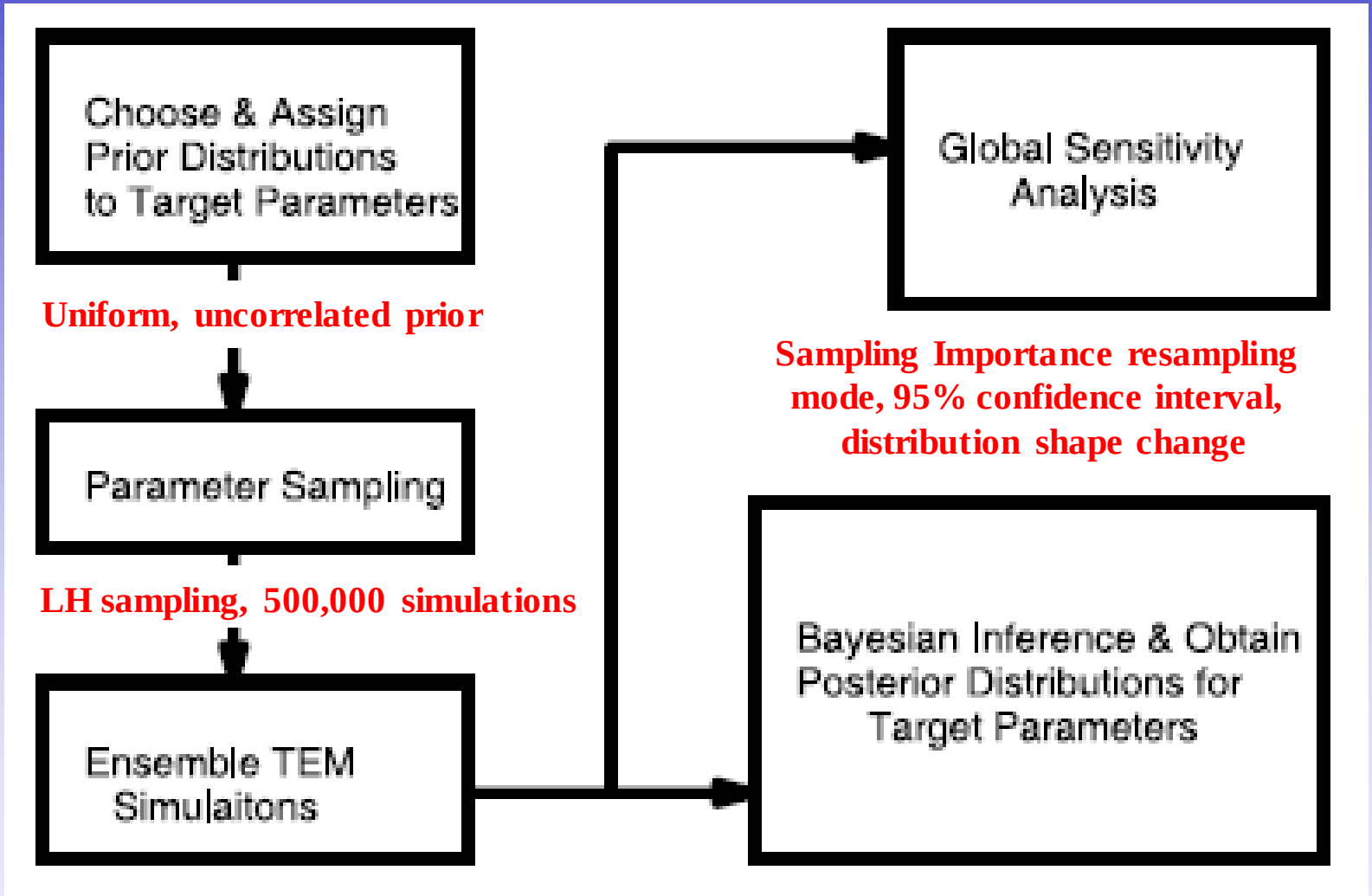
The Prior Fluxes



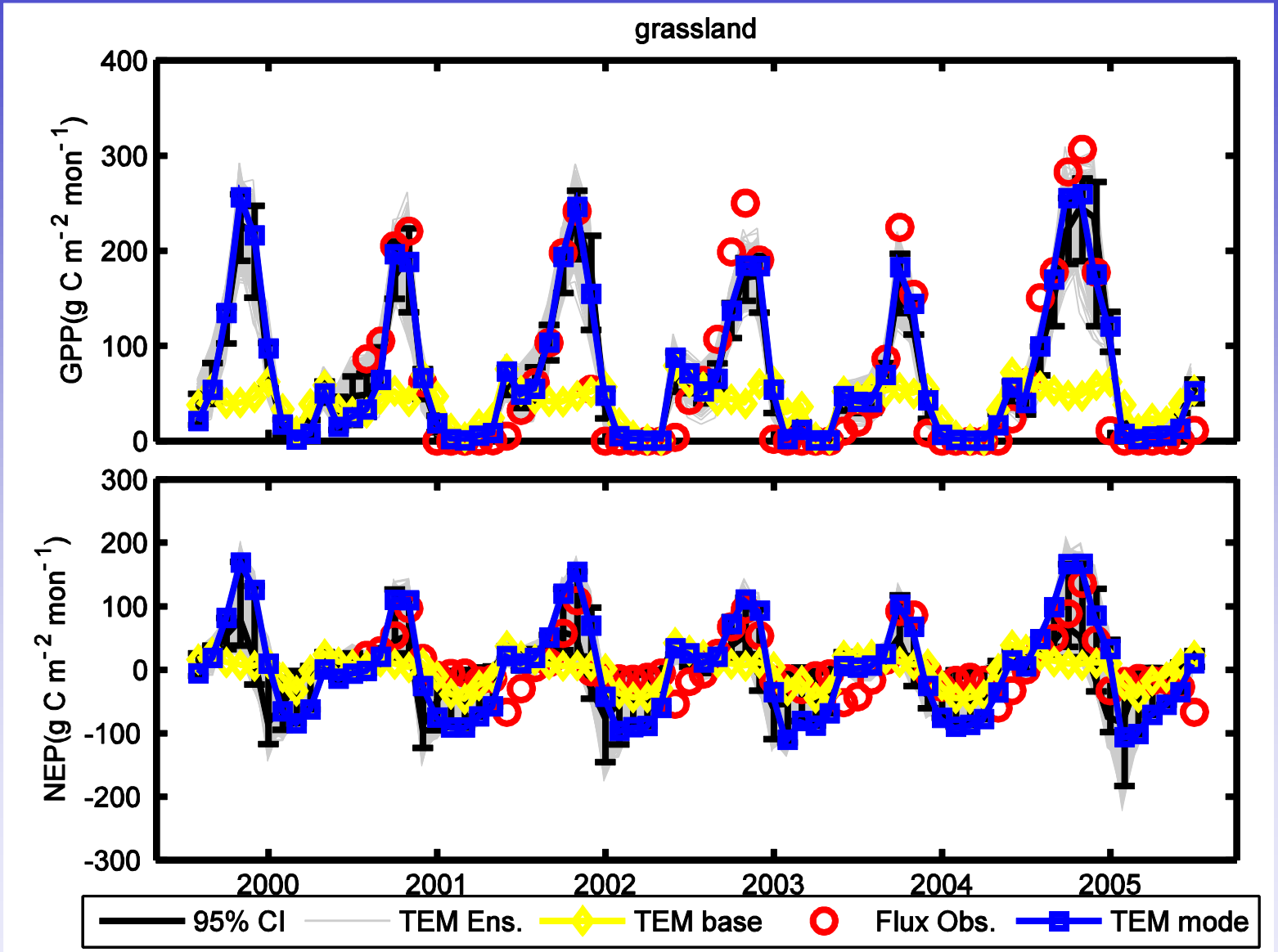
