Global parameter sensitivity analysis

Ying-Ping Wang

CSIRO Marine and Atmospheric Research

Background

- **Objectives:**
  - Improve the performance of land surface models by tuning model parameters and quantifying model errors and their structure
- **Challenges:**
  - Too many parameters to be calibrated (20 to 100)
  - Too many model runs (10^5 to 10^6) required.
  - Model is expensive to run
  - **Technology roadmap:**
  - Use parameter screening to reduce the number of varying parameters
  - Use statistical model to characterize model errors as a function of
  - Iterate the above two until they converge!!!

Methods for parameter sensitivity analysis

Qualitative methods

1. Morris method: based on one-at-a-time sampling (MOAT)
2. Delta test: Based on nearest-neighbor analysis
3. Random Forest: Belongs to the class of tree-based methods
4. MARS: Multivariate Adaptive Regression Spline

Quantitative method

1. RSMSobol: (**R**esponse **S**urface **M**ethod based **Sobol**' method)
   Based on variance decomposition

Some results from applications to two global land surface models

- CoLM: developed by Dai et al. (2010). Sensitivities of six model output to 40 model parameters are analyzed (Li et al. 2013)
- CABLE: Sensitivities of GPP and LE to 46 parameters were analyzed (see Lu et al. 2013).

Conclusion 1: **400** model runs are enough for parameter screening based on qualitative SA methods

MARS with different sampling

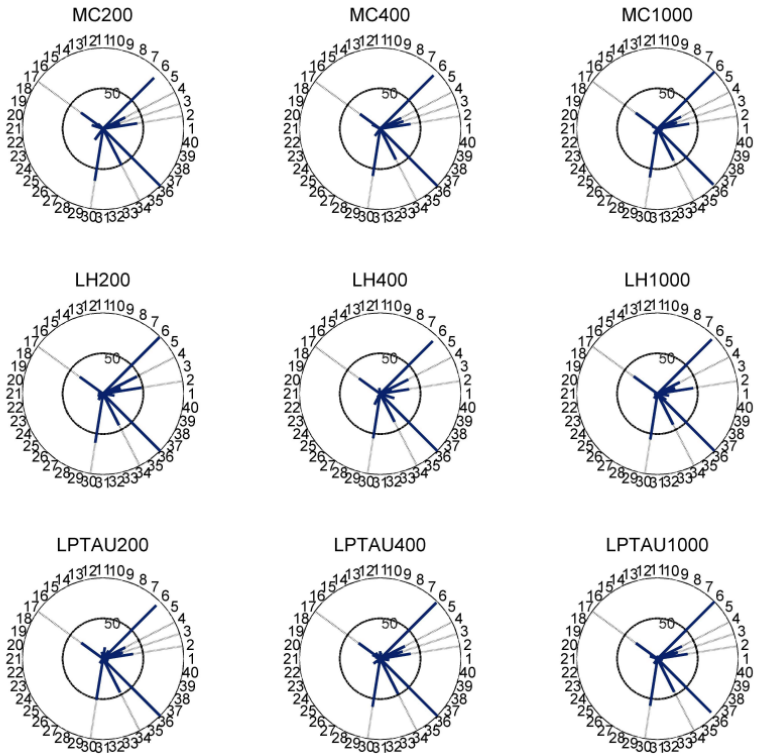
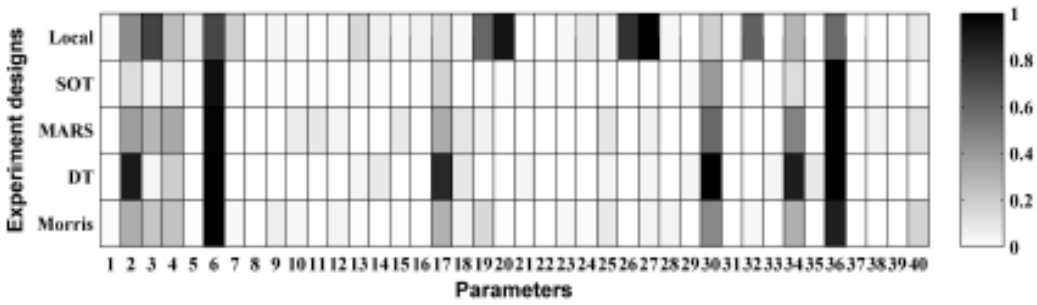


Fig. 3. The sensitivity score of sensible heat given by MARS. The length of needles represents the sensitivity score.

