Data assimilation: a family of techniques broadly based on estimation theory

Most familiar in *forecast* applications

Also used in *analysis* applications

Most carbon cycle/biogeoscientific applications are analyses

Ecosystem assimilation strategy

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$

α_i = forcing, examples are temperature and radiation.

ρ_i = time-invariant model parameters. Examples are the temperature exponent for respiration, light dependence of photosynthesis.

β_i are internal state variables of the system and are a function of prior integrated forcing (β_i can have long response times). Examples are foliar C and N, soil water, litter C and N.

Assimilation strategy

$$NEE = f(\alpha_i,\beta_i,\rho_i)$$

$$NEE = GPP(\alpha_{gpp},\beta_{gpp},\rho_{gpp}) - R\,(\alpha_r,\beta_r,\rho_r)$$

α_i = forcing, examples are temperature and radiation.

ρ_i = time-invariant model parameters. Examples are the temperature exponent for respiration, light dependence of photosynthesis.

β_i are internal state variables of the system and are a function of prior integrated forcing (β_i can have long response times). Examples are foliar C and N, soil water, litter C and N.

Assimilation strategy

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$

$$NEE = GPP(\alpha_{gpp}, \beta_{gpp}, \rho_{gpp}) - R(\alpha_r, \beta_r, \rho_r)$$

$$R(\alpha_r, \beta_r, \rho_r) = R_a(\alpha_{ra}, \beta_{ra}, \rho_{ra}) + R_h(\alpha_{rh}, \beta_{rh}, \rho_{rh})$$

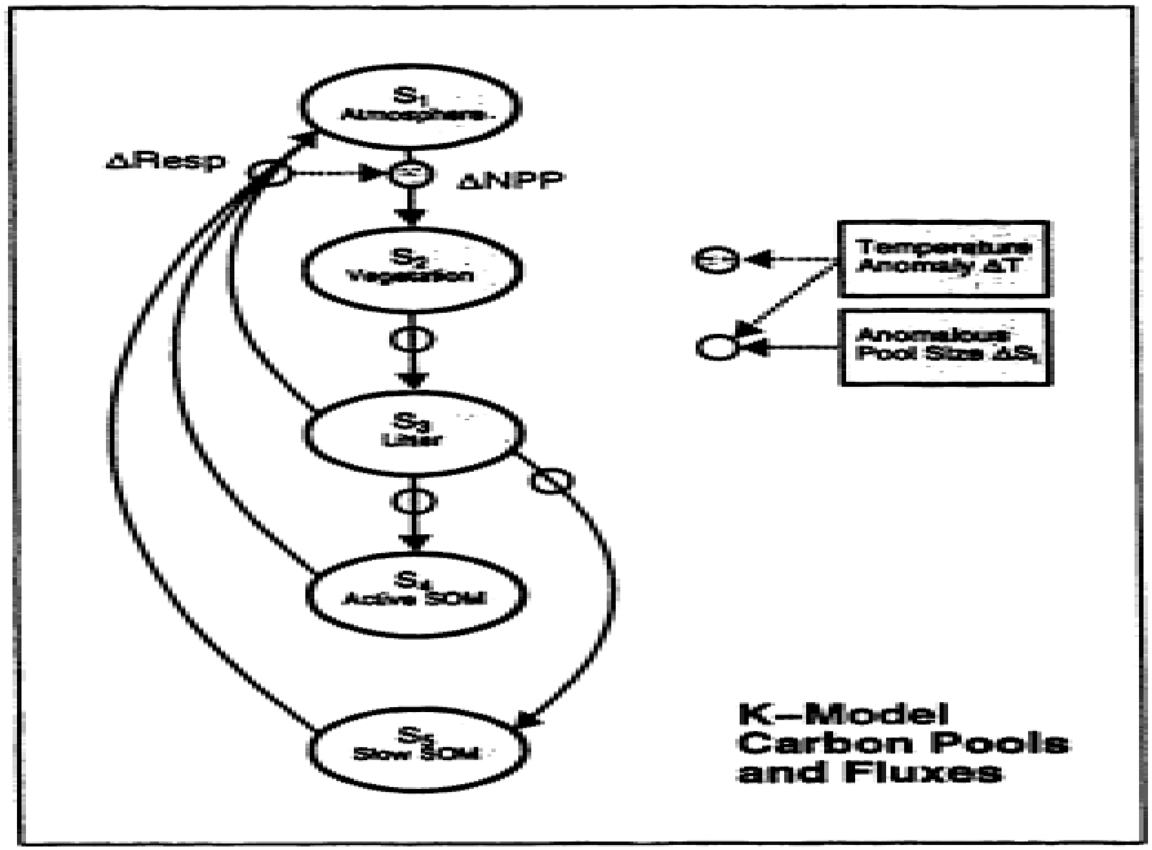
Two strategies based on estimation theory:

