Some Ideas for Land Surface Model Diagnosis

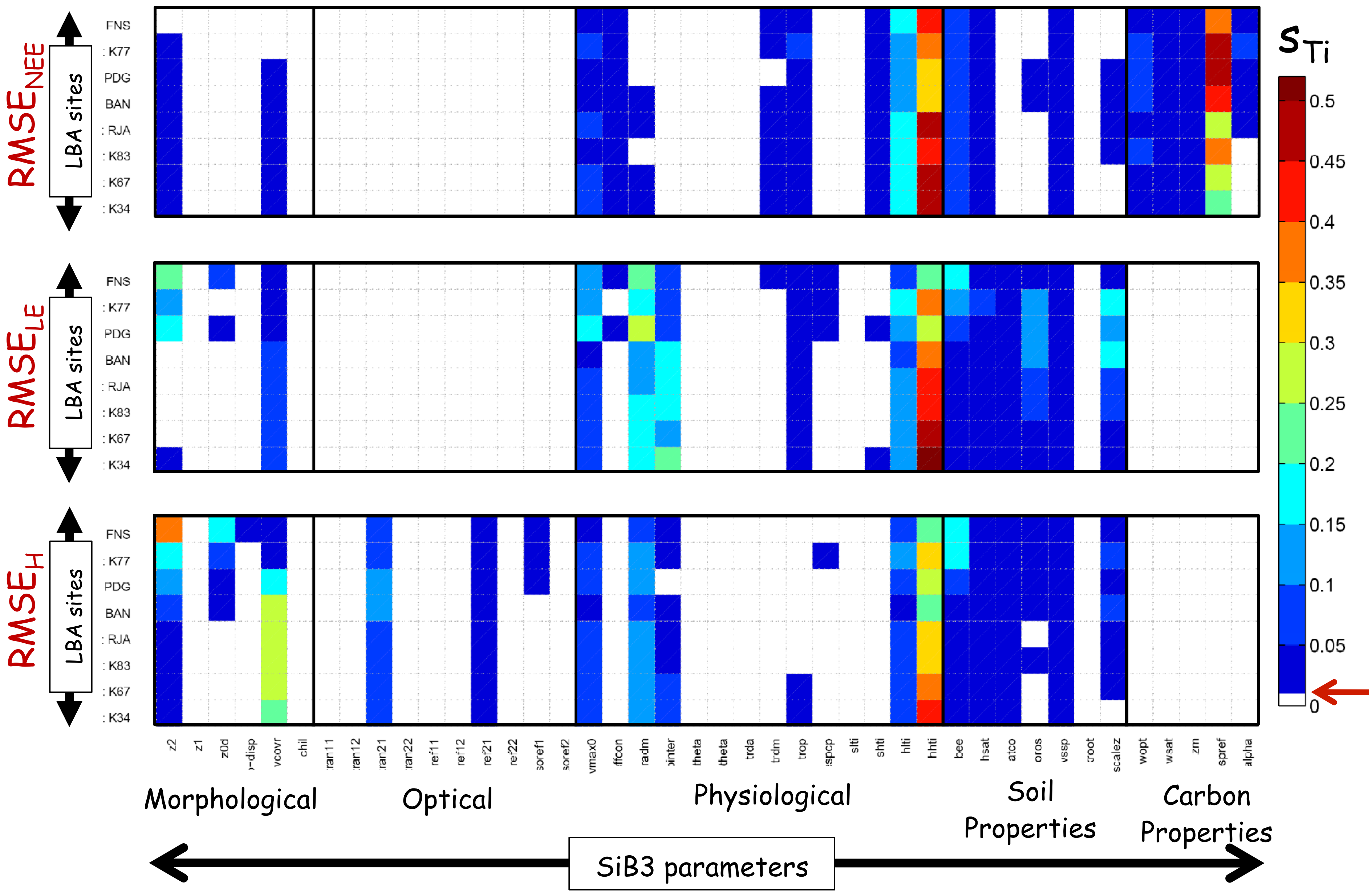
Rafael Rosolem

W. J. Shuttleworth, H. V. Gupta,
L. G. G. de Goncalves, X. Zeng

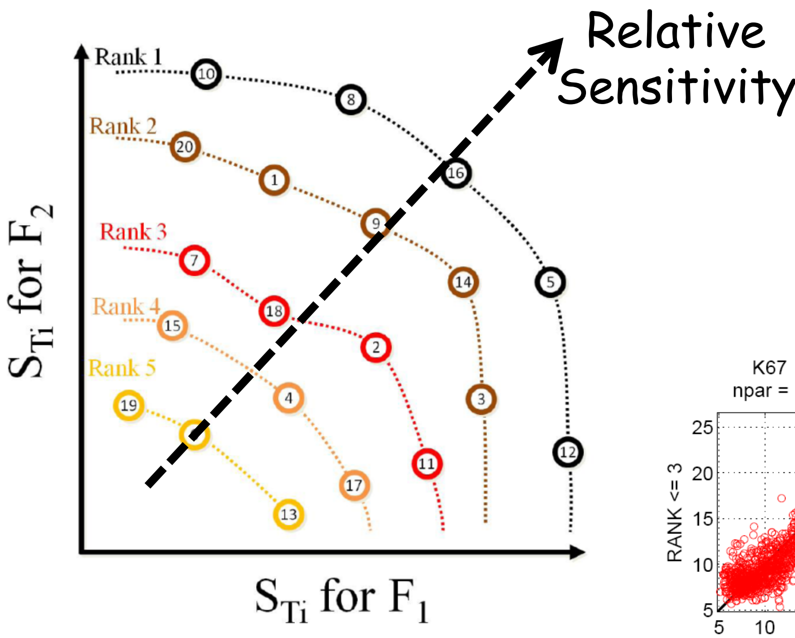
16th Annual CESM Workshop
Land Model Working Group Session
June 21, 2011