Different methods



MC:	Monte Carlo
LH:	Latin Hypercube
LPTAU:	quasi-random sequence

[Jianduo Li et.al. 2013, HESS]

Conclusion 2: The effectiveness is confirmed by quantitative Sobol' method.

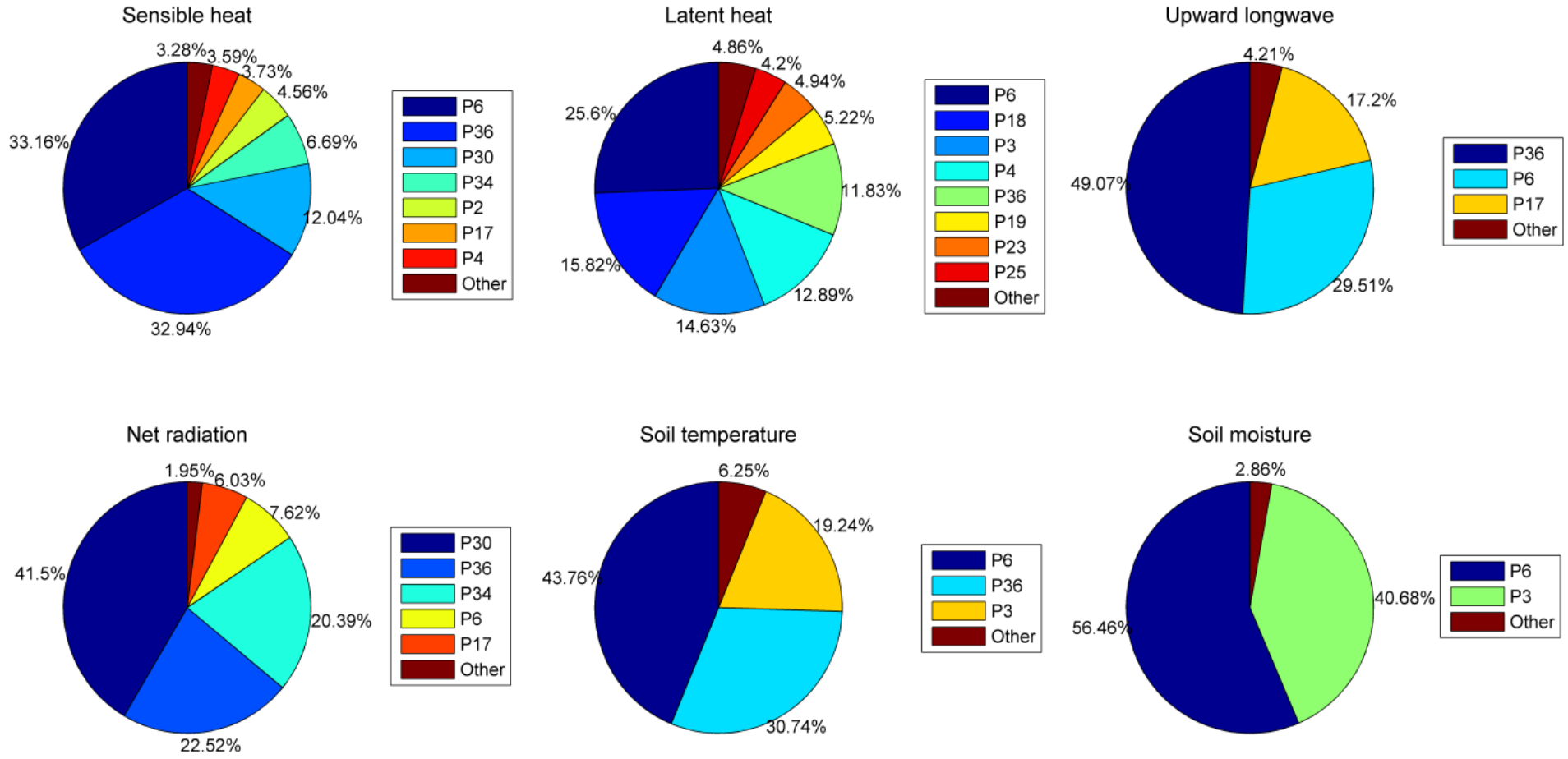


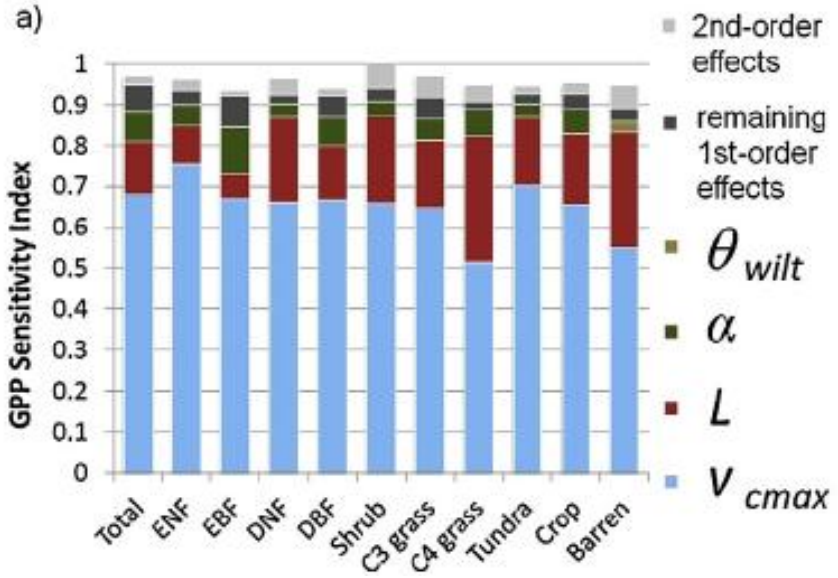
Fig. 12. The relative importance of parameters obtained by RSMSobol' total effect analysis.