Parameter estimation:

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$

State space estimation:

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$



A simple model of the sensitivity of terrestrial ecosystem physiology to temperature

Parameters of the model and prior bounds

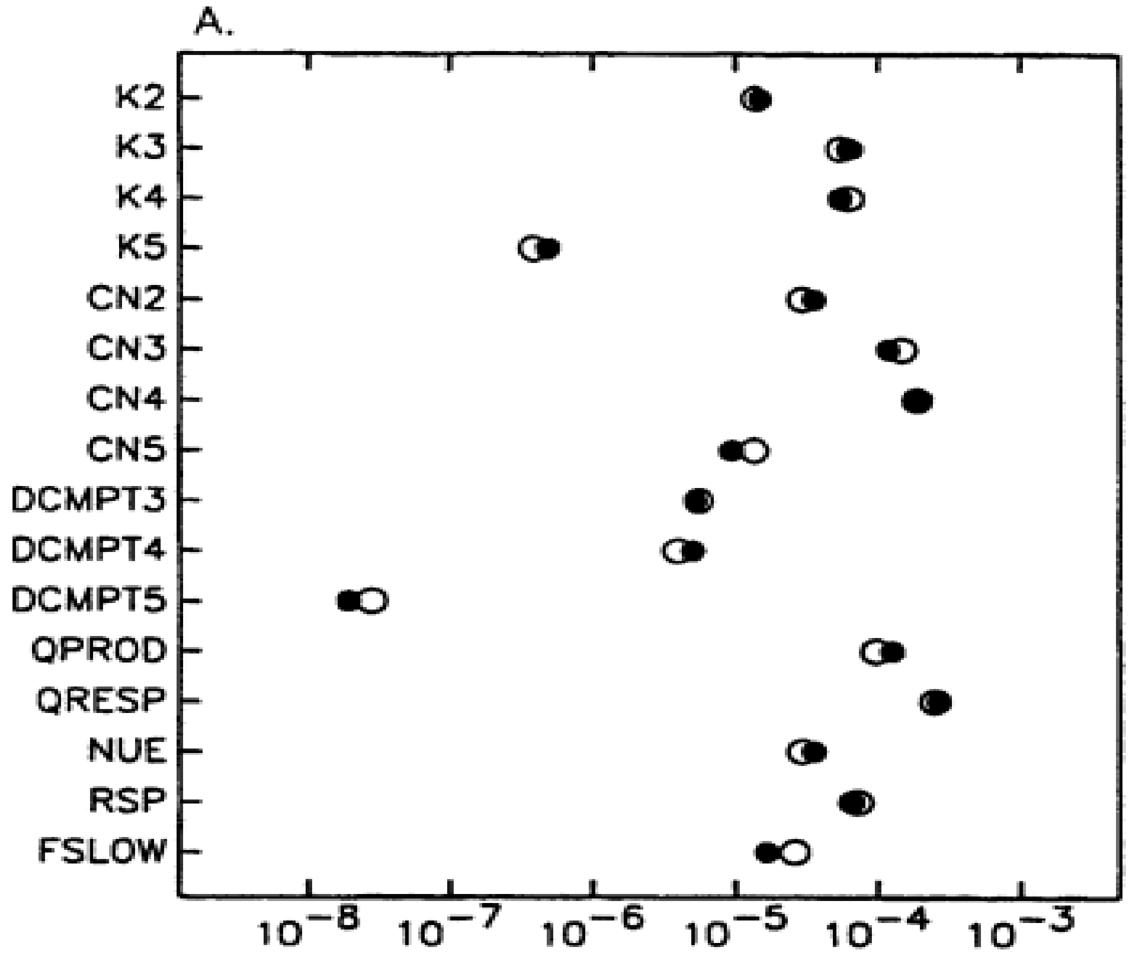
Table 1. *K-model parameters*

Name	Description	Retrieved *N*-cycle	Retrieved no-*N*-cycle	Interval of acceptable values	Units
k_2	vegetation turnover rate	0.17	0.11	0.04–0.17	$month^{-1}$
k_3	litter turnover rate	0.042	0.0083	0.0083–0.042	$month^{-1}$
k_4	active SOM turnover rate	0.014	0.031	0.014–0.042	$month^{-1}$
k_5	passive SOM turnover rate	0.00083	0.0042	0.00083–0.0042	$month^{-1}$
cn_2	C:N vegetation	66.3	N/A	25.0–75.0	gg^{-1}
cn_3	C:n litter	50.0	N/A	50.0–150.0	gg^{-1}
cn_4	C:N active SOM	37.5	N/A	12.5–37.5	gg^{-1}
cn_5	C:N passive SOM	22.5	N/A	7.5–22.5	gg^{-1}
ϕ_3	slope of k_3 versus temperature	1.0	1.0	0.001–1.00	$month^{-1}\ deg^{-1}$
ϕ_4	slope of k_4 versus temperature	1.0	1.0	0.001–1.00	$month^{-1}\ deg^{-1}$
ϕ_5	slope of k_5 versus temperature	0.43	1.0	0.001–1.00	$month^{-1}\ deg^{-1}$
ϑ_p	slope of NPP versus temperature	1.66	1.22	0.001–2.00	$gm^{-2}\ year^{-1}\ deg^{-1}$
ϑ_r	slope of resp. versus temperature	2.00	1.60	0.001–2.00	$gm^{-2}\ year^{-1}\ deg^{-1}$
$\mathcal{F}$	strength of *N*-feedback	0.59	N/A	0.001–1.00	dimensionless
$\mathcal{Q}$	resp. fraction of decomposition	0.45	0.49	0.20–0.80	dimensionless
β	slow versus passive partitioning	0.80	0.54	0.20–0.80	dimensionless

Retrieved parameter values using the K-model version with the *N*-cycle included are listed in the column labeled "retrieved *N*-cycle". The retrieved parameter values using the K-model version without the *N* cycle included are listed in the column labeled "retrieved no-*N*-cycle. The interval of acceptable values for parameters are listed in the column labeled "interval of acceptable values".