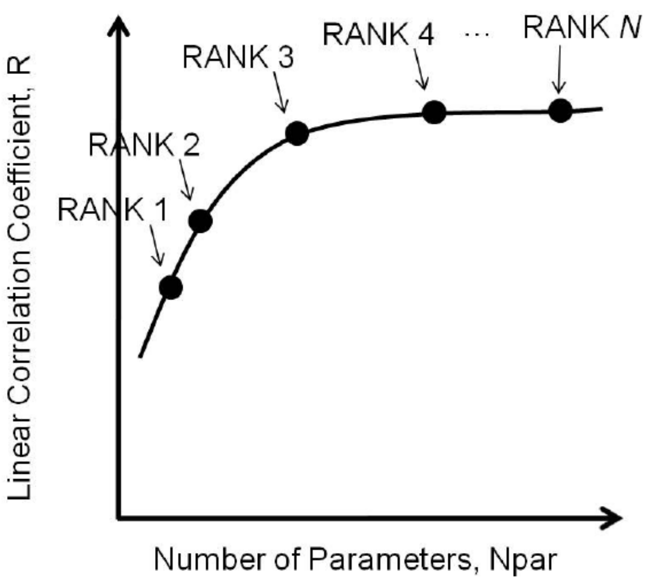
THE UNIVERSITY OF ARIZONA®

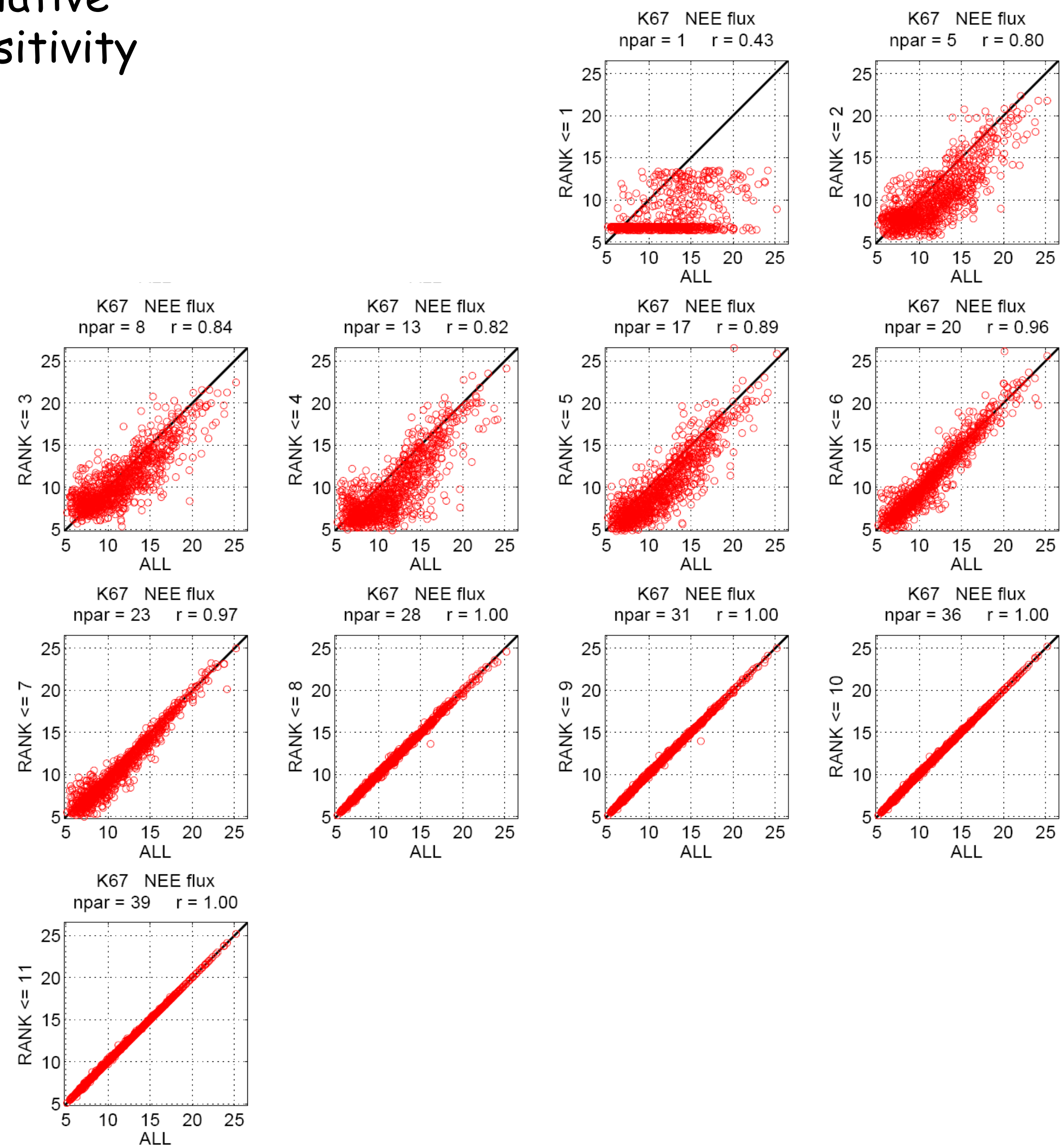
