CLM5 Matrix Model

Computational efficiency, diagnostics, and Improvement with Data

Chris Lu[1,2], Zhenggang Du[1,3], Feng Tao[4], David Lawrence[5], Erik Kluzek[5],
Keith Oleson[5], Nathan Collier[6], and Charlie Koven[7], Yiqi Luo[1,4],

[1] Northern Arizona University, USA
[2] Sun Yat-sen University, China
[3] East China Normal University, China
[4] Tsinghua University, China
[5] National Center for Atmosphere Research
[6] Oak Ridge National Laboratory
[7] Lawrence Berkeley National Laboratory

Yiqi.luo@nau.edu

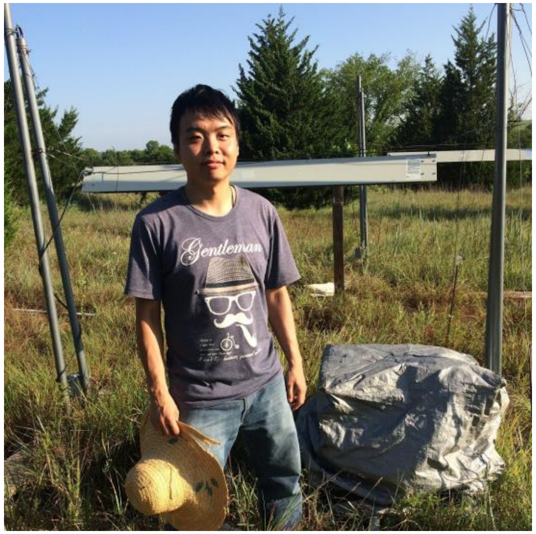
http://www2.nau.edu/luo-lab/?home

ECOS^S Center for Ecosystem Science and Society at Northern Arizona University

NCAR LMWG, Feb 11, 2019

ECOLAB OF DR. YIQI LUO

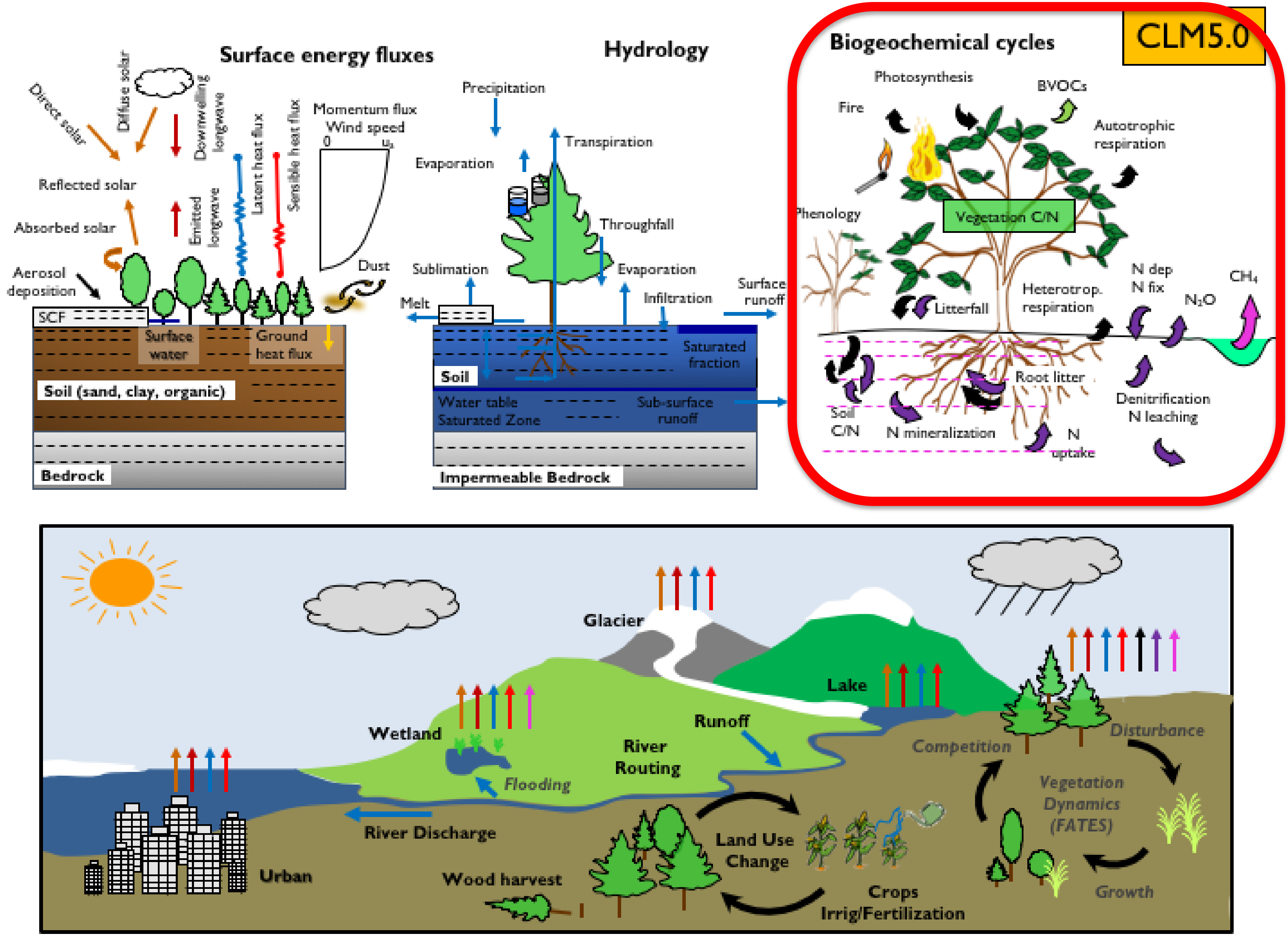
Heroes



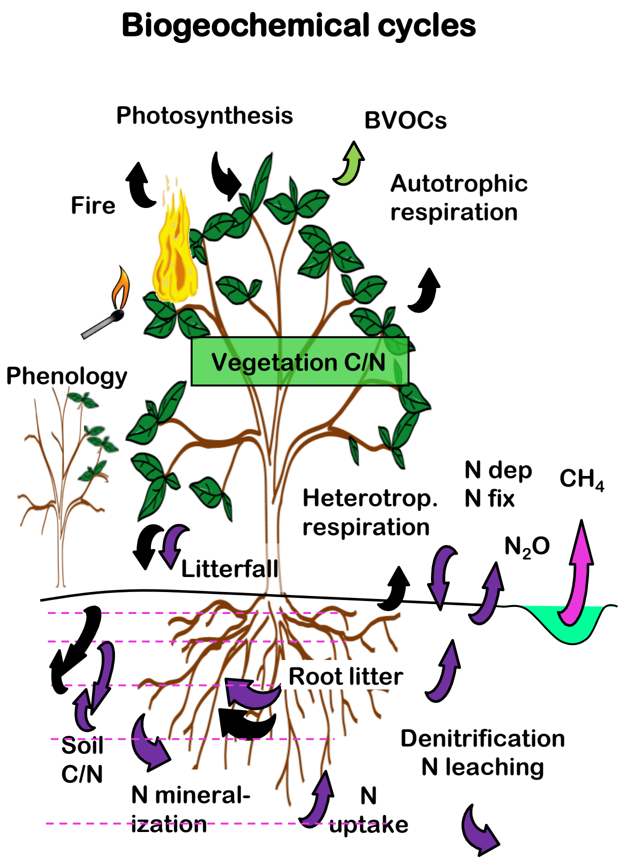
Chris Lu

Zhenggang Du

Feng Tao


Surface energy fluxes
Direct solar
Diffuse solar
Downwelling longwave
Latent heat flux
Sensible heat flux
Momentum flux
Wind speed
0
u_a
Reflected solar
Absorbed solar
Emitted longwave
Aerosol deposition
Dust
SCF
Surface water
Ground heat flux
Soil (sand, clay, organic)
Bedrock
Hydrology
Precipitation
Transpiration
Evaporation
Throughfall
Evaporation
Sublimation
Melt
Infiltration
Surface runoff
Saturated fraction
Soil
Water table
Saturated Zone
Sub-surface runoff
Impermeable Bedrock
Biogeochemical cycles
CLM5.0
Photosynthesis
BVOCs
Fire
Autotrophic respiration
Phenology
Vegetation C/N
Litterfall
Heterotrop. respiration
N dep
N fix
N_2O
CH_4
Root litter
Soil C/N
N mineralization
N uptake
Denitrification
N leaching
Glacier
Lake
Wetland
Flooding
River Routing
Runoff
River Discharge
Urban
Wood harvest
Land Use Change
Crops
Irrig/Fertilization
Competition
Disturbance
Vegetation Dynamics (FATES)
Growth

CLM5.0 matrix model



Carbon and nitrogen transfer among >892 pools

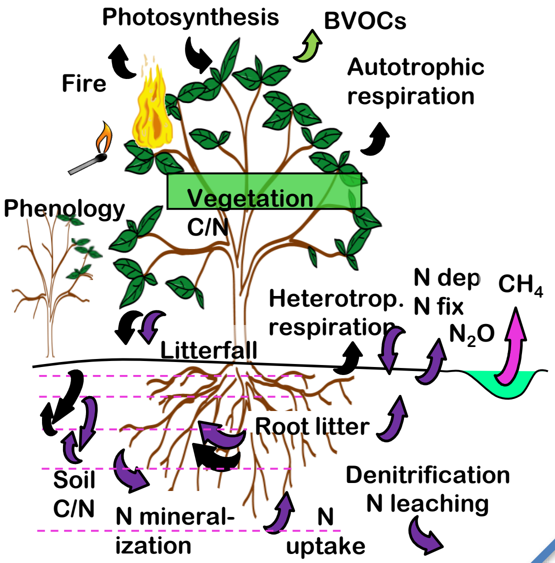
- 446 pools for carbon cycle
  - 18 x 17 = 306 plant pools
  - 140 soil and litter pools over 20 layers
- 446 pools for organic nitrogen
  - plus Inorganic nitrogen pools

$$\left\{\begin{aligned}
\frac{d}{dt}X(t) &= (\overset{\text{C transfer of phenology}}{A_{ph}(t)}\underset{\text{C turnover of phenology}}{K_{ph}(t)} + \overset{\text{C transfer of gap mortality}}{A_{gm}(t)}\underset{\text{C turnover of gap mortality}}{K_{gm}(t)} + \overset{\text{C transfer of fire}}{A_{fi}(t)}\underset{\text{C turnover of fire}}{K_{fi}(t)})\overset{\text{pool state}}{X(t)} + \underset{\text{allocation}}{B(t)}\overset{\text{input}}{F(t)} \\
\frac{d}{dt}N(t) &= \left(A_{Nph}(t)K_{Nph}(t) + A_{Ngm}(t)K_{Ngm}(t) + A_{Nfi}(t)K_{Nfi}(t)\right)N(t) + B_N(t)F_N(t)
\end{aligned}\right.$$

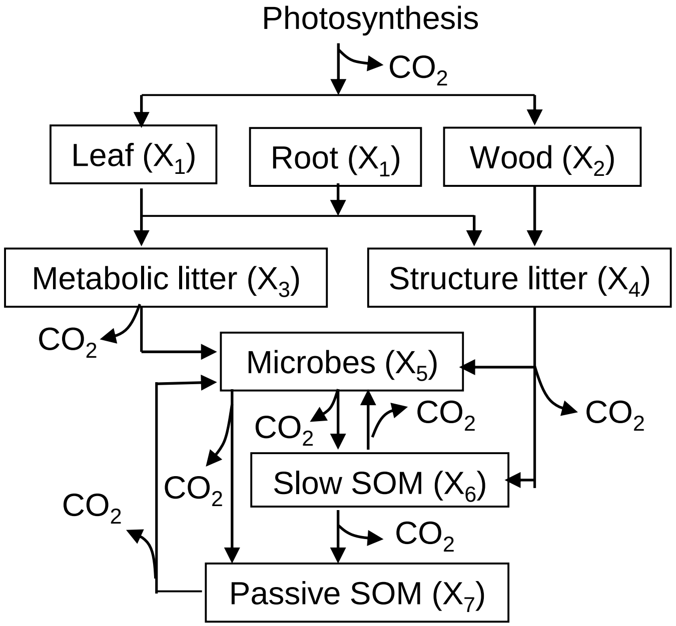
Simplicity in coding
Diagnostic capability
Computational efficiency

Model
Understanding
Evaluation
Improvement

Matrix Approach



A: Basic process


Photosynthesis

CO$_2$

Leaf (X$_1$)

Root (X$_1$)

Wood (X$_2$)

Metabolic litter (X$_3$)

Structure litter (X$_4$)

CO$_2$

Microbes (X$_5$)