Tellus 53B (2001), 2

Model sensitivity from the adjoint: change in NEE for a 1% parameter change



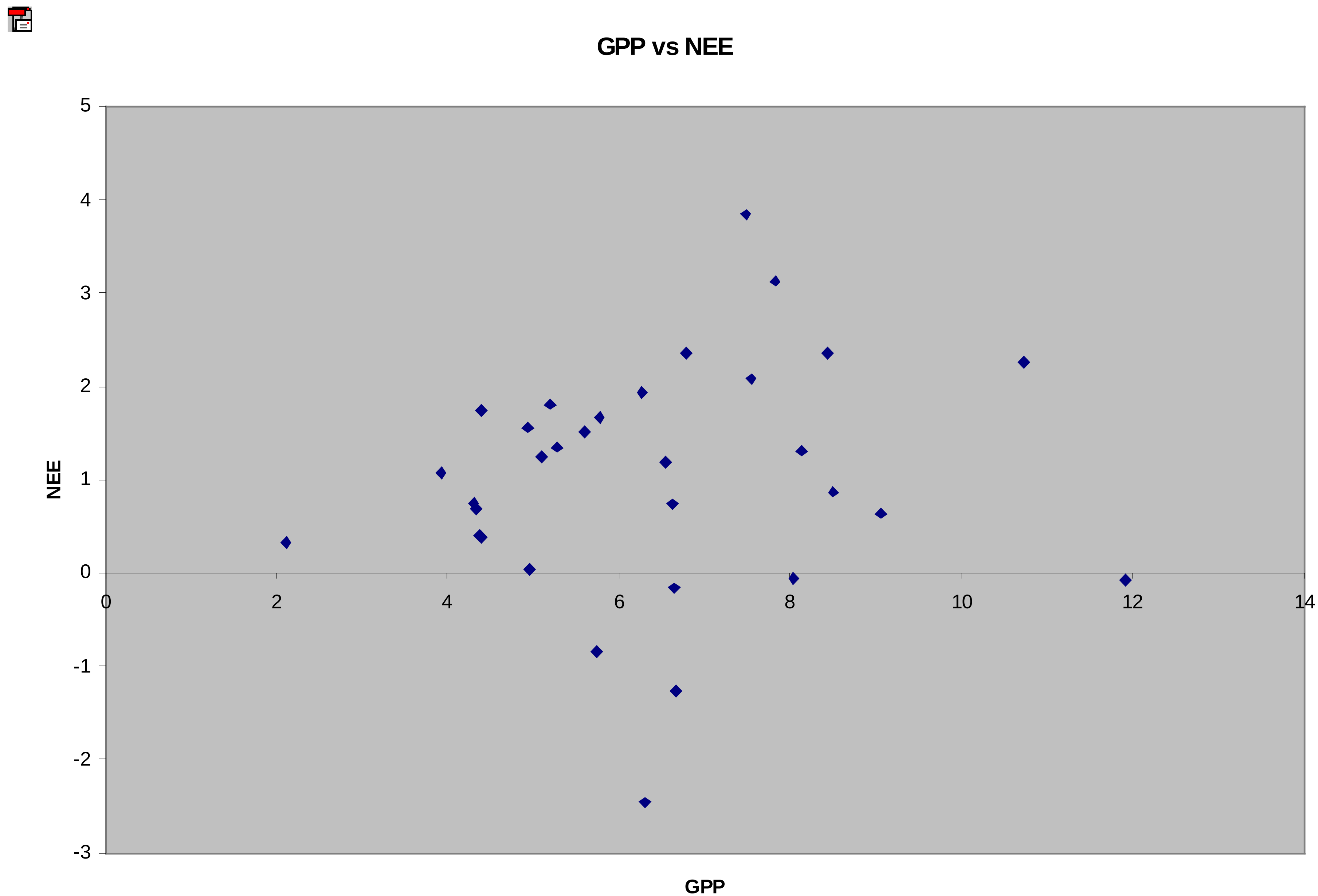
Two strategies based on estimation theory:

Parameter estimation:

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$

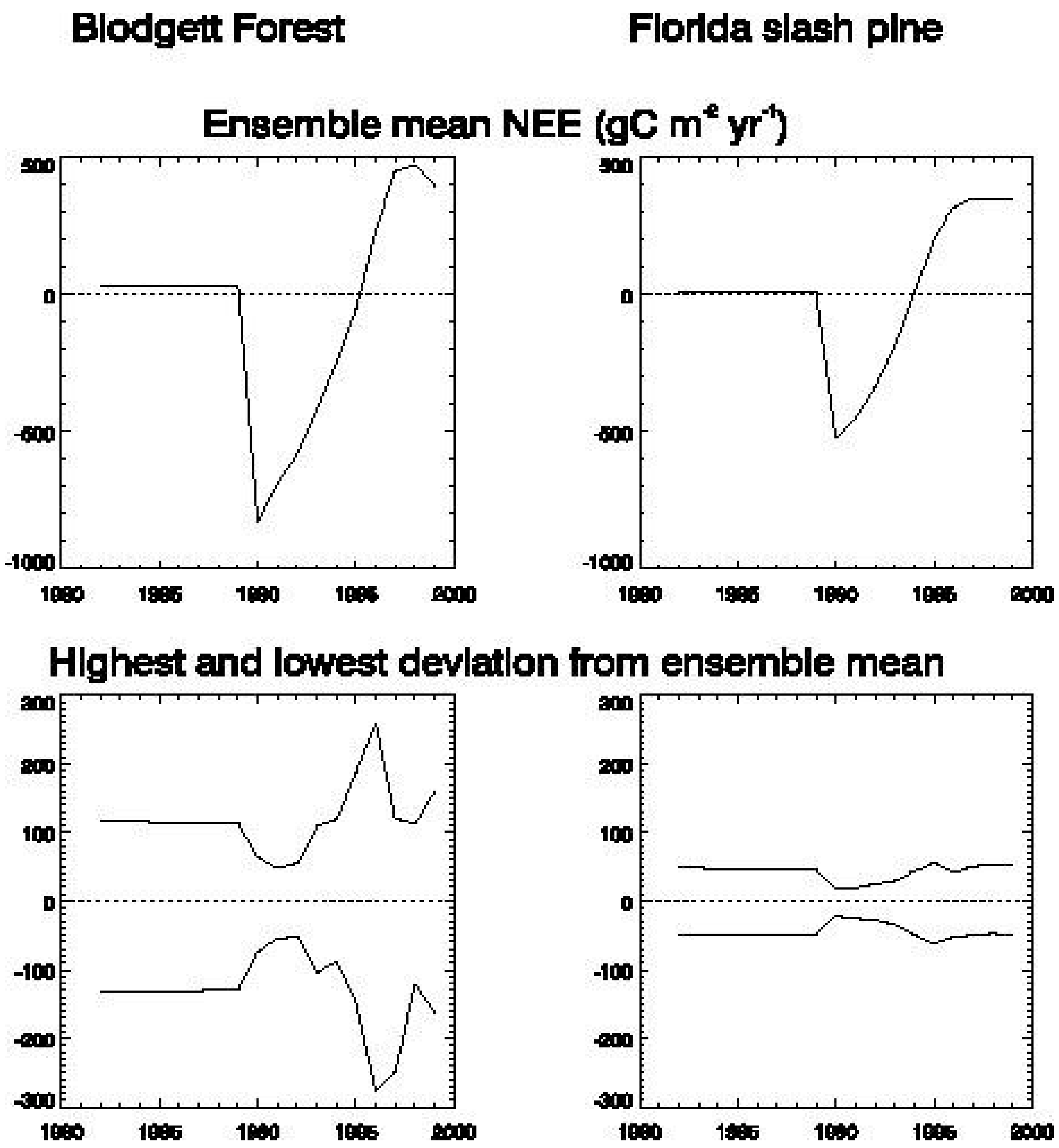
State space estimation:

$$NEE = f(\alpha_i, \beta_i, \rho_i)$$



Average daily NEE and GPP are uncorrelated: On this time scale NEE reflects longer time scale processes and prior forcing

NEE depends
On ecosystem
state, on
multiple
time scales



Kalman Filter

(forward pass)

Observation equation:

NEE_{obs} = GPP-R_a – R_h

Gain Matrix

σ(model) from validation

σ(data) from u*

Kalman Smoother

(backwards in time)