Sobol application: Conventional approach



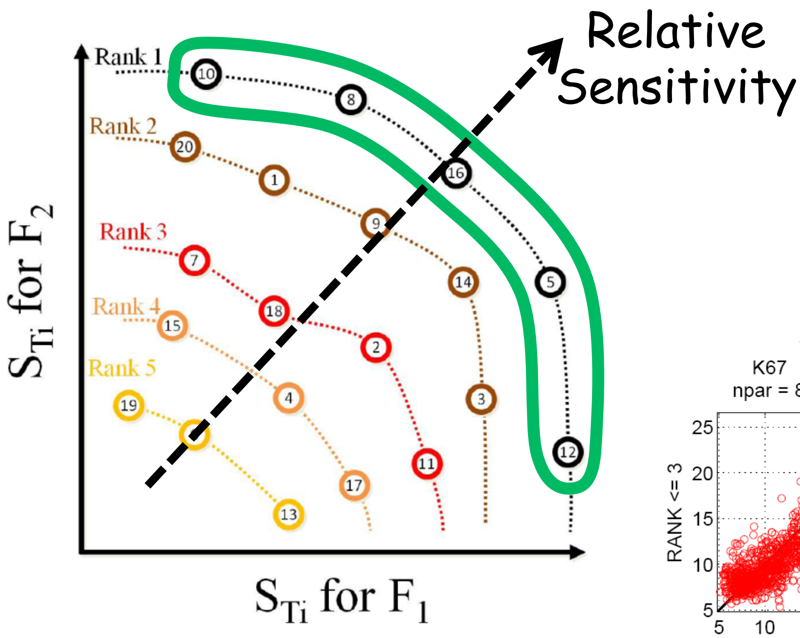
Sobol application: New multi-objective approach