CO$_2$

CO$_2$

CO$_2$

CO$_2$

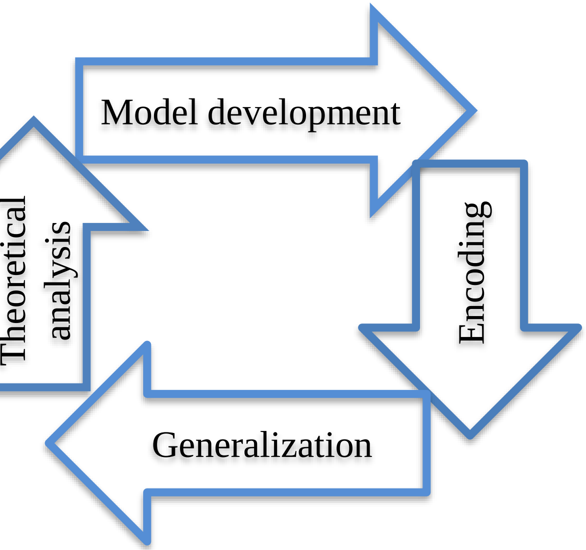
Slow SOM (X$_6$)

CO$_2$

CO$_2$

Passive SOM (X$_7$)

B: Shared model structure


Model development

Encoding

Generalization

Theoretical analysis

C: Similar algorithm

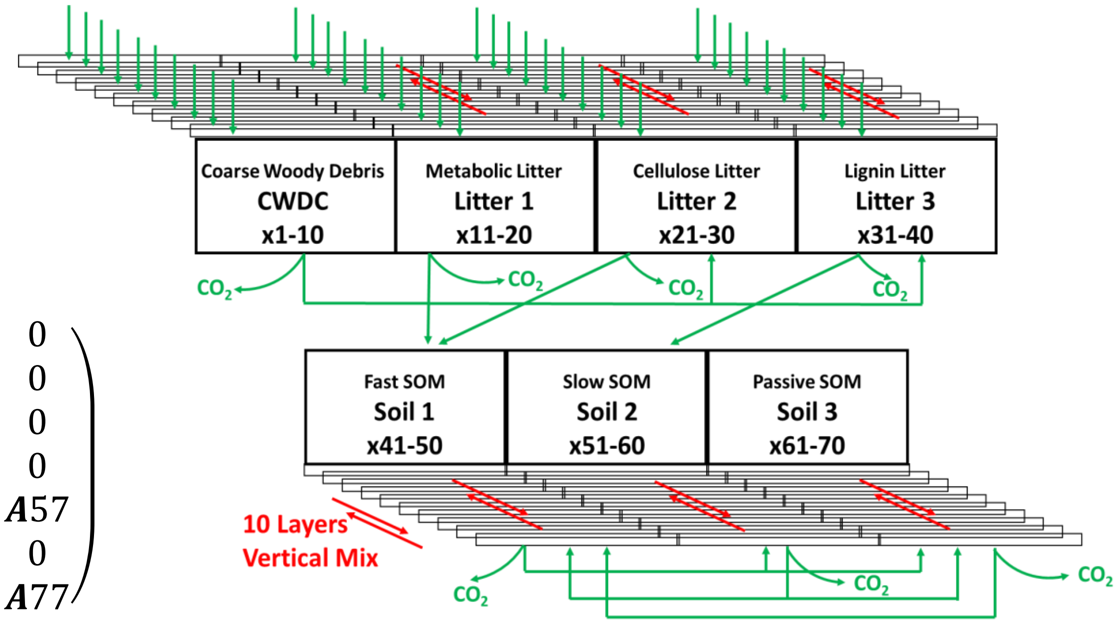
$$\left.\begin{array}{l}
Plant\begin{cases} dX_1(t)/dt = b_1U(t) - \xi c_1X_1(t) \\ dX_2(t)/dt = b_2U(t) - \xi c_2X_2(t) \\ dX_3(t)/dt = b_3U(t) - \xi c_3X_3(t) \end{cases} \\
Litter\begin{cases} dX_4(t)/dt = \xi[c_1a_{41}x_1(t) + c_3a_{43}x_3(t) - c_4X_4(t)] \\ dX_5(t)/dt = \xi[c_1a_{51}x_1(t) + c_2x_2(t) + c_3a_{53}x_3(t) - c_5X_5(t)] \end{cases} \\
SOM\begin{cases} dX_6(t)/dt = \xi[c_4a_{64}x_4(t) + c_5a_{65}x_5(t) + c_7a_{67}x_7(t) + c_8a_{68}x_8(t) - c_6X_6(t)] \\ dX_7(t)/dt = \xi[c_5a_{75}x_5(t) + c_6a_{76}x_6(t) - c_7X_7(t)] \\ dX_8(t)/dt = \xi[c_6a_{86}x_6(t) + c_7a_{87}x_7(t) - c_8X_8(t)] \end{cases}
\end{array}\right\}$$

D: General model

$$\begin{cases} \dfrac{dX(t)}{dt} = \xi ACX(t) + BU(t) \\ X(t=0) = X_0 \end{cases}$$

Luo et al. 2003 GBC
Luo and Weng 2011 TREE
Luo et al. 2012
Luo et al. 2015
Luo et al. 2017

Matrix equation of CLM4.5



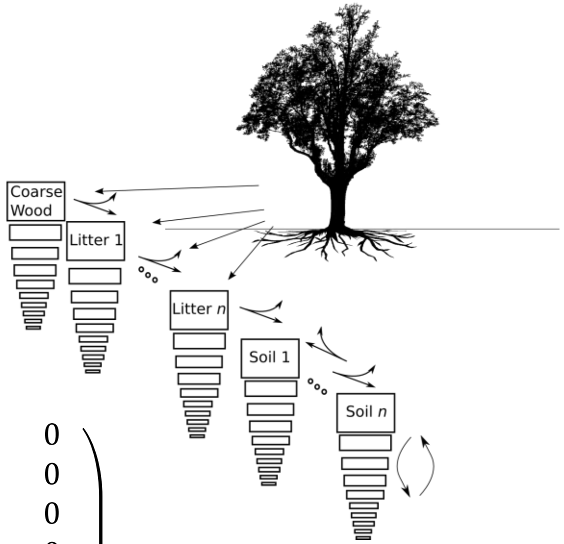
$$\frac{d\boldsymbol{X}(\boldsymbol{t})}{d\boldsymbol{t}} = \boldsymbol{B}(\boldsymbol{t})\boldsymbol{I}(\boldsymbol{t}) - \boldsymbol{A}\boldsymbol{\xi}(\boldsymbol{t})\boldsymbol{K}\boldsymbol{X}(\boldsymbol{t}) - \boldsymbol{V}(\boldsymbol{t})\boldsymbol{X}(\boldsymbol{t})$$

$$X(t) = (X_1(t), X_2(t), X_3(t), \dots, X_{70}(t))^T$$

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{A}11 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{A}22 & 0 & 0 & 0 & 0 & 0 \\ \boldsymbol{A}31 & 0 & \boldsymbol{A}33 & 0 & 0 & 0 & 0 \\ \boldsymbol{A}41 & 0 & 0 & \boldsymbol{A}44 & 0 & 0 & 0 \\ 0 & \boldsymbol{A}52 & \boldsymbol{A}53 & 0 & \boldsymbol{A}55 & \boldsymbol{A}56 & \boldsymbol{A}57 \\ 0 & 0 & 0 & \boldsymbol{A}64 & \boldsymbol{A}65 & \boldsymbol{A}66 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{A}75 & \boldsymbol{A}76 & \boldsymbol{A}77 \end{pmatrix}$$

$$\mathbf{A}_{31} = diag(-f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31}, -f_{31})$$

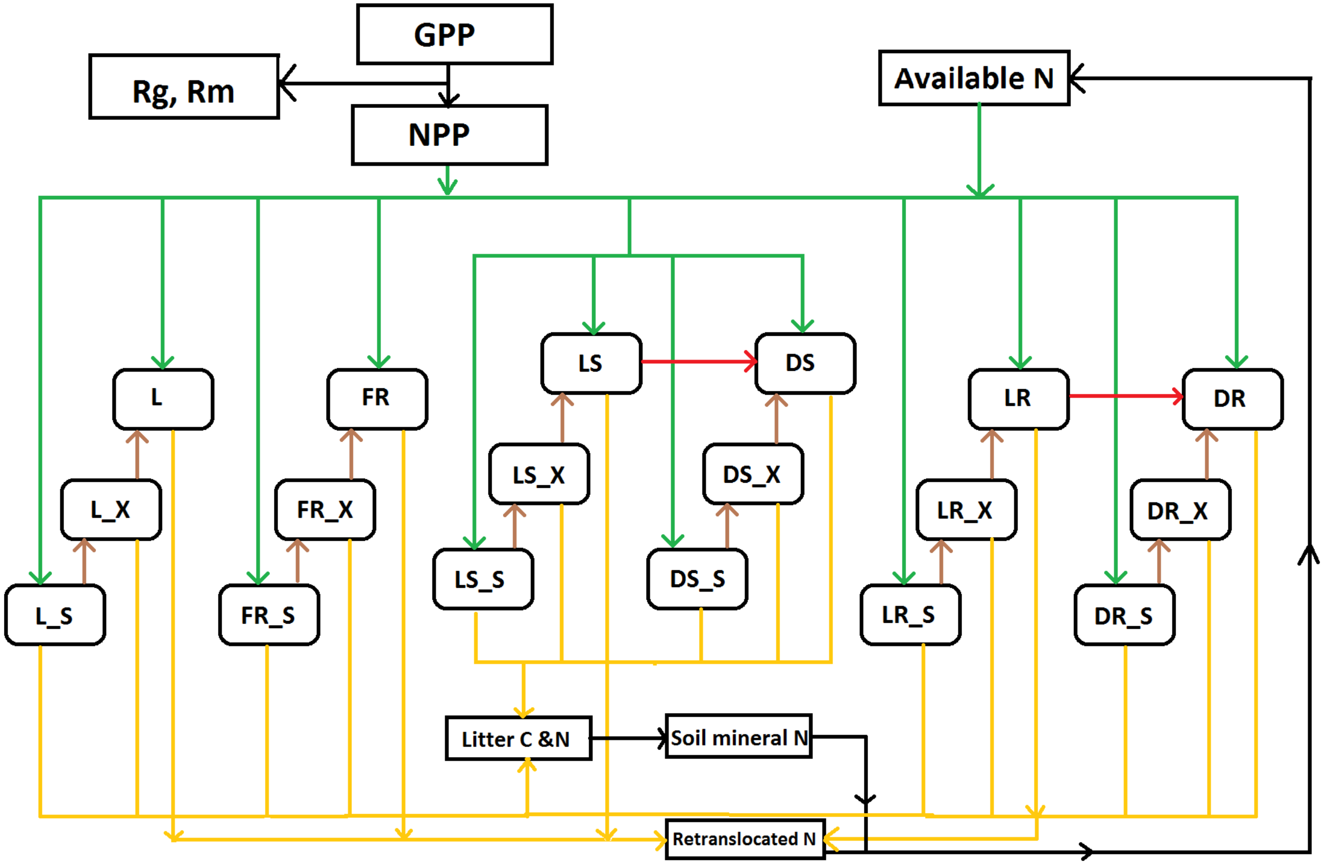
$$\mathbf{V}(\mathbf{t}) = \begin{pmatrix} \boldsymbol{V}11 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{V}22(\boldsymbol{t}) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{V}33(\boldsymbol{t}) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{V}44(\boldsymbol{t}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{V}55(\boldsymbol{t}) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{V}66(\boldsymbol{t}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boldsymbol{V}77(\boldsymbol{t}) \end{pmatrix}$$



$$\boldsymbol{V}22 = \boldsymbol{diag}(\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{10})^{-1} \begin{pmatrix} \boldsymbol{g}_1 & -\boldsymbol{g}_1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\boldsymbol{h}_2 & \boldsymbol{h}_2 + \boldsymbol{g}_2 & -\boldsymbol{g}_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\boldsymbol{h}_3 & \boldsymbol{h}_3 + \boldsymbol{g}_3 & -\boldsymbol{g}_3 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\boldsymbol{h}_4 & \boldsymbol{h}_4 + \boldsymbol{g}_4 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \boldsymbol{h}_8 + \boldsymbol{g}_8 & -\boldsymbol{g}_8 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -\boldsymbol{h}_9 & \boldsymbol{h}_9 + \boldsymbol{g}_9 & -\boldsymbol{g}_9 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\boldsymbol{h}_{10} & \boldsymbol{h}_{10} \end{pmatrix}$$

Huang et al. 2018
Global Change Biology

CLM vegetation C&N: phenology, fire etc.