Uses

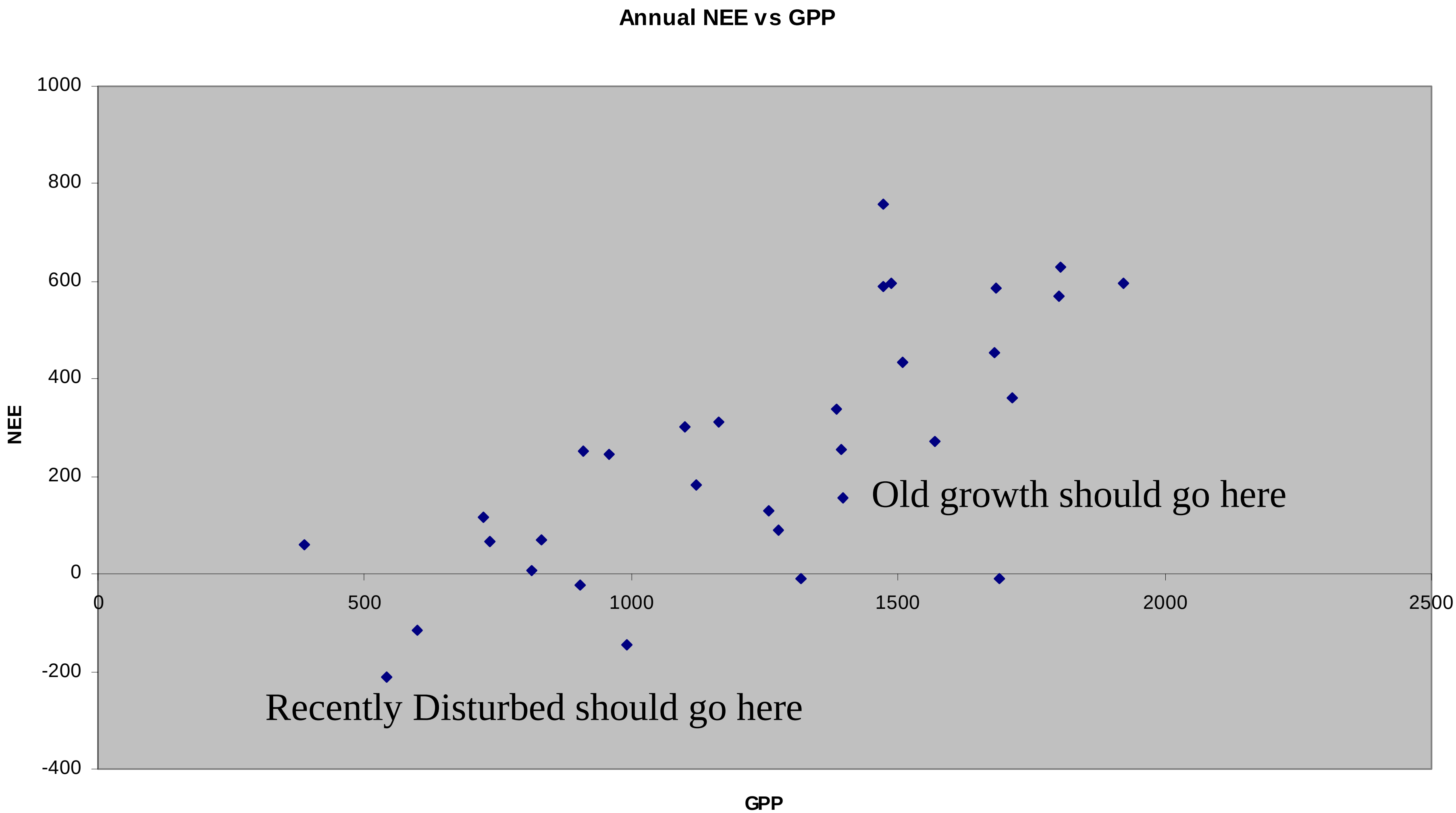
δNEE

δ β_i

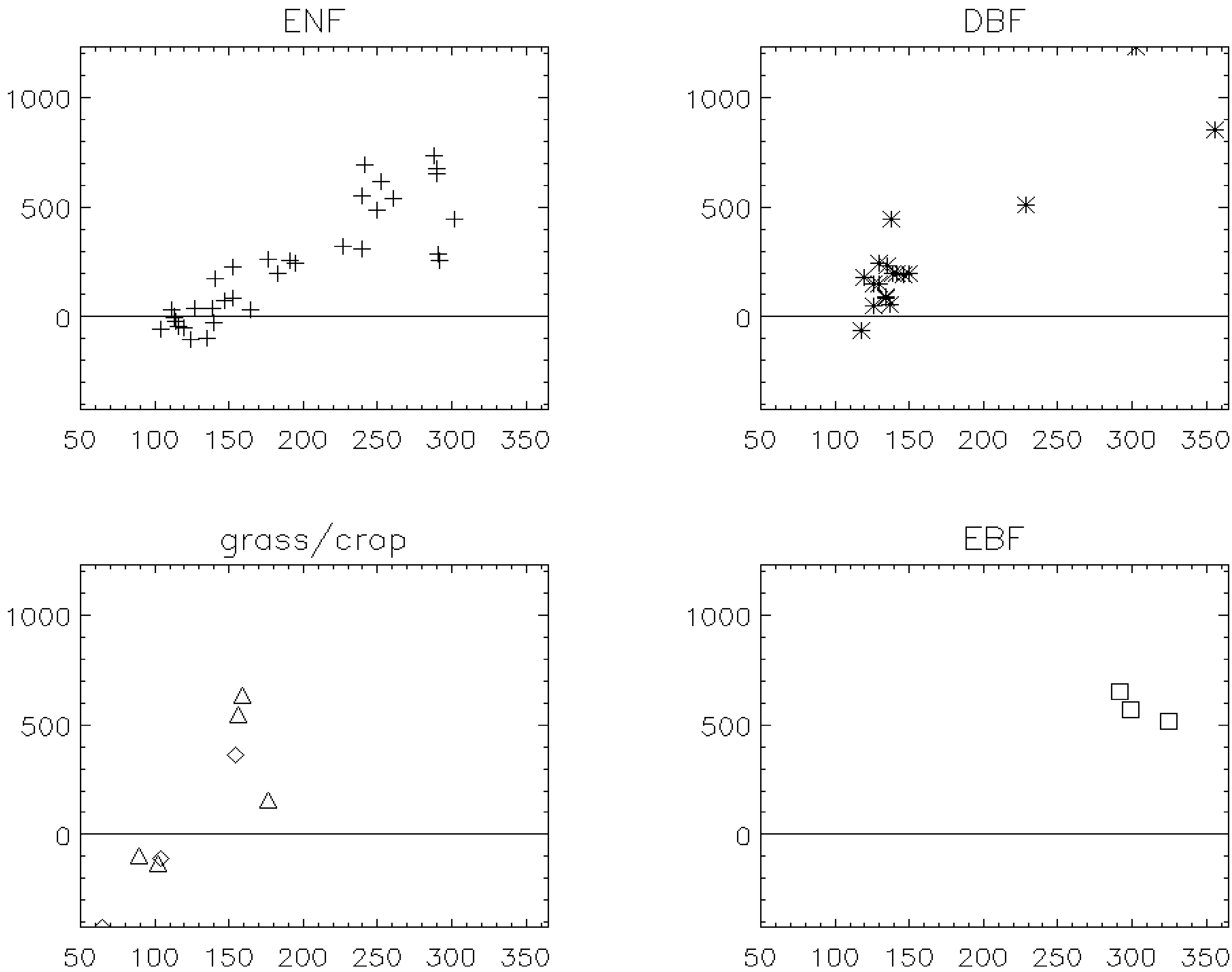
State space estimation can help with both *model error* and *aggregation error*

The state space approach can be generalized to work with fluxes estimated from atmospheric and satellite measurements

State space and parameter estimation techniques can be combined using ensemble data assimilation approaches