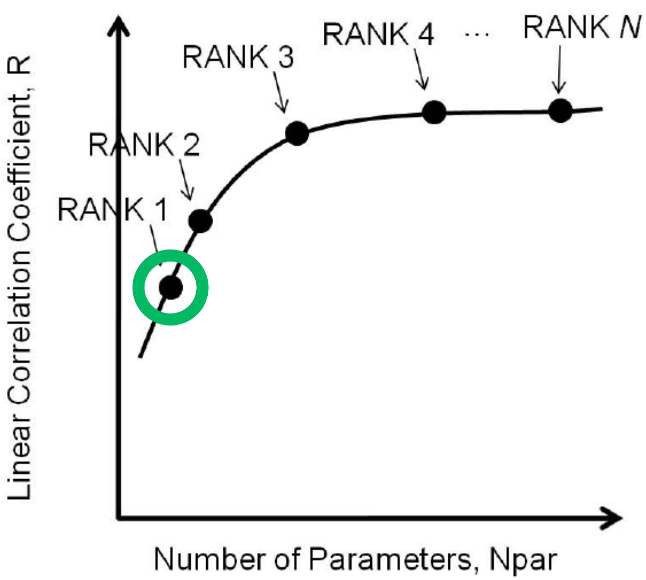


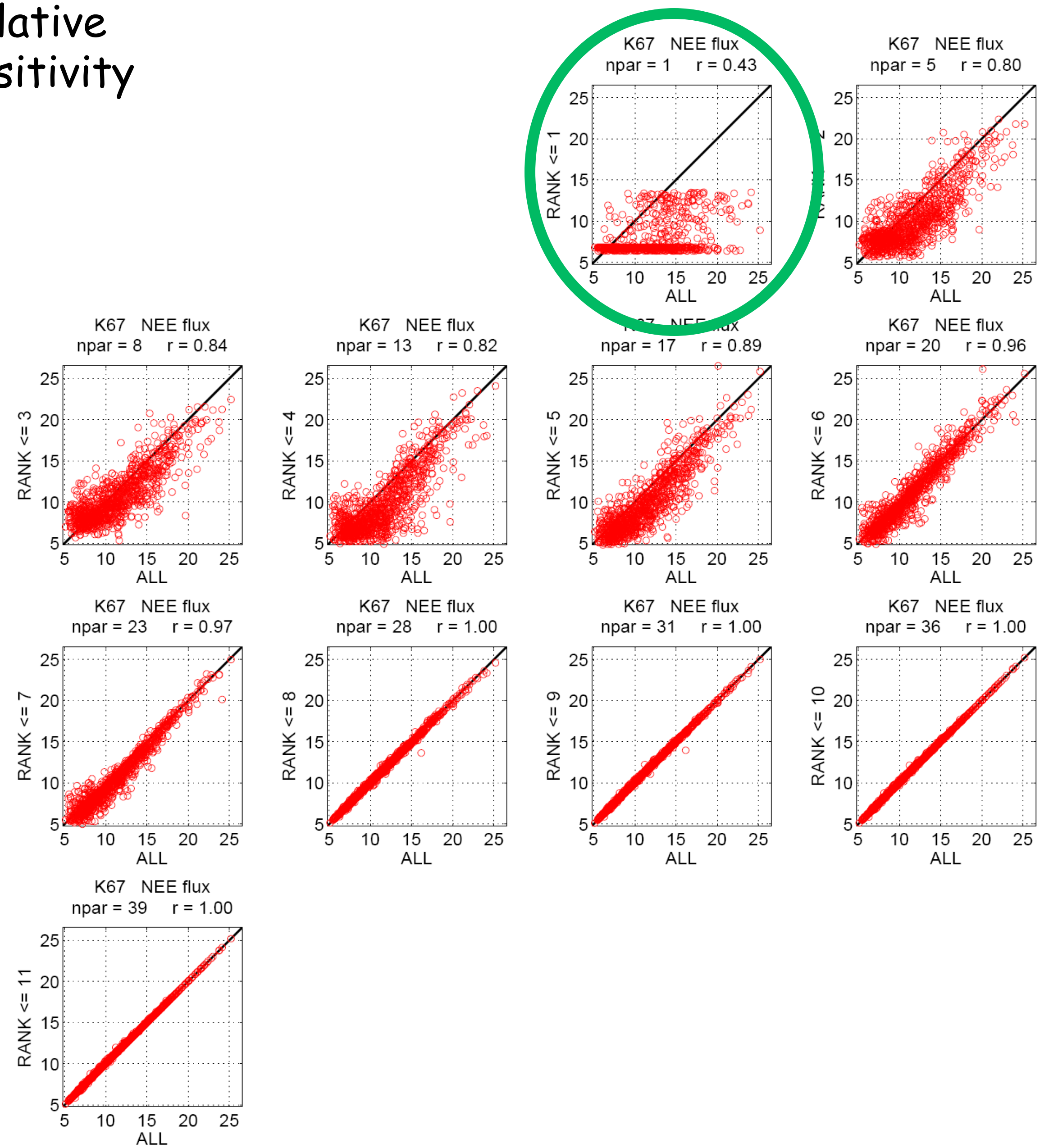




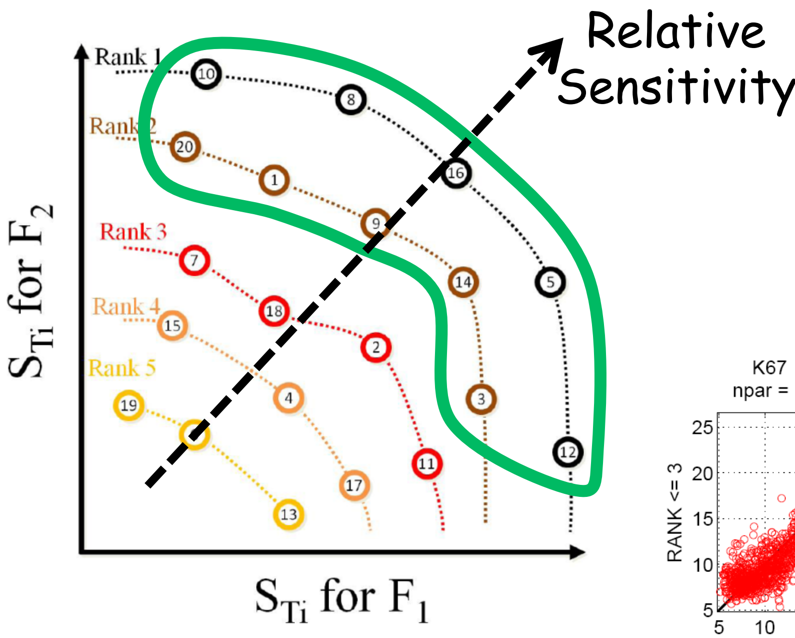
Sobol application: New multi-objective approach