Controling Procedure

🟩 Allocation, Phenology

🟨 Phenology offset, Background turnover, Gap mortality, Fire

🟥 Phenology, Fire

🟫 Phenology

L: leaf;
FR: fine root;
LS: live stem;
DS: dead stem;
LR: live coarse root;
DR: dead coarse root;

L_X: leaf transfer;
FR_X: fine root transfer;
LS_X: live stem transfer;
DS_X: dead stem transfer;
LR_X: live coarse root transfer;
DR_X: dead coarse root transfer;

L_S: leaf storage
FR_S: fine root storage
LS_S: live stem storage
DS_S: dead stem storage
LR_S: live coarse root storage
DR_X: dead coarse root storage

Matrix equation of vegetation carbon dynamics

$$\frac{d}{dt}X(t) = \left(\overset{\text{C transfer of phenology}}{A_{ph}}(t)\underset{\text{C turnover of phenology}}{K_{ph}}(t) + \overset{\text{C transfer of gap mortality}}{A_{gm}}(t)\underset{\text{C turnover of gap mortality}}{K_{gm}}(t) + \overset{\text{C transfer of fire}}{A_{fi}}(t)\underset{\text{C turnover of fire}}{K_{fi}}(t)\right)\overset{\text{pool state}}{X}(t) + \underset{\text{allocation}}{B}(\mathrm{t})\overset{\text{input}}{F}(t)$$

Vegetation nitrogen dynamics

$$\frac{d}{dt}N(t) = \left(A_{Nph}(t)K_{Nph}(t) + A_{Ngm}(t)K_{Ngm}(t) + A_{Nfi}(t)K_{Nfi}(t)\right)N(t) + B_{N}(\mathrm{t})F_{N}(t)$$

Mathematical framework

Vertical profile

$$\frac{d\boldsymbol{X}(\boldsymbol{t})}{d\boldsymbol{t}} = (\boldsymbol{A}\xi(\boldsymbol{t})\boldsymbol{K} + \boldsymbol{V}(\boldsymbol{t}))\boldsymbol{X}(\boldsymbol{t}) + \boldsymbol{B}(\boldsymbol{t})u(\boldsymbol{t})$$

$$\frac{dX(t)}{dt} = A\xi(t)KX(t) + Bu(t)$$

$$\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{X}(\boldsymbol{t}) = (\overset{\text{C transfer of phenology}}{\boldsymbol{A}_{\boldsymbol{ph}}}(\boldsymbol{t})\underset{\text{C turnover of phenology}}{\boldsymbol{K}_{\boldsymbol{ph}}}(\boldsymbol{t}) + \overset{\text{C transfer of gap mortality}}{\boldsymbol{A}_{\boldsymbol{gm}}}(\boldsymbol{t})\underset{\text{C turnover of gap mortality}}{\boldsymbol{K}_{\boldsymbol{gm}}}(\boldsymbol{t}) + \overset{\text{C transfer of fire}}{\boldsymbol{A}_{\boldsymbol{fi}}}(\boldsymbol{t})\underset{\text{C turnover of fire}}{\boldsymbol{K}_{\boldsymbol{fi}}}(\boldsymbol{t}))\overset{\text{pool state}}{\boldsymbol{X}(\boldsymbol{t})} + \underset{\text{allocation}}{\boldsymbol{B}}(\mathbf{t})\boldsymbol{F}\overset{\text{input}}{(\boldsymbol{t})}$$

Vegetation dynamics

Developing models at different levels of complexity under one overarching theory

Unifying land carbon cycle models

Matrix models

1. CLM 3.5
2. CLM4.0
3. CLM4.5
4. CLM5.0
5. CABLE
6. LPJ-GUESS
7. ORCHIDEE
8. BEPS
9. TECO
10. JULES
11. IBIS

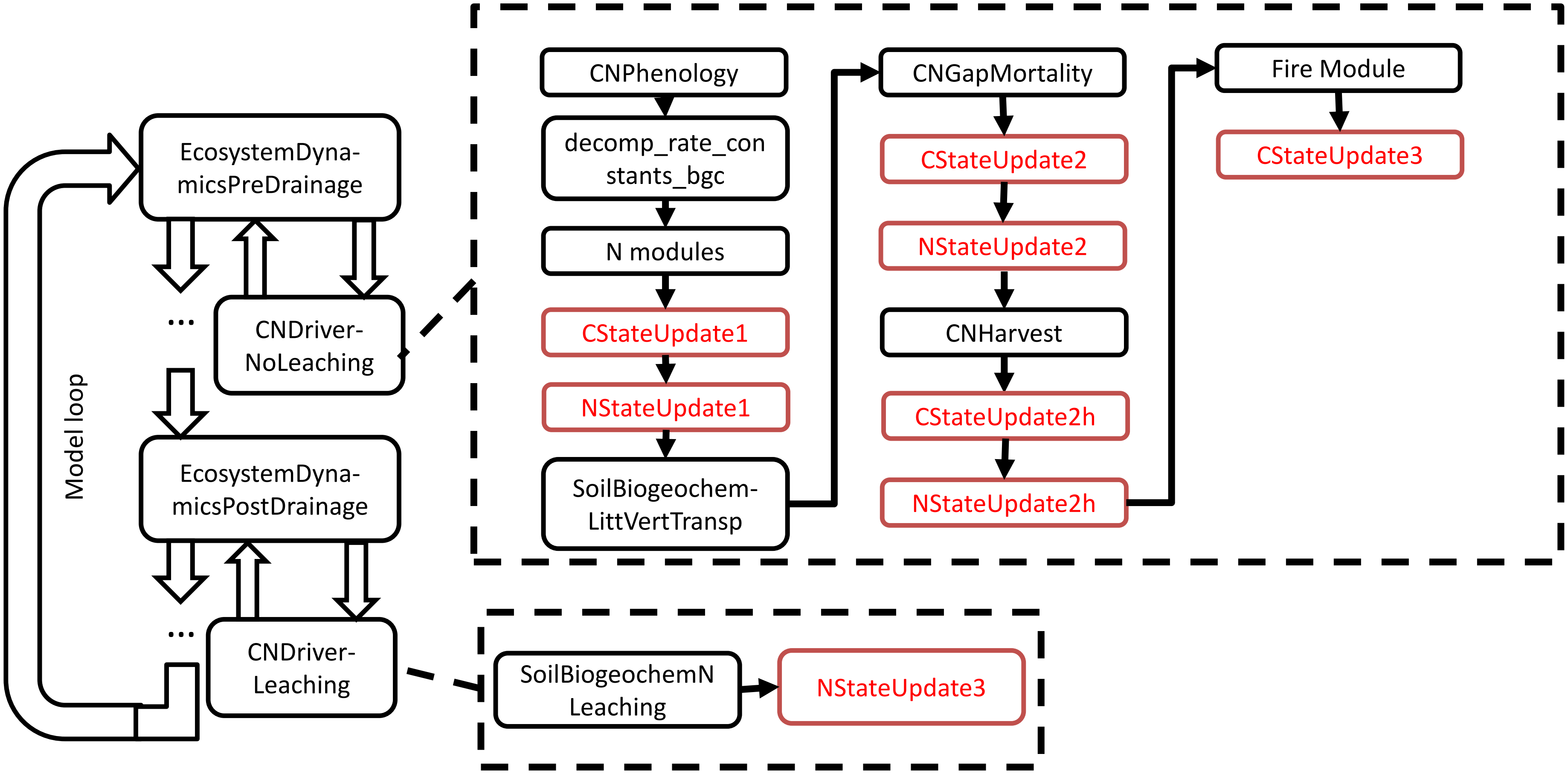
In progress

1. LM3V-N

10 nonlinear microbial models by Carlos Sierra

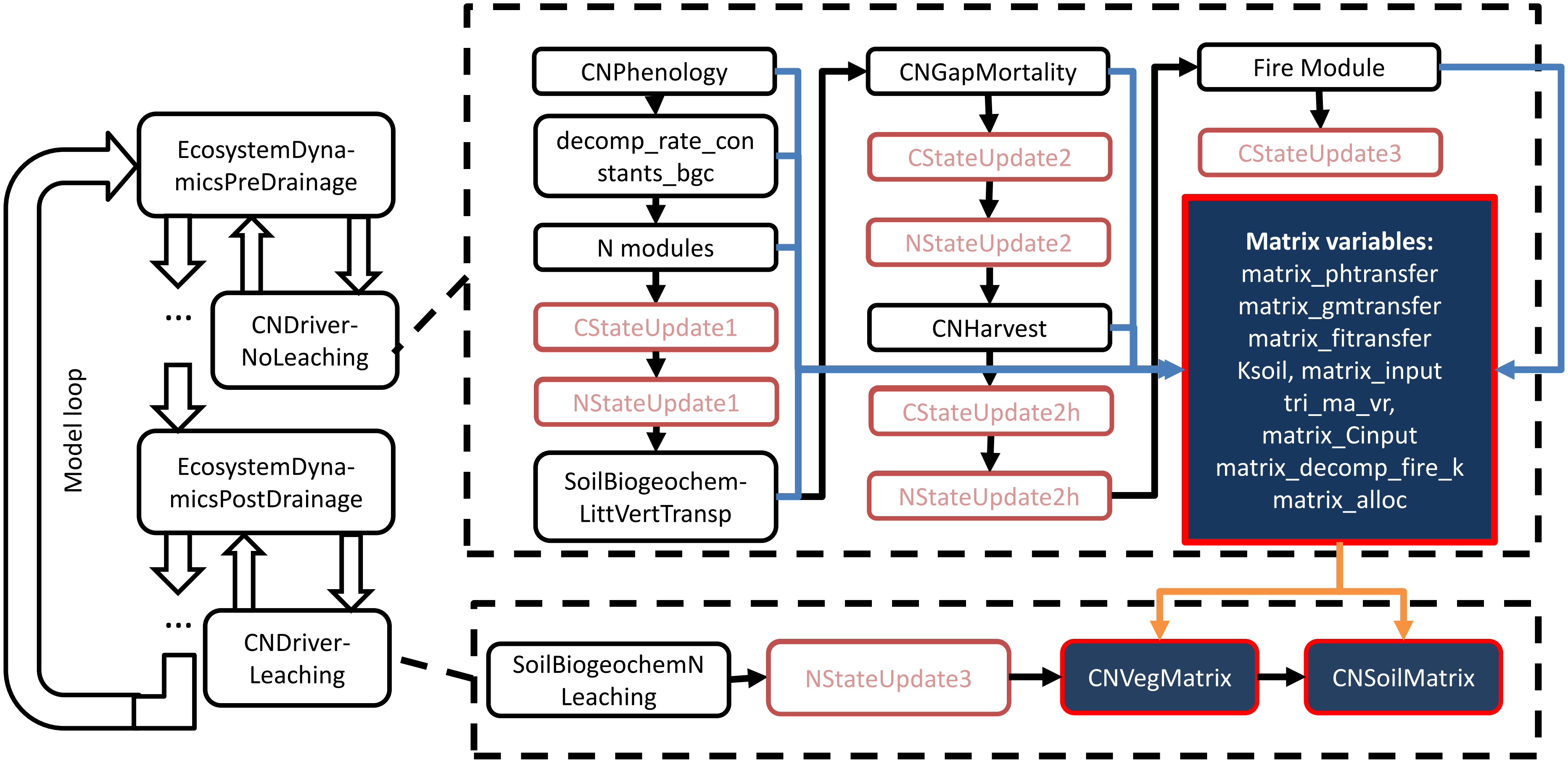
Luo et al. 2017
Luo et al. *to be submitted*

Programming



CLM5.0 biogeochem cycle

Programming



CLM5.0 biogeochem cycle with matrix

Technique test

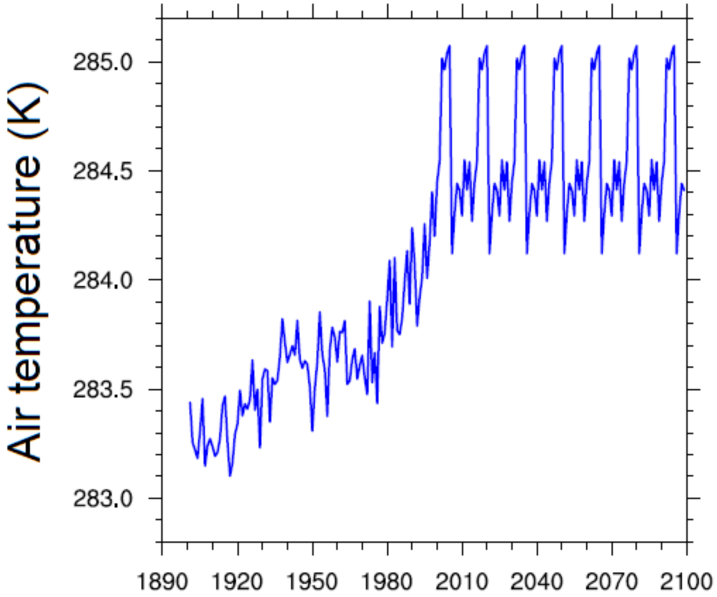
- Use automated tools (CESM test suites, use “aux_clm” testlists) to test CLM5.0 code both with matrix on and matrix off.