The Posterior Fluxes



Workflow of the Bayesian and Ecosystem Modeling Approach



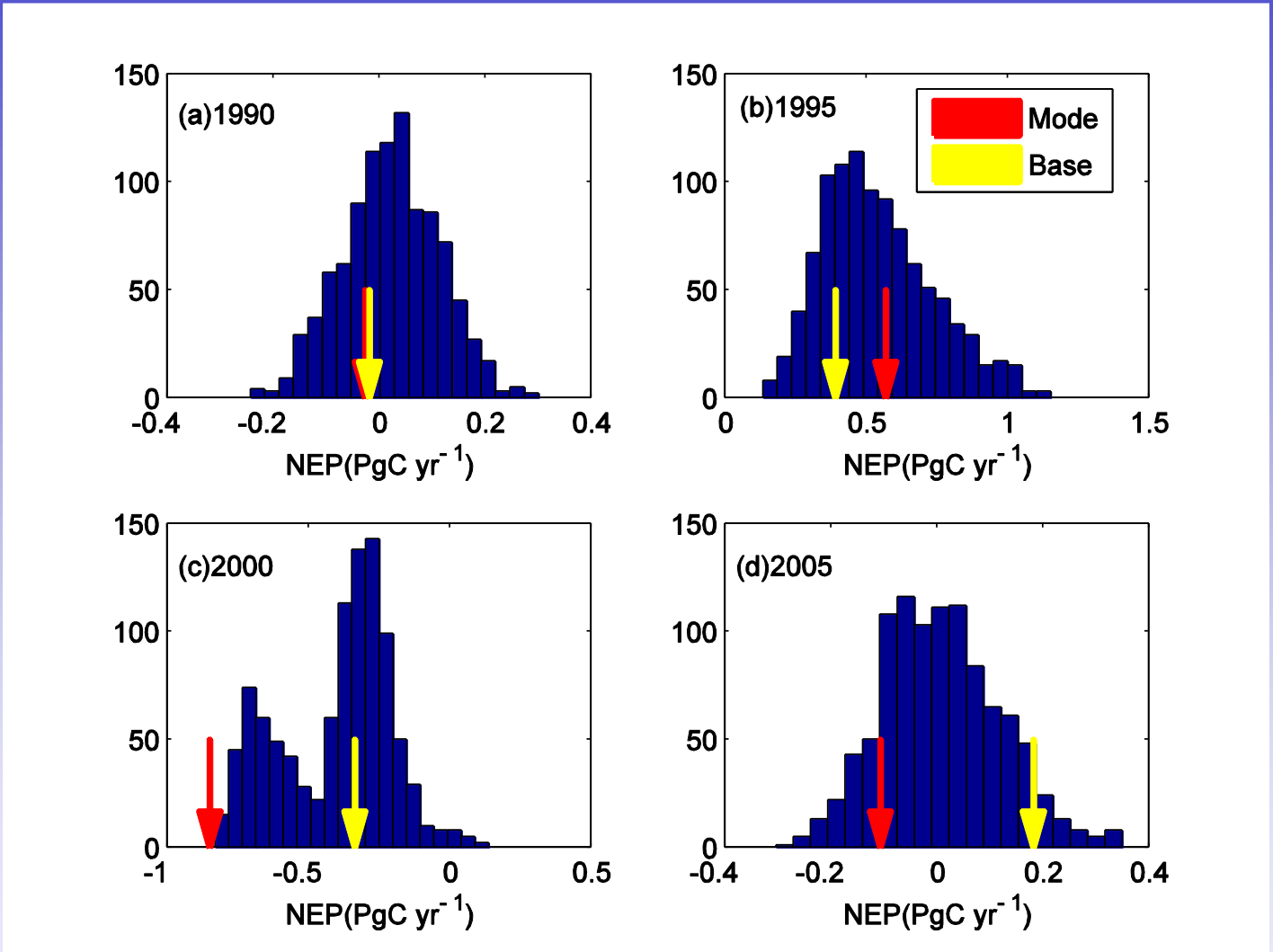
Application of the Bayesian and Ecosystem Modeling Approach to a Grassland Ecosystem