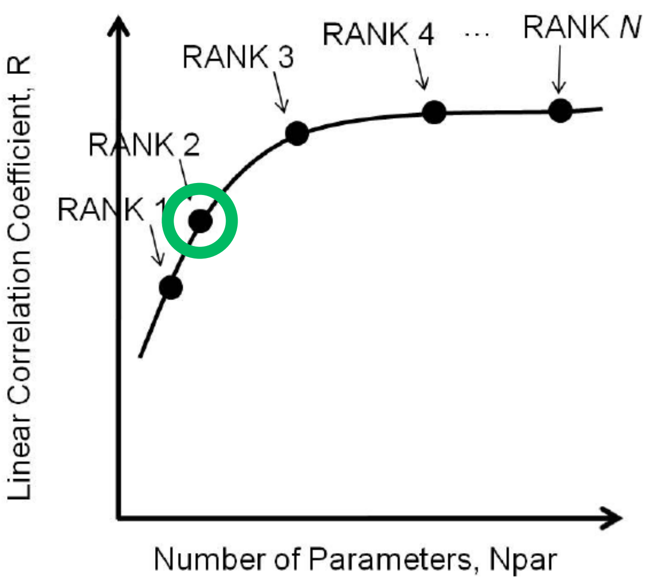


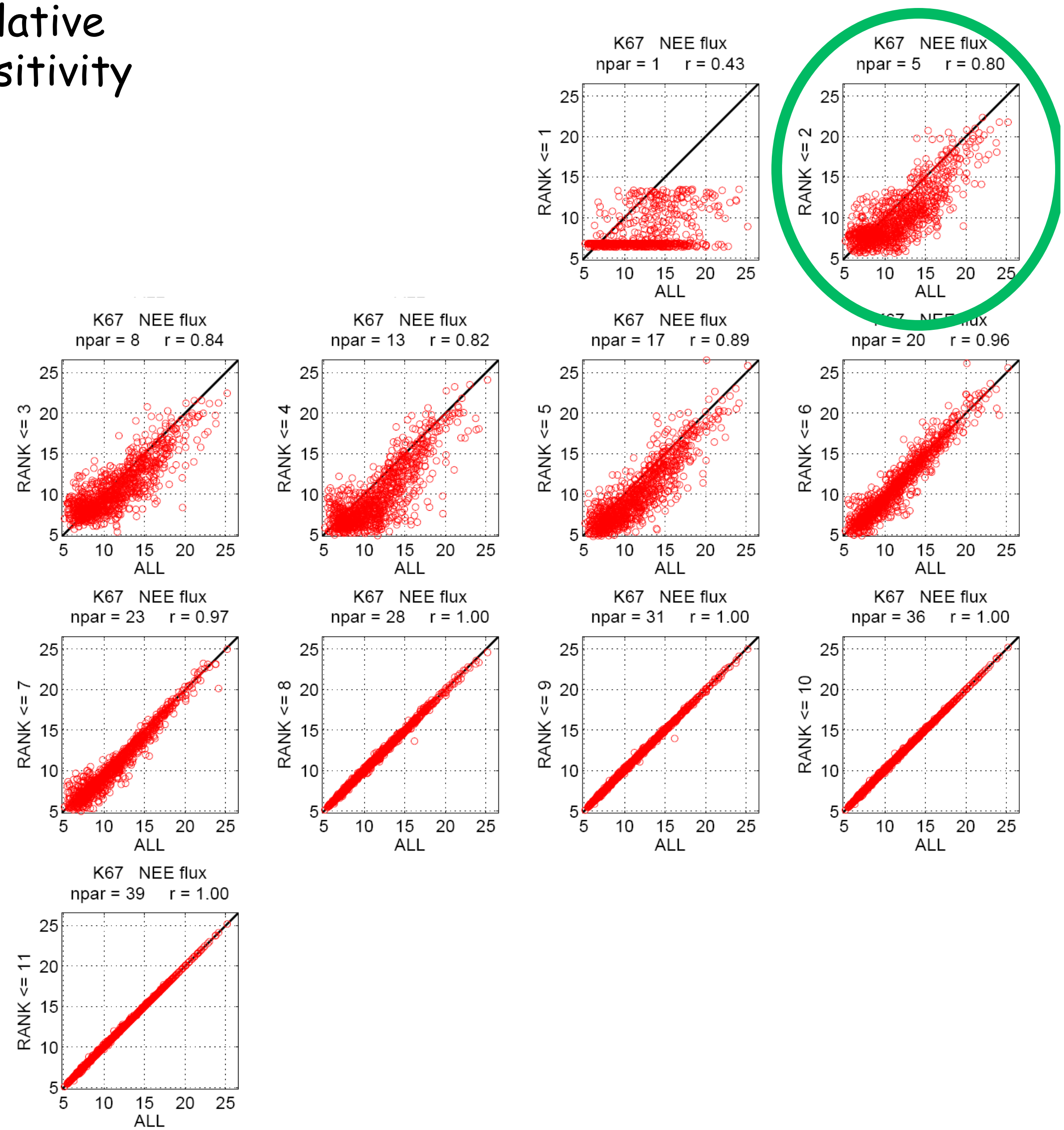