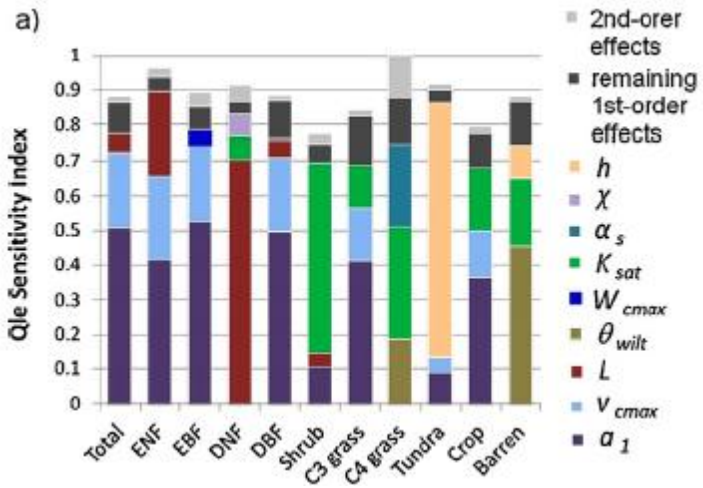
[Jianduo Li et.al. 2013, HESS]

Applications to CABLE

- Sensitivities of annual GPP and LE to 43 model parameters
- For Morris Method
  - CABLE was run to steady state (20 years) for each set of 43 parameter values. A total of 256, 512, 1024 parameter sets we tried.
- For Sobol method
  - Only ran CABLE for the important parameters selected using Morris method