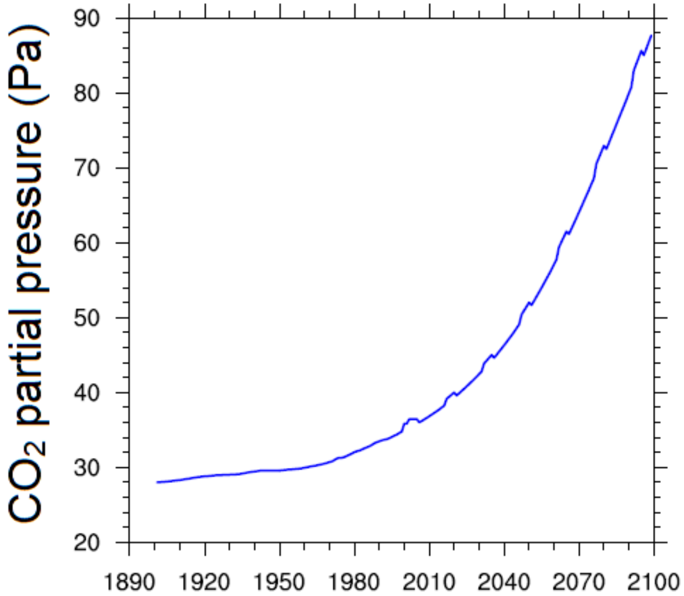
- Code with matrix on:
- In total 159 tests, 149 tests are passed, 2 tests are expected fails, 6 tests are due to restart inconsistency, 1 is from C balance issue, 1 is from wrong initial file address pointed to old address.

- Code with matrix off:
- In total 159 tests, 156 tests are passed, 2 tests are expected fails, 1 is from wrong initial file address pointed to old address..

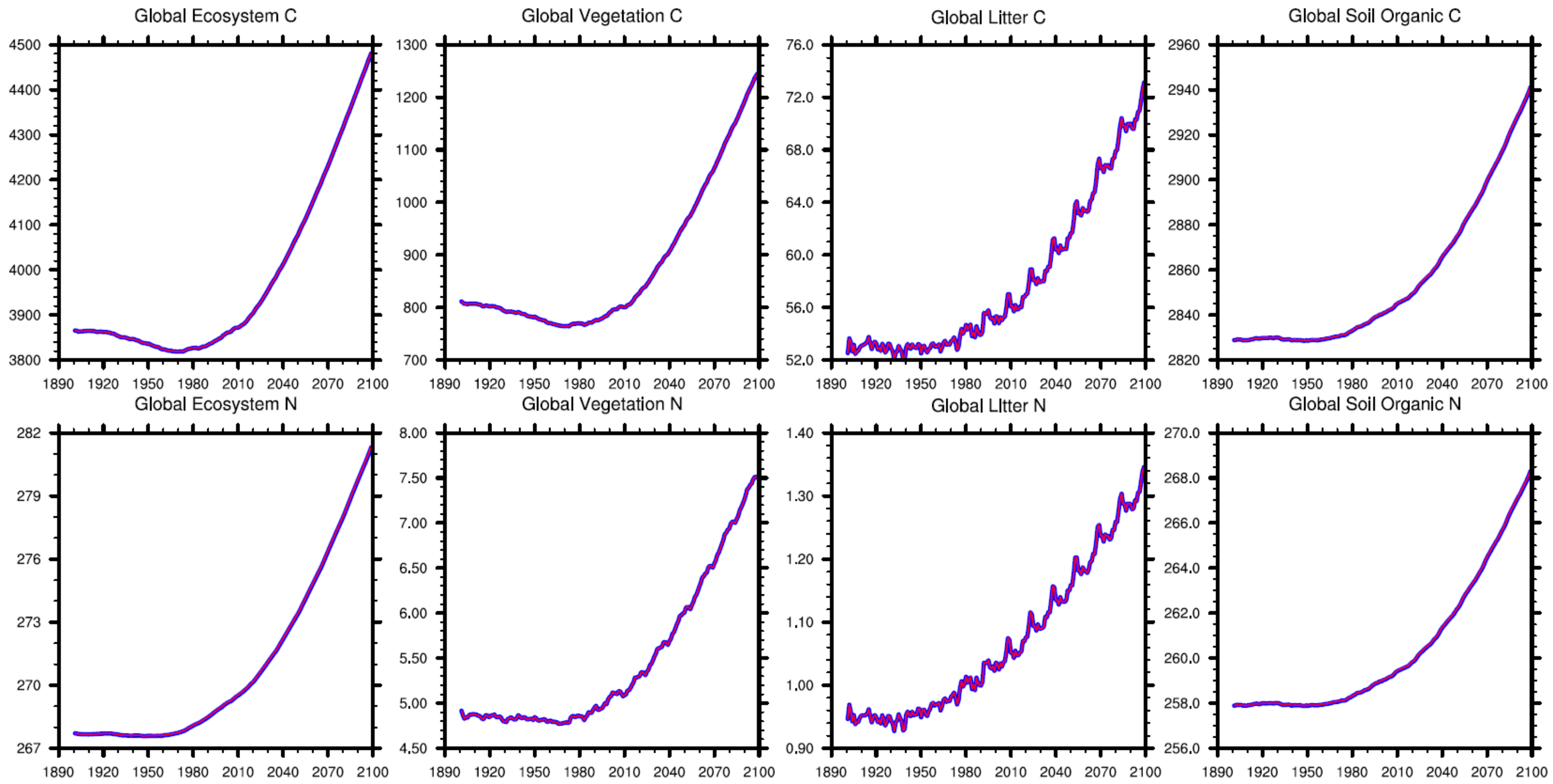
Scientific test: Simulations

Time	Resolution	Compset	Climate	[CO_2]
1901-2099	4x5	IHistClm50Bgc	GSWP3	CMIP5

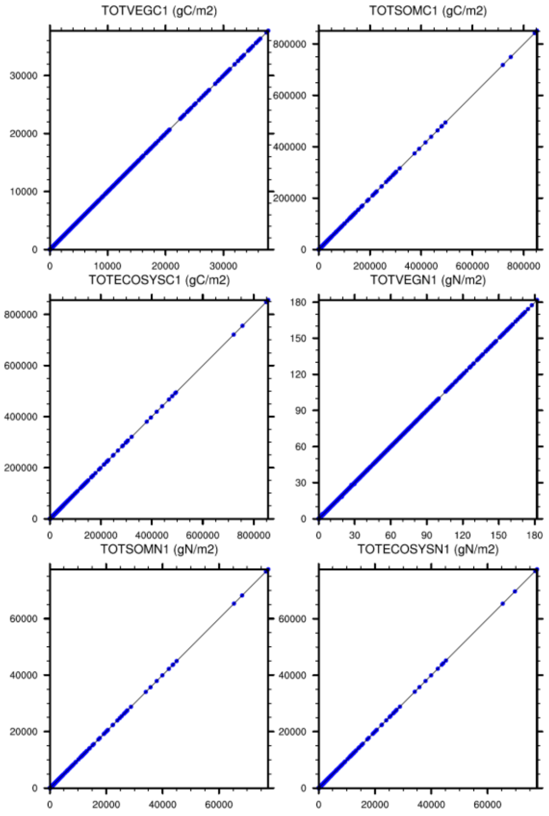




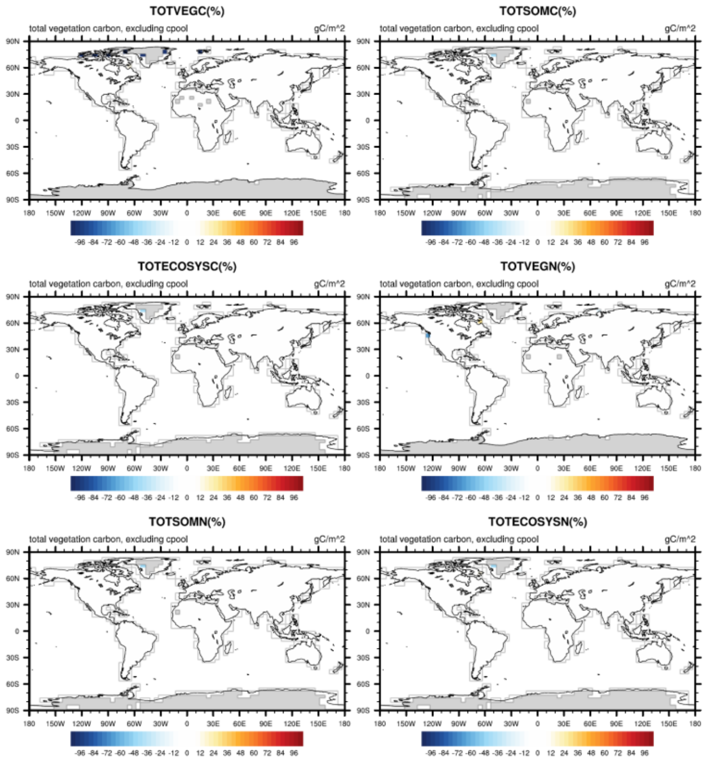
Matrix Vs default RCP 8.5 (global-level)



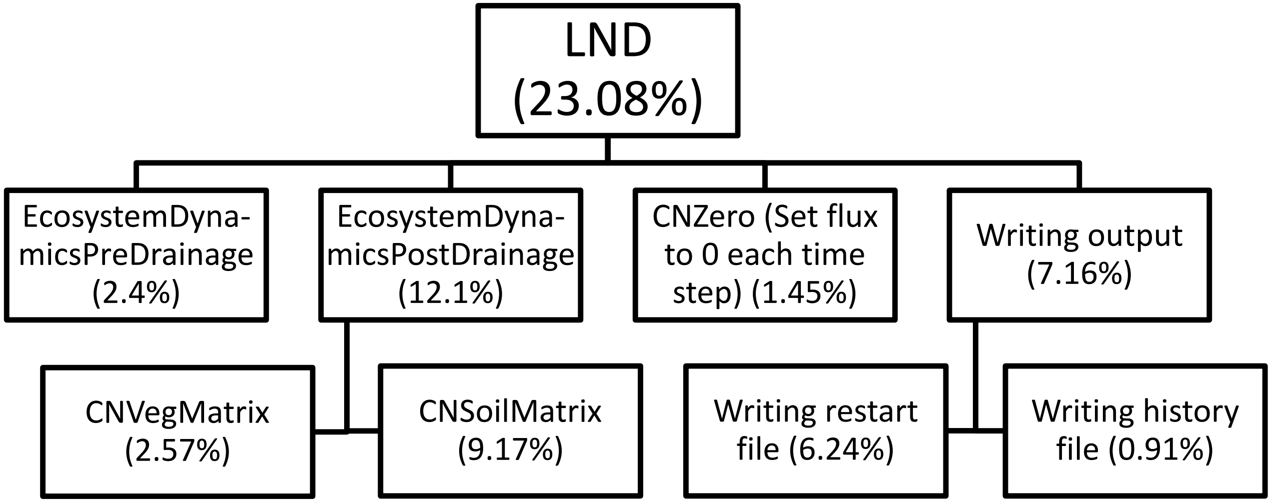
Matrix Vs default (Global)



(Matrix – Default)/Default * 100%



Computational efficiency for forward modeling



LND (23.08%)

EcosystemDynamicsPreDrainage (2.4%)

EcosystemDynamicsPostDrainage (12.1%)

CNZero (Set flux to 0 each time step) (1.45%)

Writing output (7.16%)

CNVegMatrix (2.57%)

CNSoilMatrix (9.17%)

Writing restart file (6.24%)

Writing history file (0.91%)

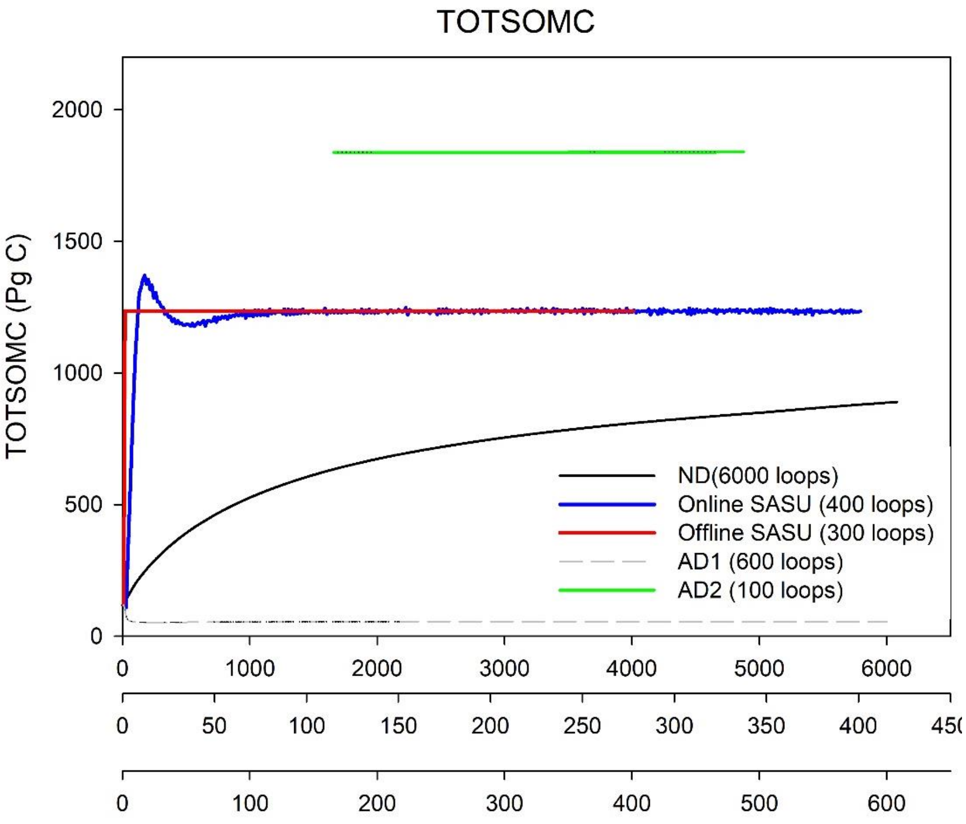
- The fraction in the bracket is a relative increase of computational time:

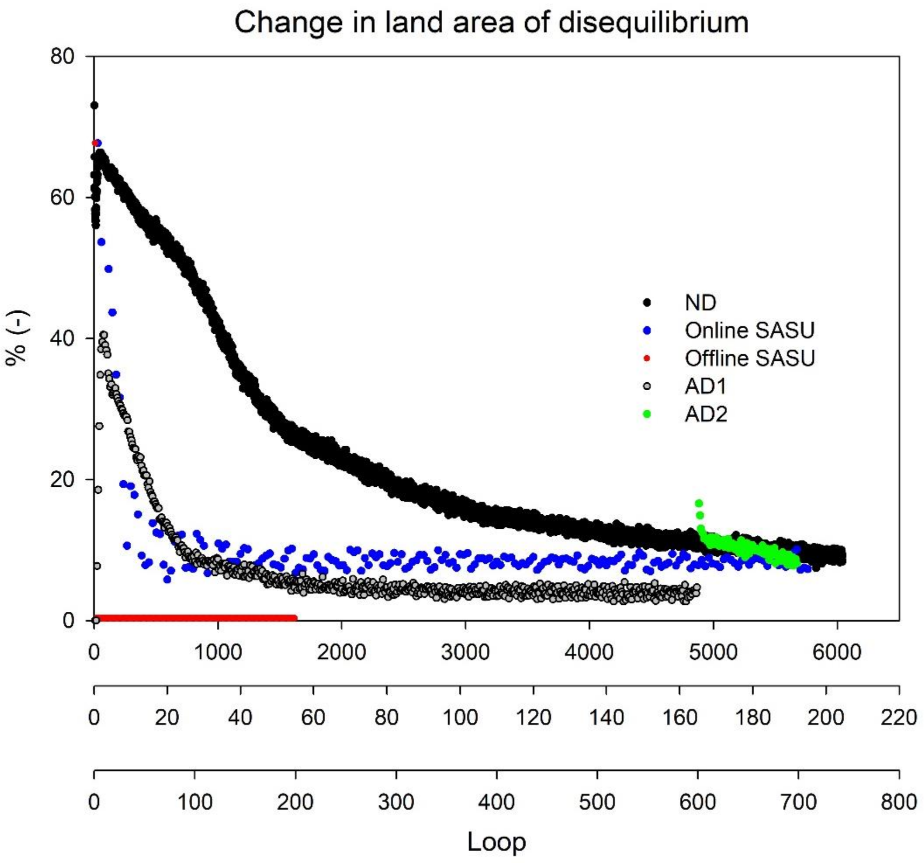
$$(T_{i,\ matrixon} - T_{i,\ matrixoff}) / T_{total,\ matrixoff} * 100\%$$

$T_{i,\ matrixon}$ is the computational time consumed by subroutine i, in matrix simulation. $T_{i,\ matrixoff}$ is the computational time consumed by subroutine i, in control simulation, $T_{total,\ matrixoff}$ is total computational time in control simulation.

- Comparison is between one-month global simulation (2^ox2^o) with matrix on and off.

