# *Global Sensitivity Analysis a* SAND2013-10386C *Calibration for an Ecosystem Carbon Model*

C. Safta[1],D. Ricciuto[2],K. Sargsyan[1], B.Debusschere[1],H.N. Najm[1],P. Thornton[2]

[1]Sandia National Laboratories
Livermore, CA, USA

[2]Oak Ridge National Laboratory
Oak Ridge, TN, USA

B24B - Quantifying Uncertainty in Biogeochemical Studies
2013 AGU Fall Meeting
San Francisco, CA
December 10, 2013

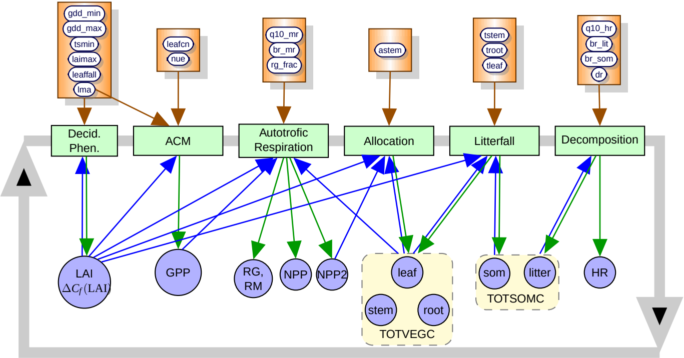
# Acknowledgement


- This work was supported by the US Department of Energy, Office of Science, under the project “Climate Science for a Sustainable Energy Future”, under the Biological and Environmental Research (BER) program.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

# Outline

# Ecosystem Carbon Model




- 19 parameters - arranged according to the process they control.
  - Effectively 18 parameters since the Aggregate Canopy Model (ACM) depends on the ratio $(nue/leafcn)$

# Ecosystem Carbon Model - Process Description

- Modified version of data assimilation linked ecosystem carbon (DALEC) model (Williams et al, 2005)
  - 3 vegetation carbon pools (leaf, stem, root)
  - 2 soil carbon pools (soil organic matter, litter)
  - Temperature-based deciduous phenology
  - Daily time step for all processes
  - Photosynthesis driven by ACM, which itself is calibrated to the Soil-Plant-Atmosphere (SPA) model (Williams et al., 1996)
- Key limitations
  - Single layer for decomposition calculations
  - Does not account for effects of soil moisture, texture
  - Does not account for effects of nutrient limitations

# Model setup

- Site: Harvard Forest, Massachusetts, USA
  - Temperate deciduous forest  70 years old
  - Eddy covariance, meteorological measurements from 1991-present
- Model drivers
  - Observed air temperature, solar radiation, vapor pressure deficit, CO2 concentration at flux tower site
  - Gap-filled using monthly mean
- Model constraints
  - Daily net ecosystem exchange (NEE) as measured by site PIs
  - NEE uncertainty estimates provided by North American Carbon Program (NACP) interim synthesis (Barr et al., 2009)
  - Incorporates u* (low turbulence threshold) and sampling uncertainties. Generally scales with flux magnitude.

# Variance-based Global Sensitivity Analysis

*ANOVA decomposition*: the model output $M(\theta)$ with $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is expressed in terms of a hierarchy of functions that account for the interaction between model parameters

$$M(\theta) = E[M(\theta)] + \sum_{i=1}^{n} f_i(\theta_i) + \sum_{1\leq i<j\leq n} f_{ij}(\theta_i, \theta_j) + \ldots$$

or

$$V = \sum_{i=1}^{n} V_i + \sum_{1\leq i<j\leq n} V_{ij} + \ldots$$

where

$$V = V[M(\theta)],\ V_i = V_{\theta_i}\left[E\left[M(\theta)|\theta_i\right]\right],\ V_{ij} = V_{\{\theta_i,\theta_j\}}\left[E\left[M(\theta)|\theta_i,\theta_j\right]\right]$$

# Variance-based Global Sensitivity Analysis

... normalized effects:

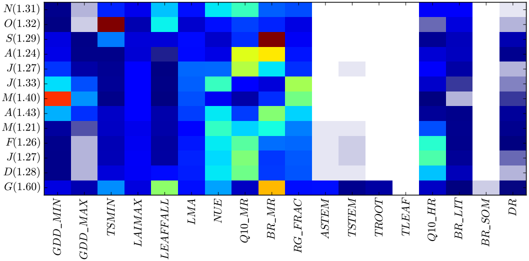
$$S_i = \frac{V_i}{V}, \quad S_{ij} = \frac{V_{ij}}{V} \quad \text{(Sobol indices)}$$

- ANOVA decomposition is unique if variables are independent
- For dependent variables we also employ the total effect index

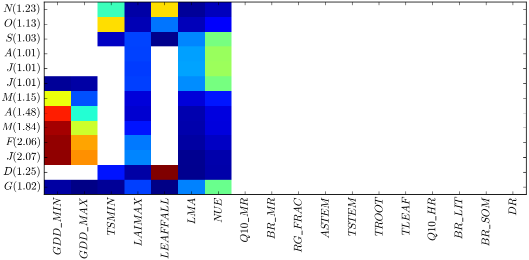
$$S_i^T = \frac{E_{\theta_{-i}}[V[M(\theta)|\theta_{-i}]]}{V}$$

where $\theta_{-i} = \{\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n\}$