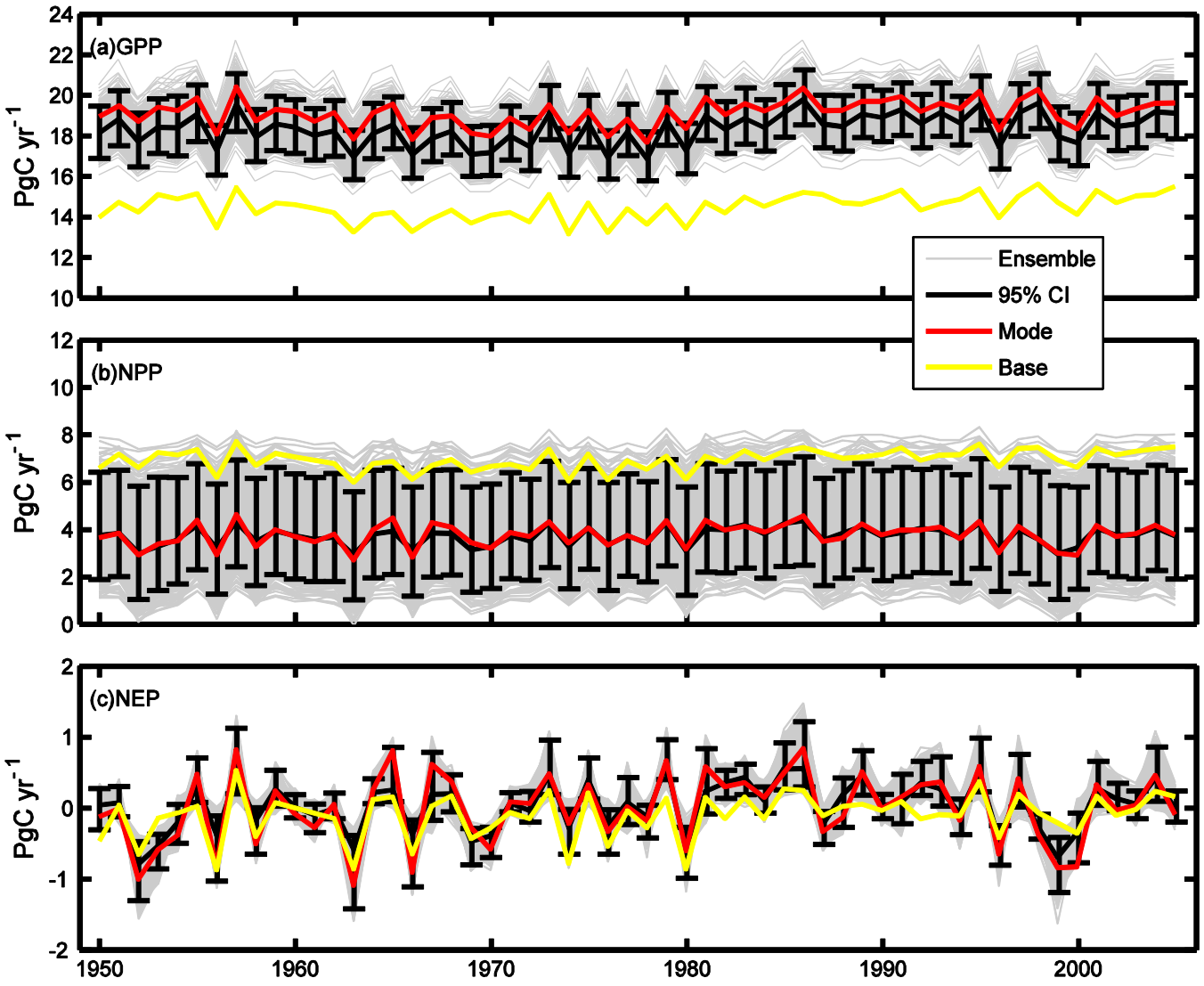
Parameterization for Eight Ecosystem Types with the Approach

1. **Grassland - Vaira Ranch (*Baldocchi et al., 2001*)**
2. **Boreal forest - Niwot Ridge (*Monson et al. 2002*)**
3. **Xeric woodland - Tonzi Ranch (*Ma et al., 2007*)**
4. **Temperate broadleaf forest - Kennedy Space Center Scrub Oak (*Bracho et al., 2008*)**
5. **Temperate coniferous forest - Howland Forest West Tower (*Hollinger et al., 1999*)**
6. **Temperate deciduous forest - Harvard Forest (*Urbanski et al., 2007*)**
7. **Wet tundra – Atqasuk (*Oberbauer et al., 2007*)**
8. **Xeric shrubland - Tonzi Ranch (*Ma et al., 2007*)**

Simulated Carbon Fluxes in North America



Simulated Carbon Fluxes in North America



Summary

- Bayesian approach with eddy flux data and associated error information provides a new way to parameterize ecosystem model

- With multiple-site data, the Bayesian approach estimates the “mode” regional sink or source, differing from the quantification with the traditional calibration method based on annual flux data

Next Step

- Use the Bayesian approach to improve a finer time step ecosystem model with finer resolution observational data (e.g., daily)
- Use satellite data to constrain the regional uncertainty based on ensemble ecosystem model simulations

Acknowledgments

1. Eddy flux measurement community

2. Supported in part by:
   NASA LCLUC Program
   NSF Carbon and Water Program