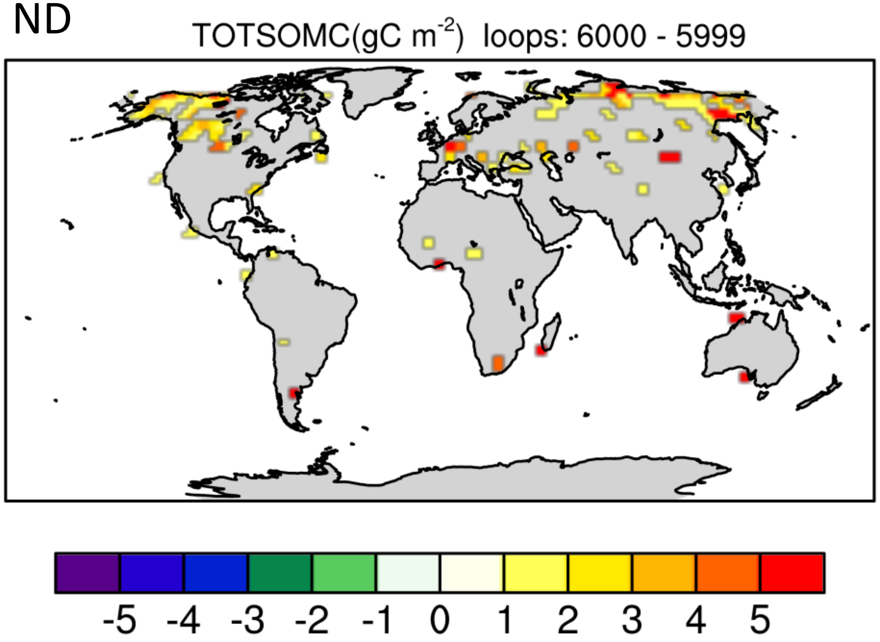
Computational efficiency for Spin-up



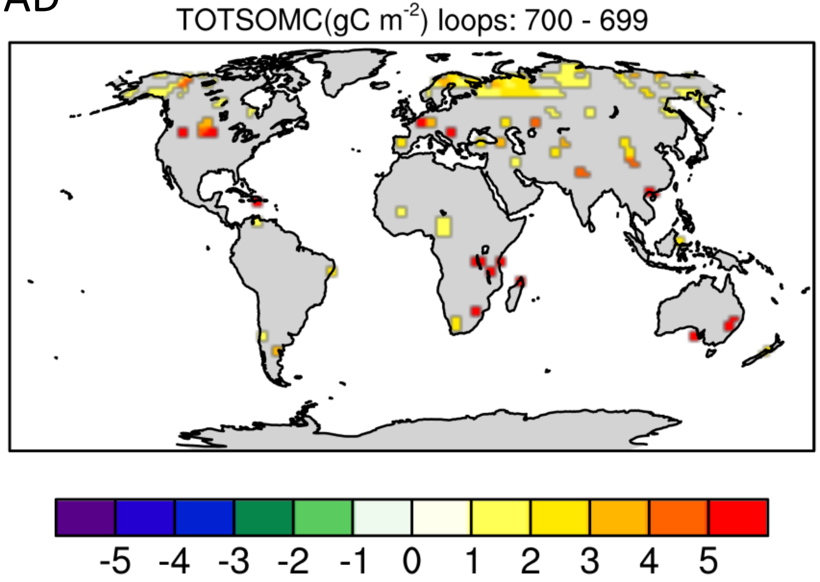


Forcing data: One-year (1911) GSWP3V

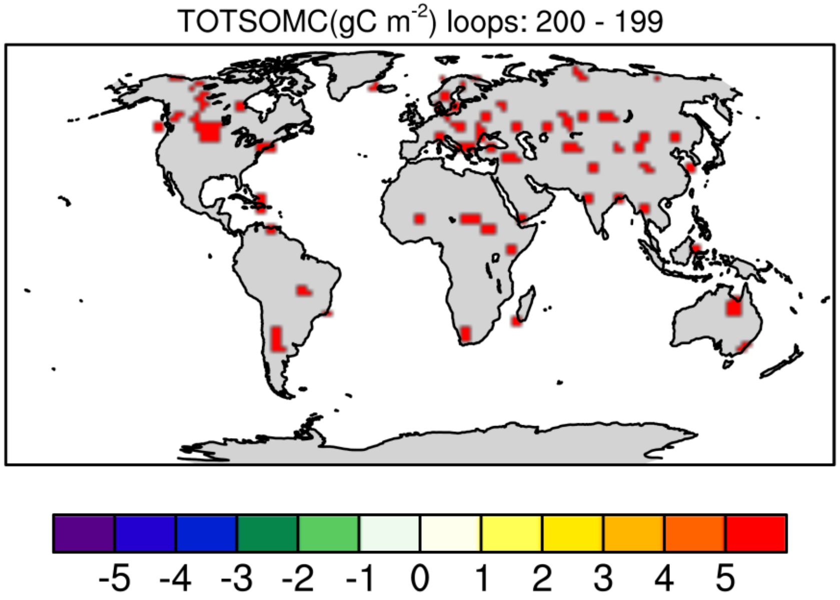
ND


TOTSOMC(gC m^{-2}) loops: 6000 - 5999

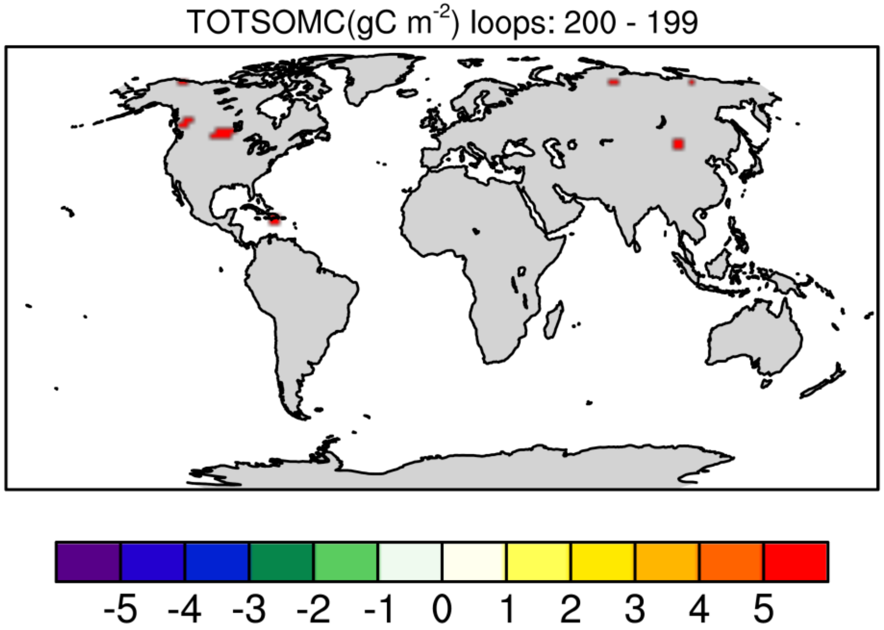
AD


TOTSOMC(gC m^{-2}) loops: 700 - 699

Online SASU

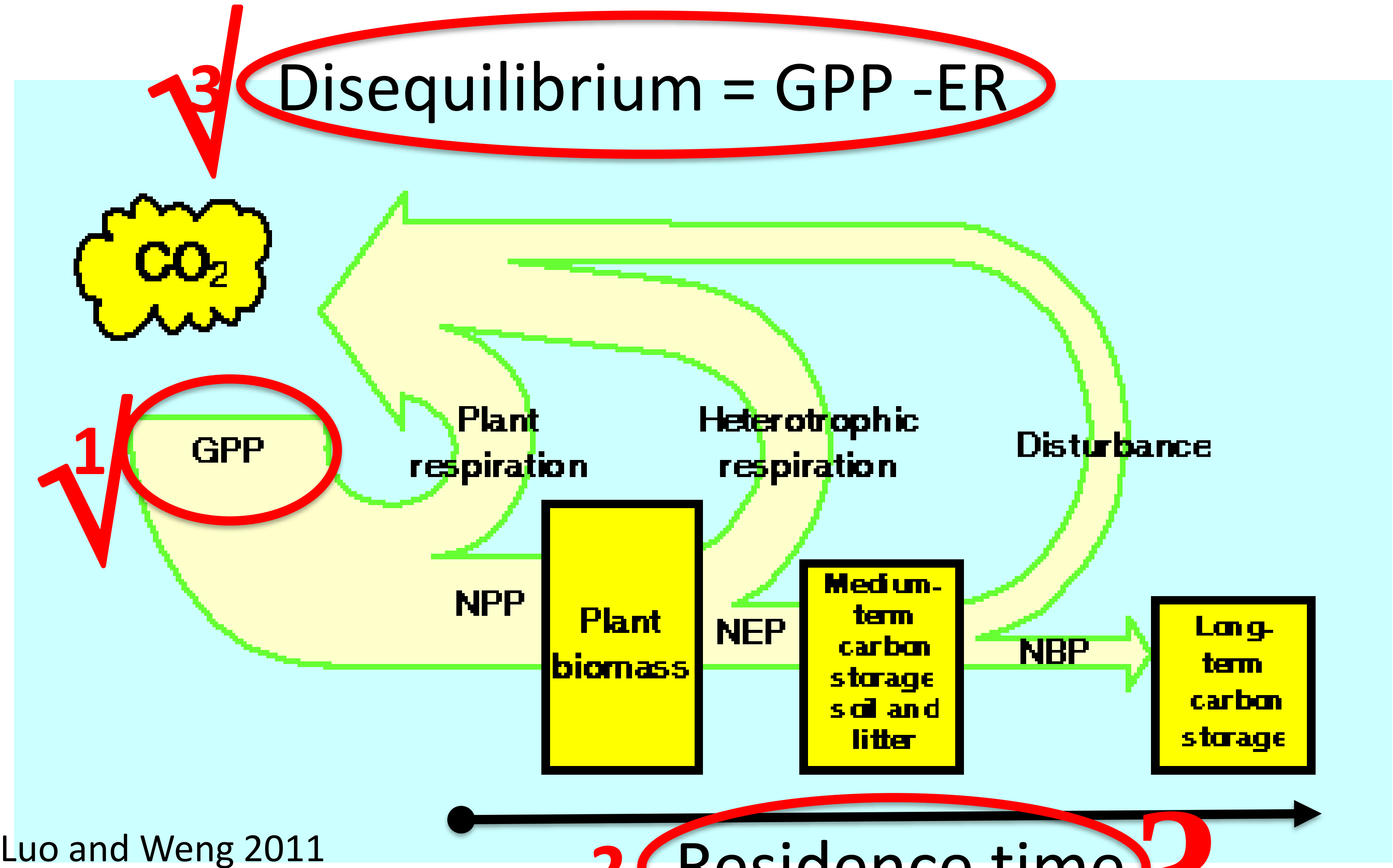

TOTSOMC(gC m^{-2}) loops: 200 - 199

Offline SASU


TOTSOMC(gC m^{-2}) loops: 200 - 199

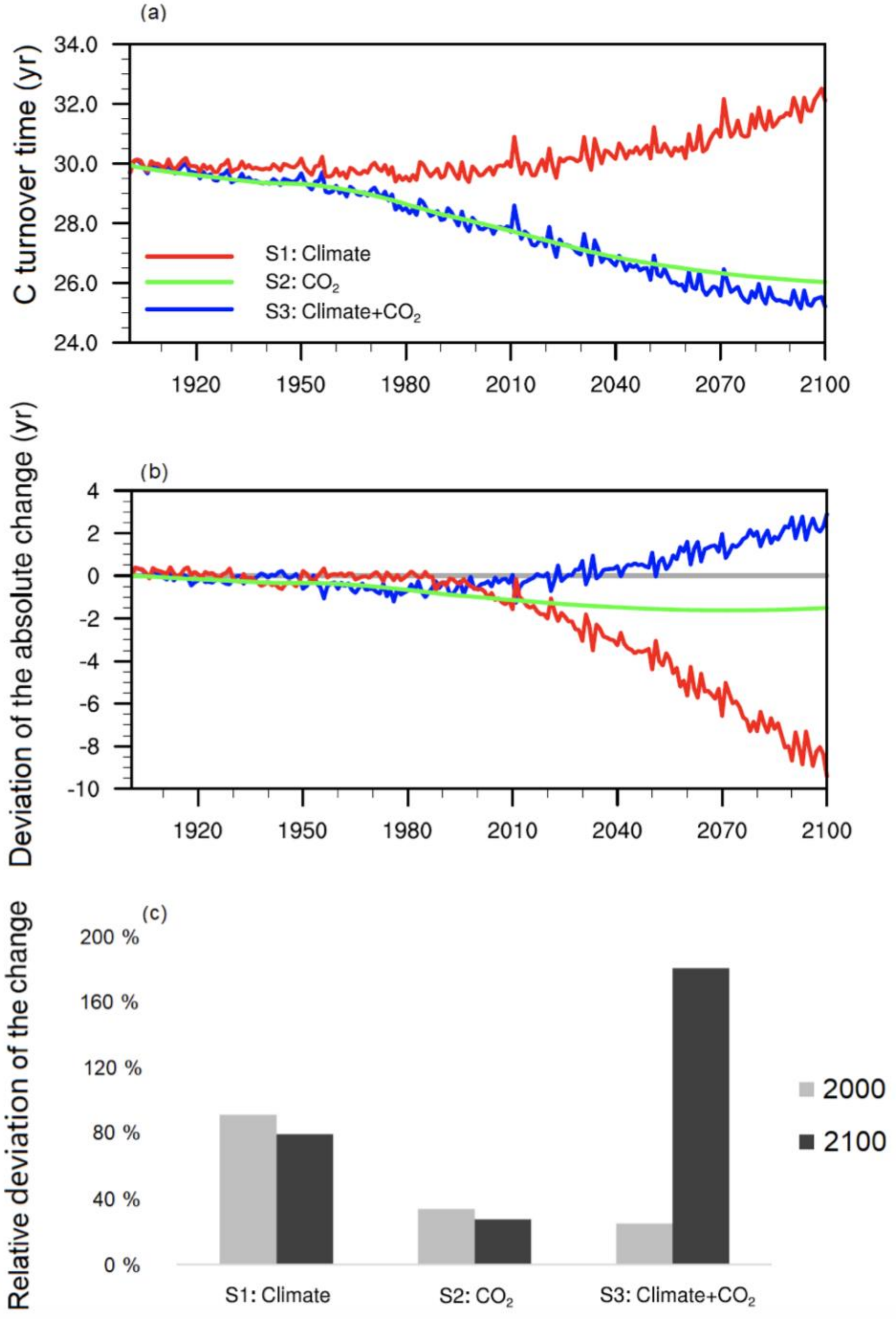
Diagnostics

3 √ Disequilibrium = GPP -ER



2 Residence time ?

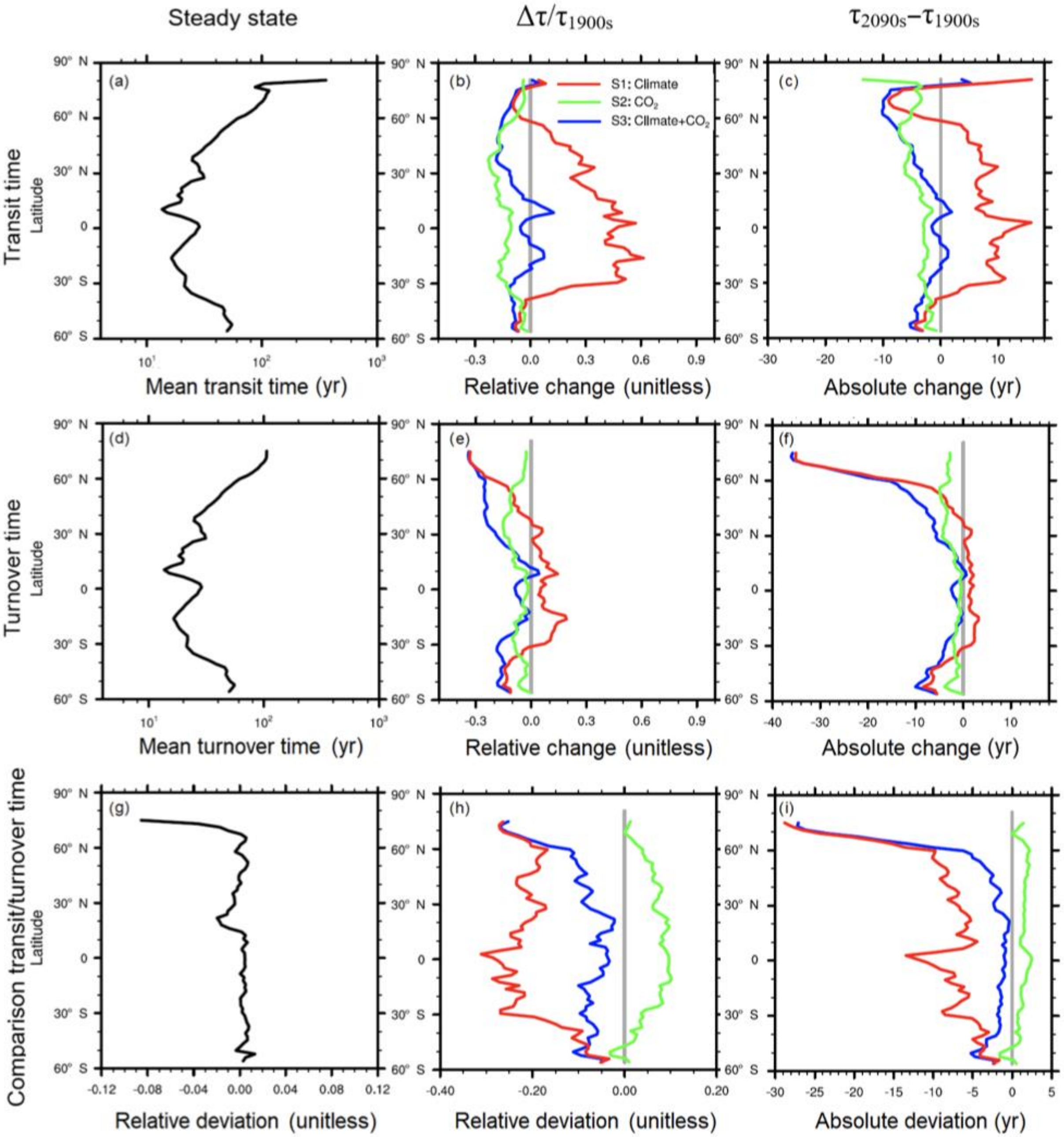
Luo and Weng 2011
Luo et al. 2017



Residence (or transit) time = time spent by carbon from entry to exit in a network of multiple pools