# GSA Results-Total Effects for Monthly and Global Avg.

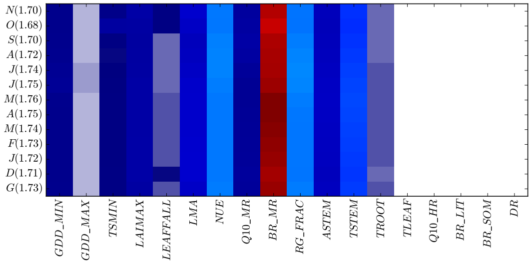



**NEE**

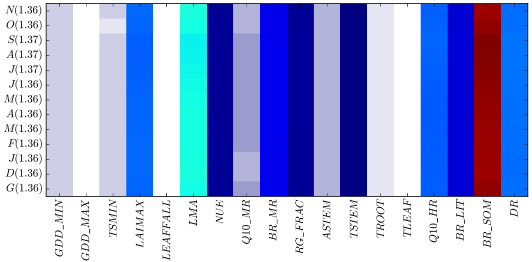



**GPP**

# GSA Results-Total Effects for Monthly and Global Avg.



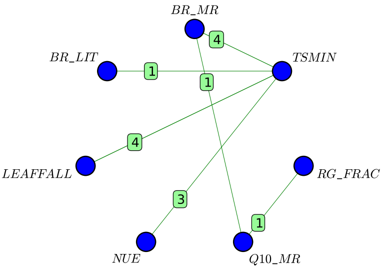

**TVC**



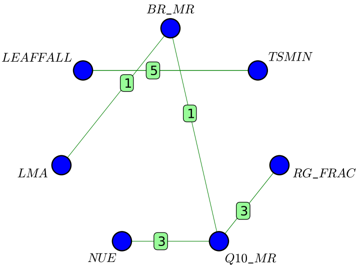

**TSC**

# GSA Results-Joint Sensitivity Indices

## NEE - October Average




## NEE - November Average



# Model Calibration using Daily NEE Observations

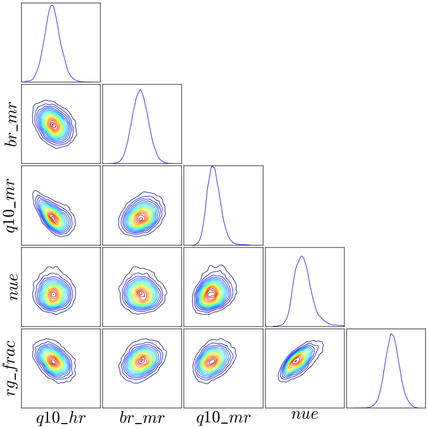
*Employ a Bayesian framework to calibrate model parameters based on available NEE observations*

$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$$

- prior probability $P(\theta)$, posterior probability $P(\theta|\mathcal{D})$, likelihood $P(\mathcal{D}|\theta)$, and evidence $P(\mathcal{D})$.
- uniform priors for all parameters with bounds based on expert opinion
- likelihood based on a zero-mean Gaussian discrepancy between model and observations; diagonal covariance matrix with standard deviations derived from observations

*The posterior probability is sampled using a Markov Chain Monte Carlo algorithm*

# Posterior Densities for Model Parameters



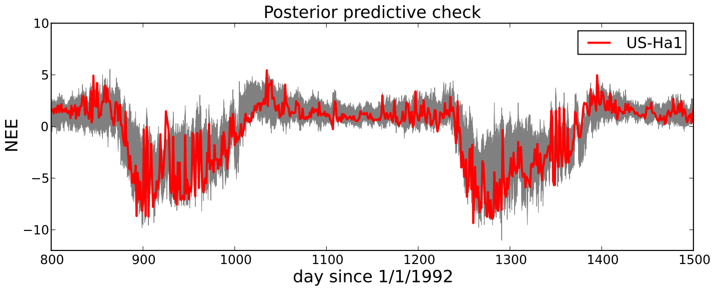

Distance correlation:

$$dCorr(X,Y) = \frac{dCov(X,Y)}{\sqrt{dVar(X)dVar(Y)}}$$

| $(X,Y)$ | $dCorr(X,Y)$ |
|---|---|
| (rg_frac, nue) | 0.67 |
| (q10_mr, q10_hr) | 0.51 |
| (rg_frac, q10_hr) | 0.34 |
| (br_mr, nue) | 0.05 |

# Predictive Capabilities of the Calibrated Model

$$p(y^{rep}|y) = \int p(y^{rep}|\theta)p(\theta|y)d\theta$$



# Summary

- Variance-based Global Sensitivity Analysis identifies parameters and interactions relevant to Quantities of Interest and seasons
- Posterior distributions for model parameters are inferred based on NEE observations. Correlations are observed for some parameters while others are not impacted by the data.
- Predictive skill of the model assessed through posterior predictive checks.

Future work:

- Transfer this experience to studies based on CLM
  - Tackle dimensionality: discover unimportant parameters
  - Tackle computational expense: surrogate models
  - Multi-site calibration