The Morris method

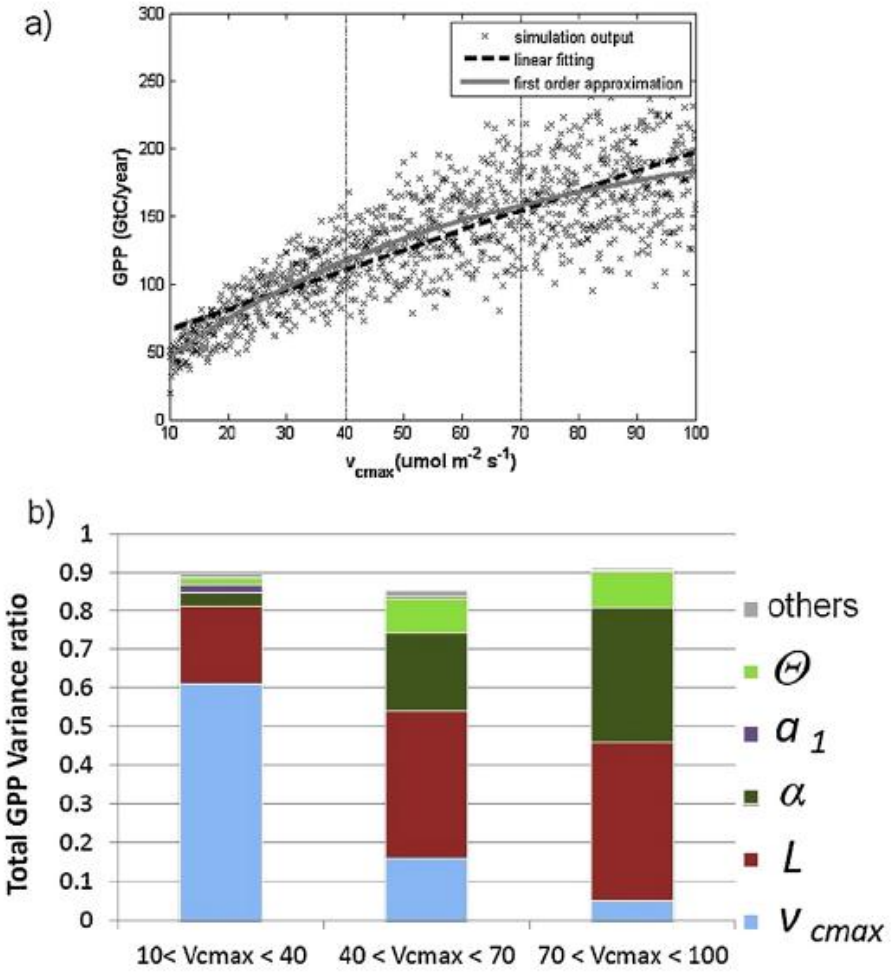




Key findings of CABLE

- Only 960 simulations are required to identify 5 most sensitive parameters for simulated annual GPP and LE out of 22.
- Over 80% of the variance of GPP can be explained by 5 most sensitive parameters to simulation of annual GPP.
- Over 70% of the variance of LE can be explained by 5 most sensitive parameters to simulation of annual LE.
- Maximum carboxylation rate (v_{cmax}), canopy leaf area index (LAI) and quantum use efficiency (α) are 3 most important parameters for GPP of most PFTs.
- An empirical model in the leaf stomatal conductance model (a_1), maximum carboxylation rate (v_{cmax}) and LAI are 3 most important parameters for annual LE of most PFTs.

Parameter sensitivity depends on its prescribed range



Some key references

Ziehn, T and Tomlin, AS (2009) A software tool for global sensitivity analysis of complex models. *Environmental Modelling Software*, **24**:775-785.
Li, J et al. (2013) Assessing parameter importance of the common land model based on qualitative and quantitative sensitivity analysis. *Hydrol. Earth Syste. Sci*, **17**:3279-3293
Lu et al. (2013) An efficient method for global parameter sensitivity analysis and its applications to the Australian community land surface model (CABLE). *Agric. For. Meteorol*., **182-183**:292-303.

Acknowledgment

The following people have contributed to this presentation:

- Dr Wei Gong, Beijing Normal University
- Mr Jianduo Li, Beijing Normal University
- Dr Chris Lu, Beijing Normal University/CSIRO
- Dr Tilo Ziehn, CSIRO

Topics for the working group

1: Parameter estimation/data assimilation
Who uses it? What package or implementation do you use? And what are your application?
2: what are the most significant limiting factor in your applications
3: what kinds of observations do you use? And how do you estimate the uncertainties of the observations?
4: model errors: how do you account for the non-randomness of model errors in your applications
5: how can we share the experiences?