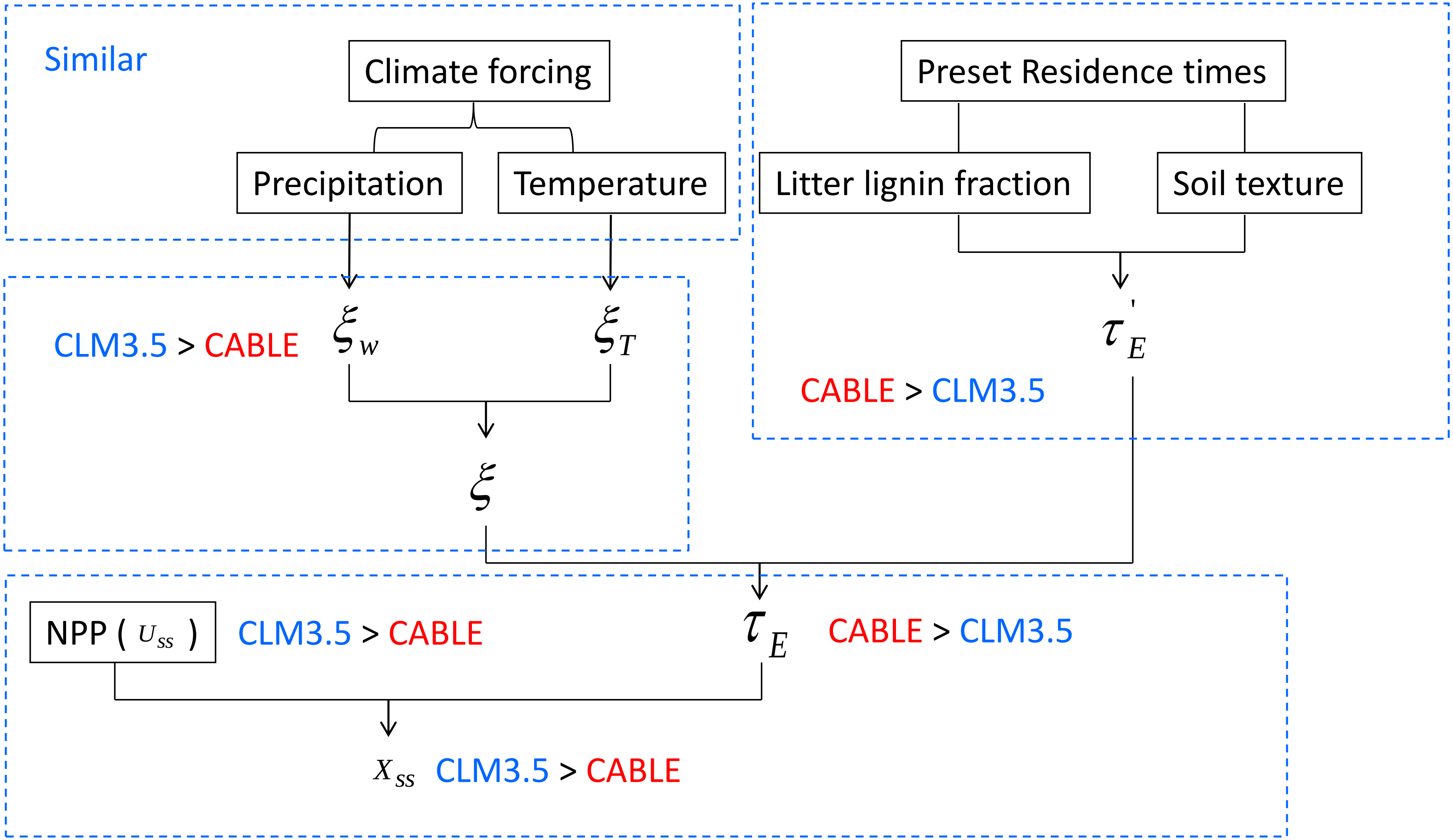
Turnover time =pool/flux

Lu et al. 2018 Biogeosciences



Lu et al. 2018 Biogeosciences

Traceability analysis for model Intercomparison



Similar

Climate forcing

Precipitation

Temperature

Preset Residence times

Litter lignin fraction

Soil texture

CLM3.5 > CABLE ξ_w ξ_T

τ'_E

CABLE > CLM3.5

ξ

NPP (u_{ss}) CLM3.5 > CABLE

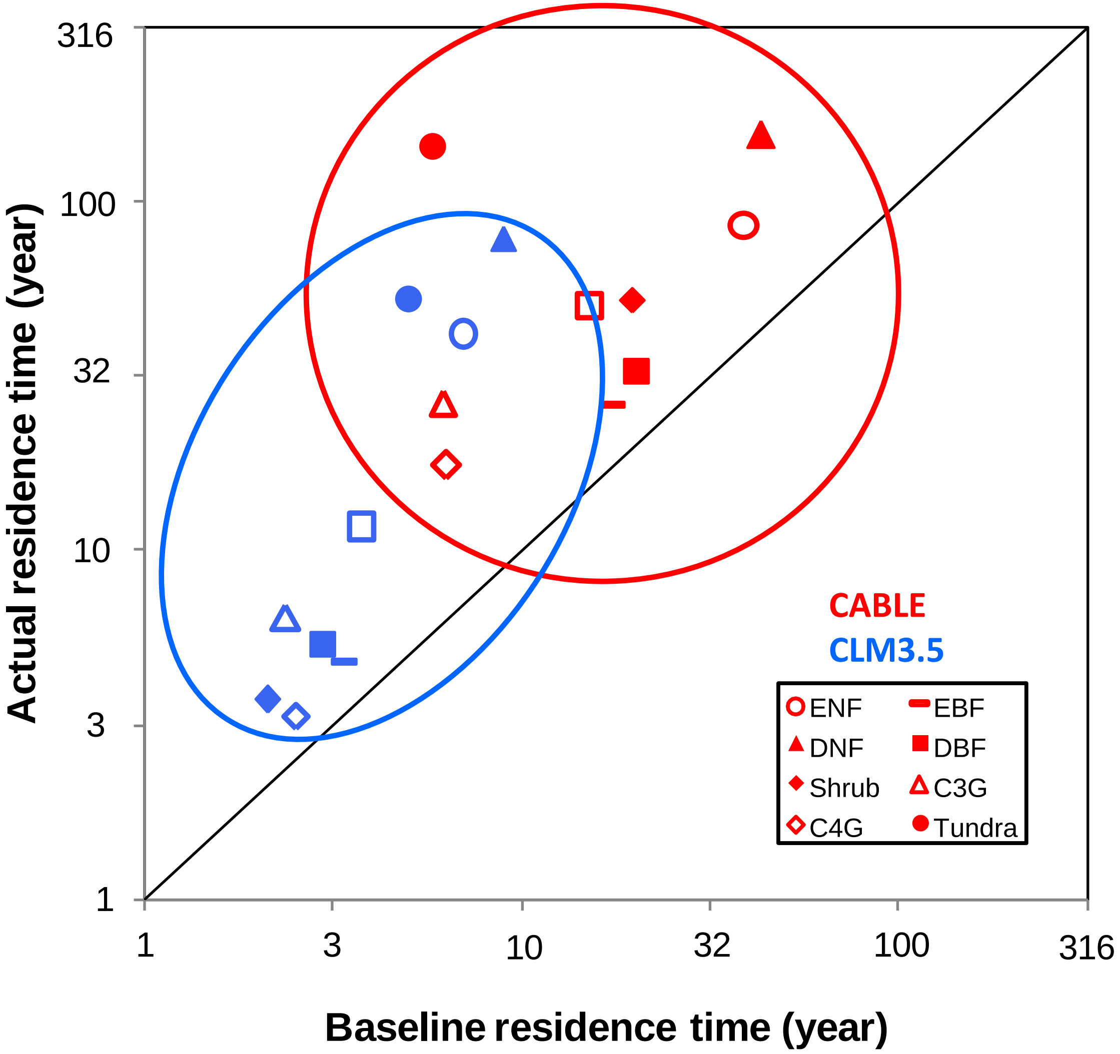
τ_E CABLE > CLM3.5

X_{ss} CLM3.5 > CABLE

Rafique *et al*. 2016 *Earth System Dynamics*

Longer residence times in CABLE than CLM3.5

$$\tau_E = \xi^{-1}\,\tau_E'$$

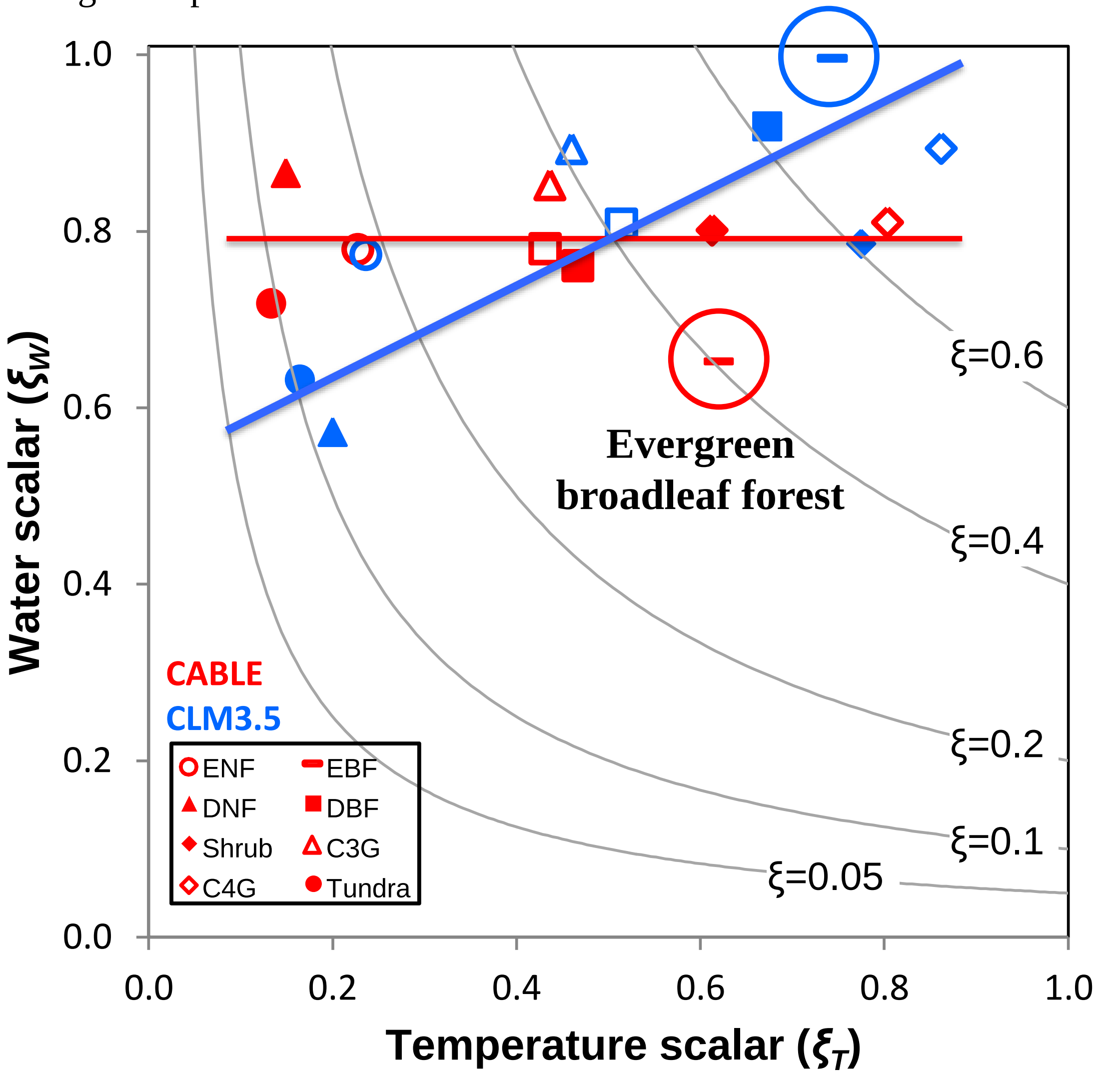


Rafique *et al*. 2016 *Earth System Dynamics*

Temperature and water scalars

$$\xi \quad = \xi_T \, \xi_W$$

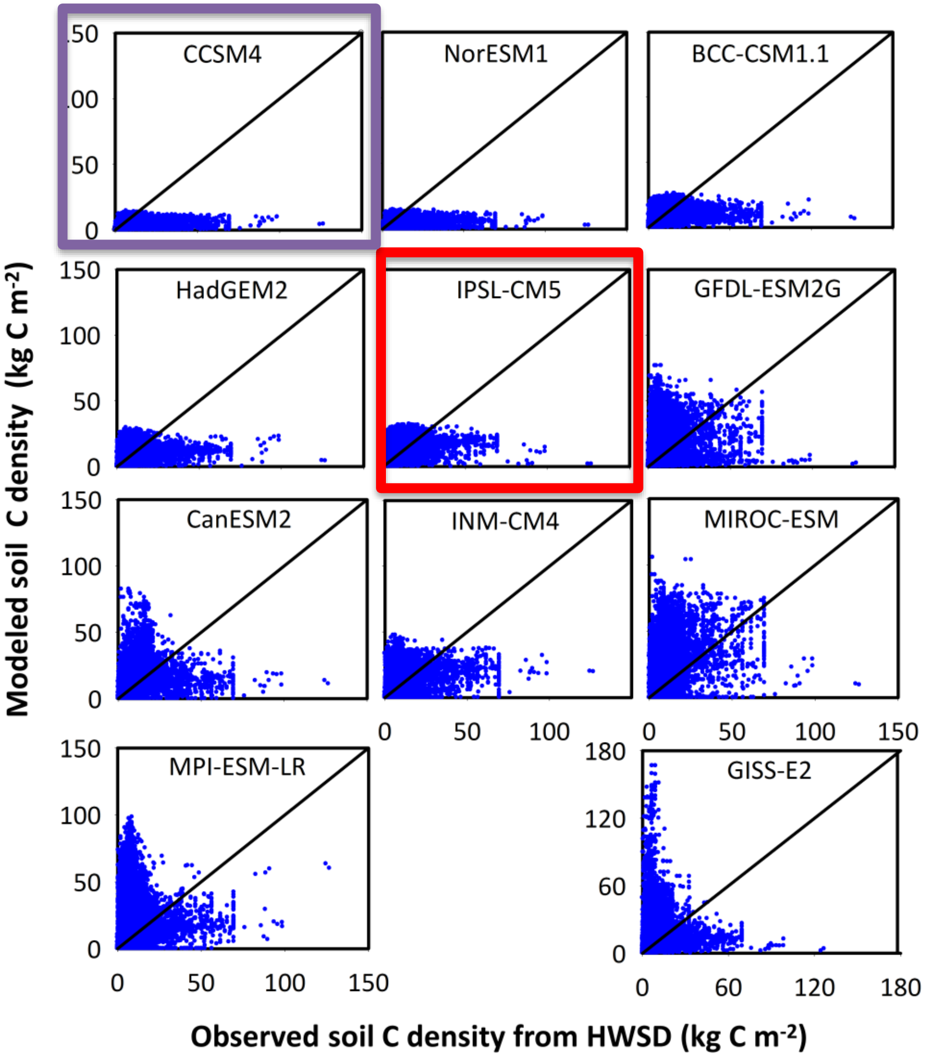
- Larger temperature and water limitations on most biomes in CABLE