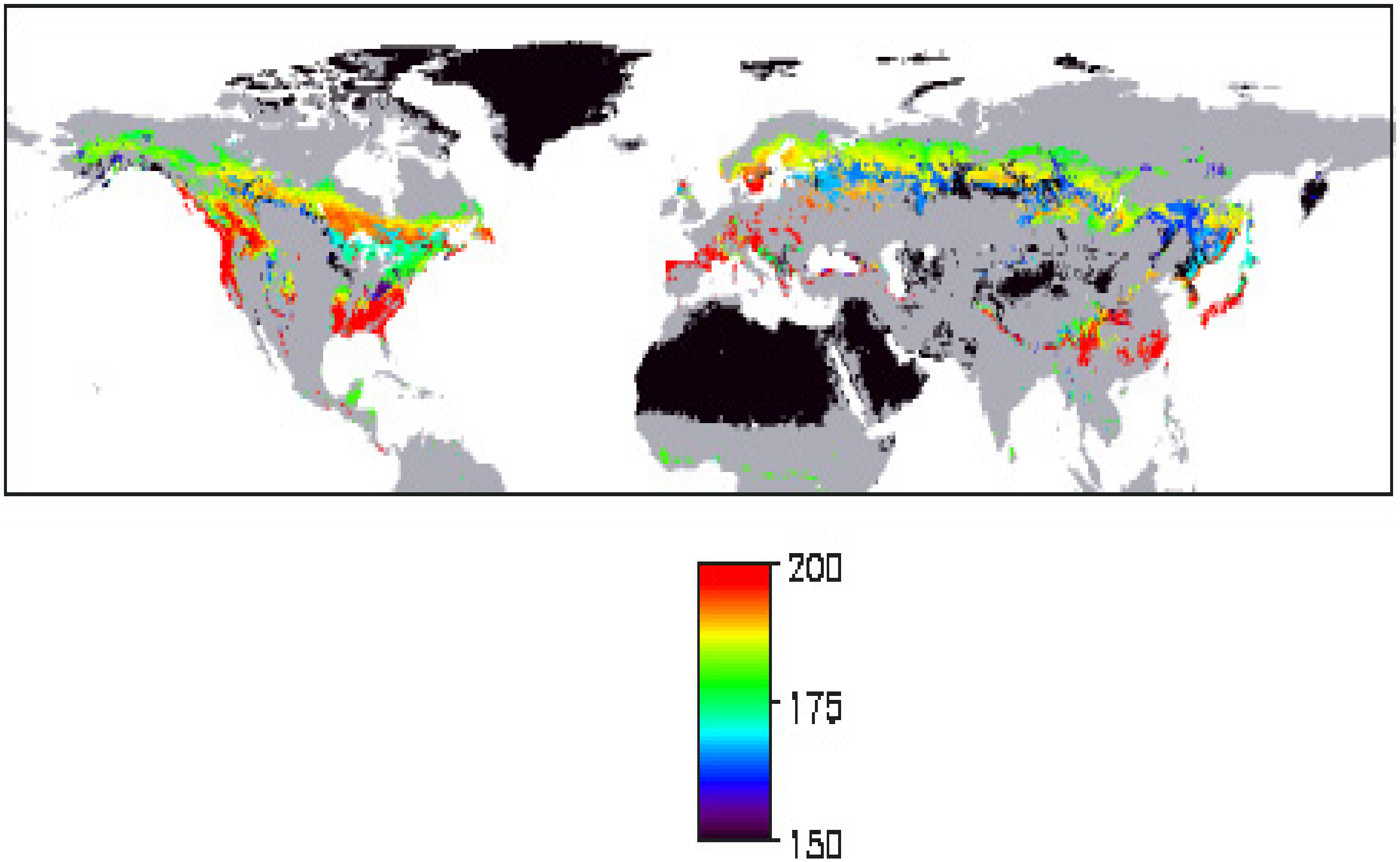


Annual NEE and GPP: also showing parts of “disturbance space” unsampled in this data set


ENF
DBF
grass/crop
EBF
1000
500
0
50
100
150
200
250
300
350

D.I.E.

Forest NEE extrapolated from CUP



Using: forest type map, separate regressions for broad and needle leafed forests and a satellite-based CUP, all aggregated to 0.5°.