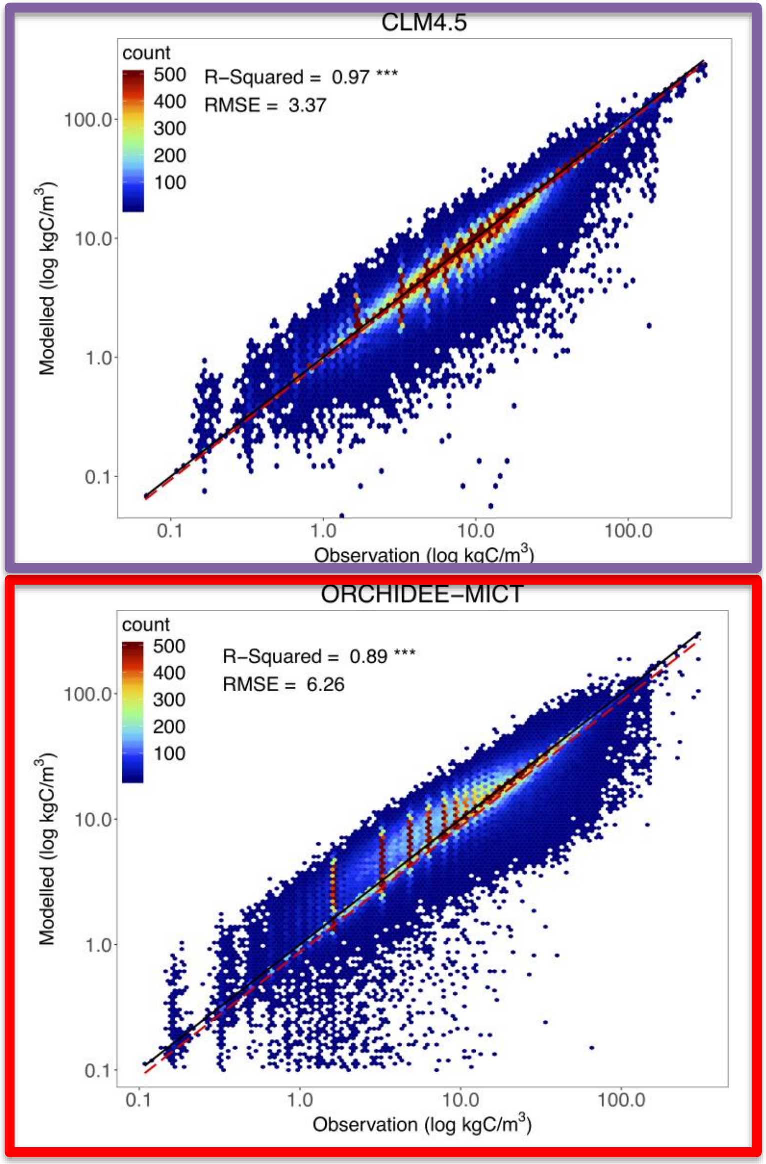
Evergreen broadleaf forest

Rafique *et al*. 2016 *Earth System Dynamics*

Data-model comparison



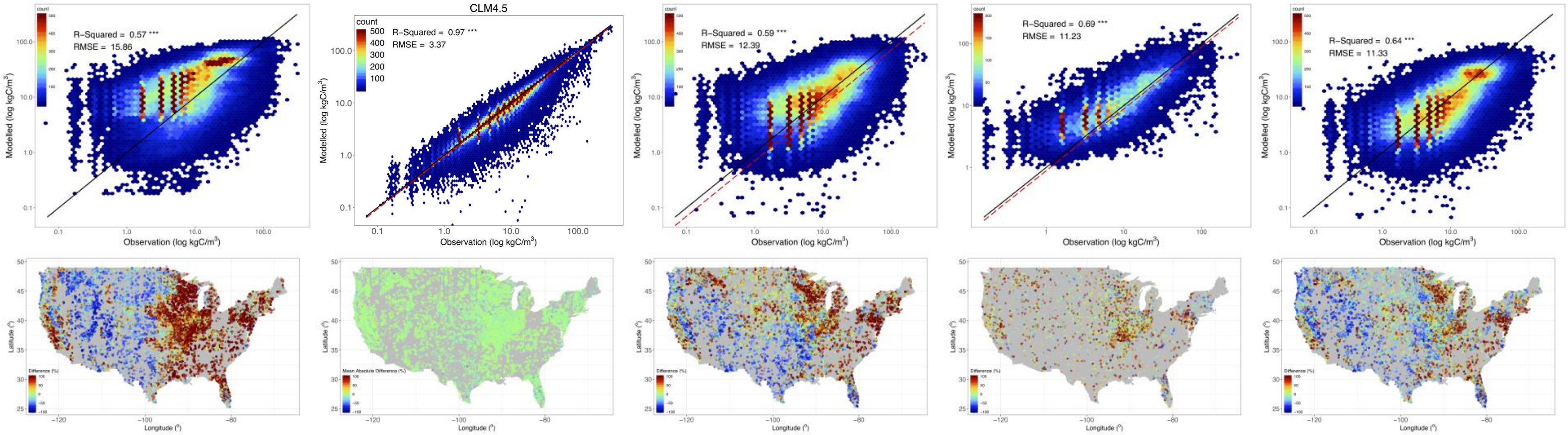
Luo et al. 2015 *Global Change Biology*



Tao et al. *In prep*.

Parameterization with data from ~30,000 soil profiles

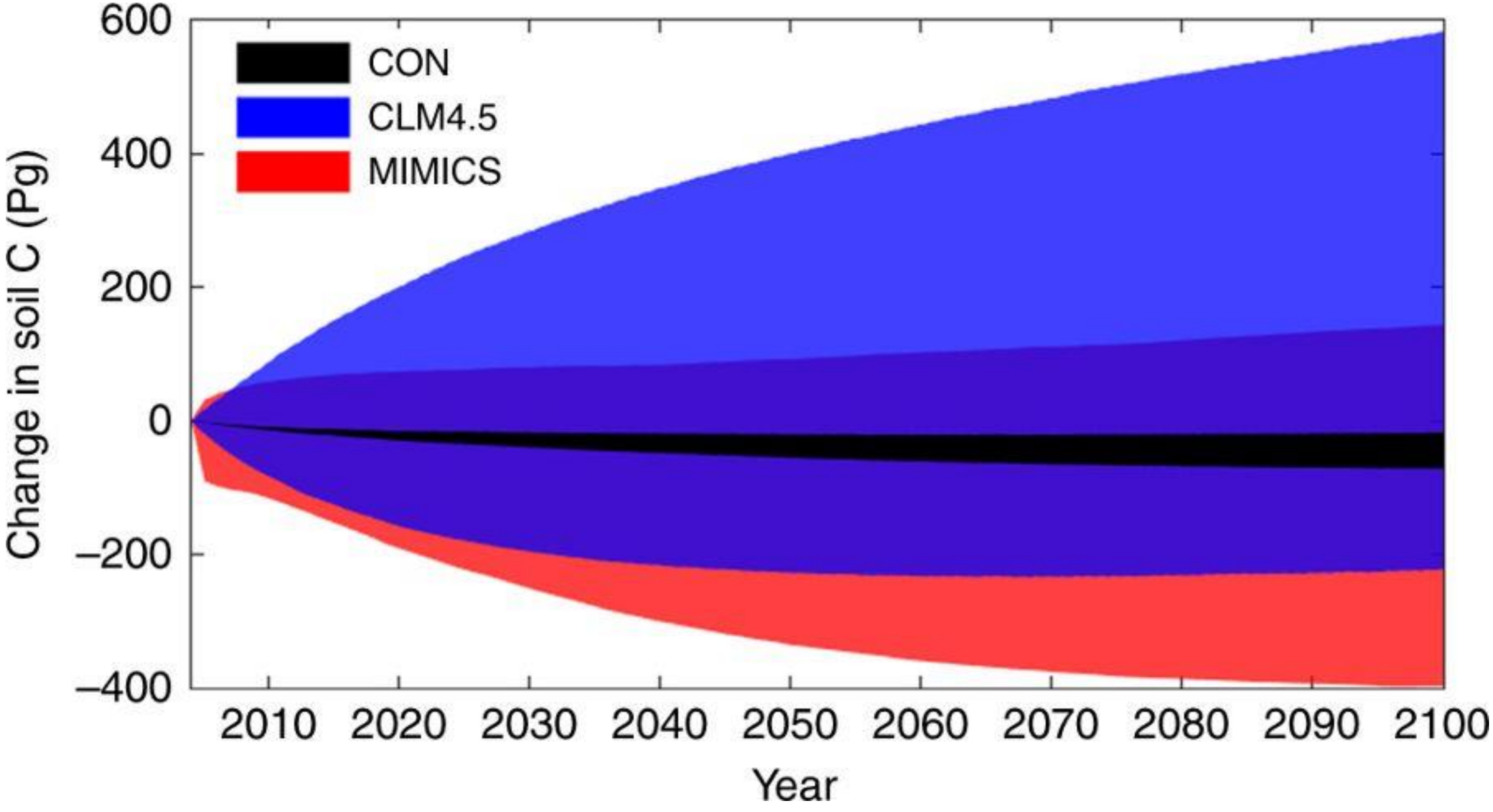
Method	Original Model	Site by Site	Random Sampling	Neural Networking	One Batch (MPI)
R-Squared	0.57	0.97	0.59	0.69	0.64
RMSE	15.86	3.37	12.39	11.23	11.33



Original Model Site by Site Random Sampling Neural Networking One Batch

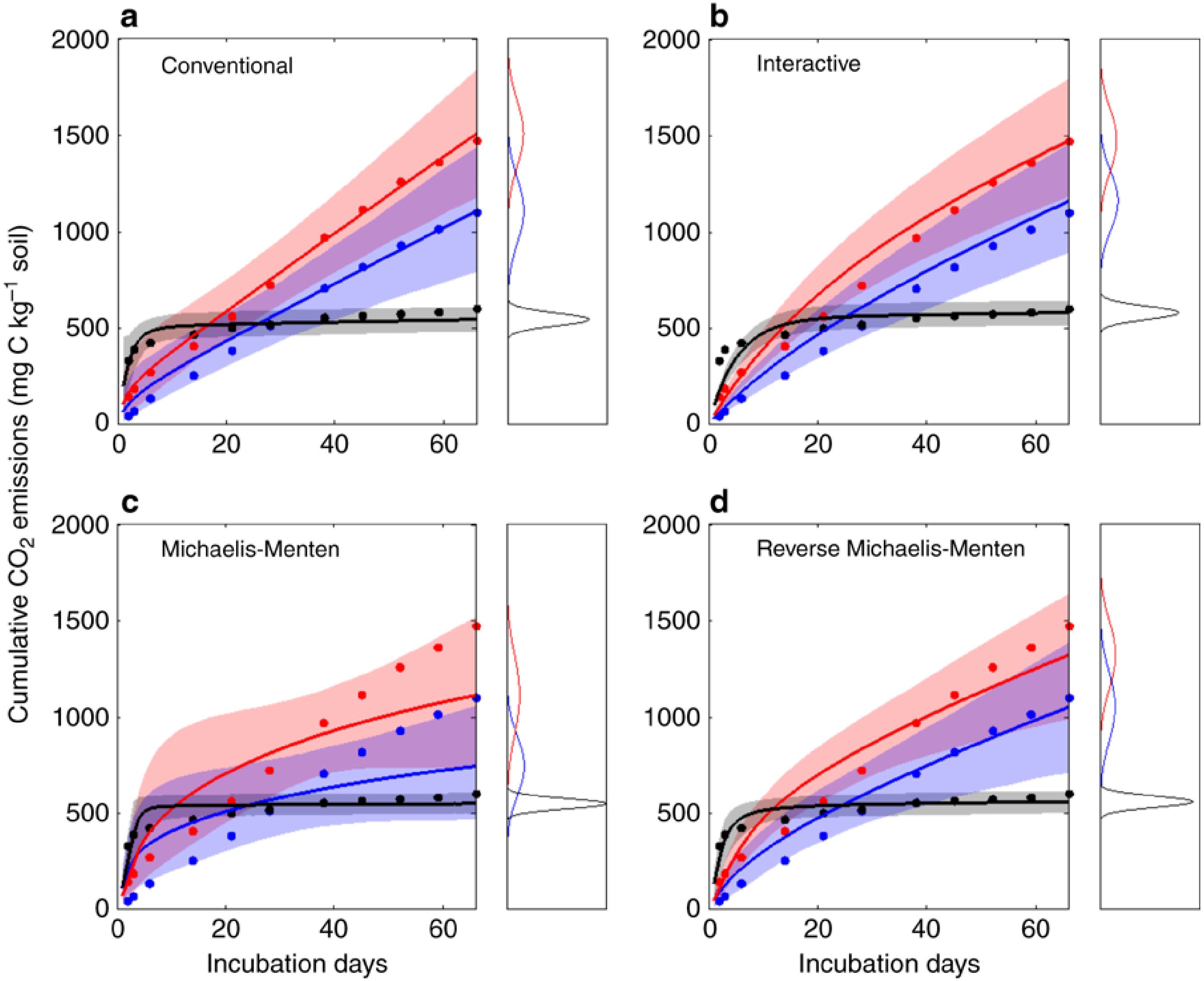
Tao et al. *In prep*.

Uncertainty in projections



Shi et al. 2018 *Nat. Comm.*

Selection of model structure for priming



Liang et al. 2018 *Nat. Comm.*

2nd Training Course on

New Advances in Land Carbon Cycle Modeling

Flagstaff, AZ, USA, May 13-24, 2019

- **Modelers** to gain simplicity in coding, diagnostic capability, computational efficiency for your models
- **Empiricists** to use your data to constrain models toward ecological forecasting
- **Scientists** who want to learn modeling, data assimilation, and ecological forecasting

Yiqi Luo, yiqi.luo@nau.edu
Lifen Jiang, lifen.jiang@nau.edu

Discussion points

1. How much details do we need in a model?
2. How can we develop an efficient model development-evaluation-improvement continuum?