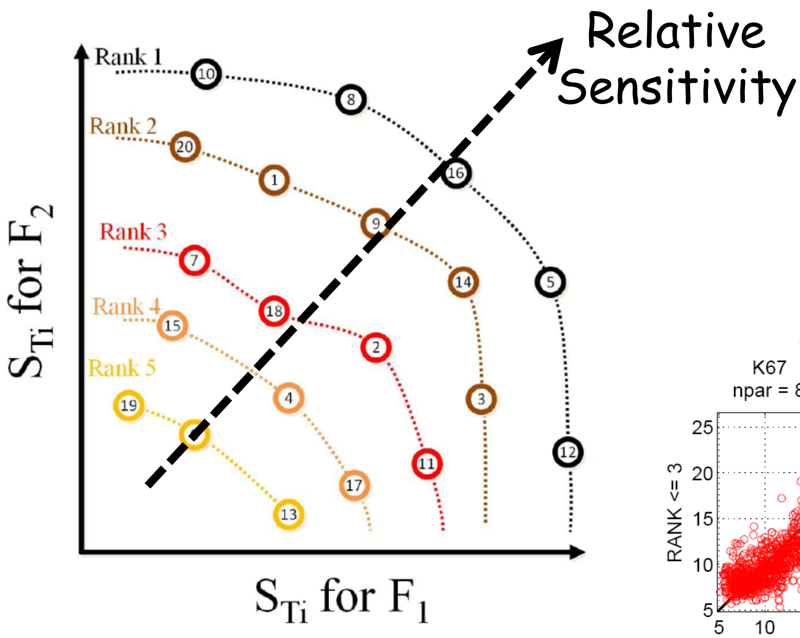
Sobol application: New multi-objective approach

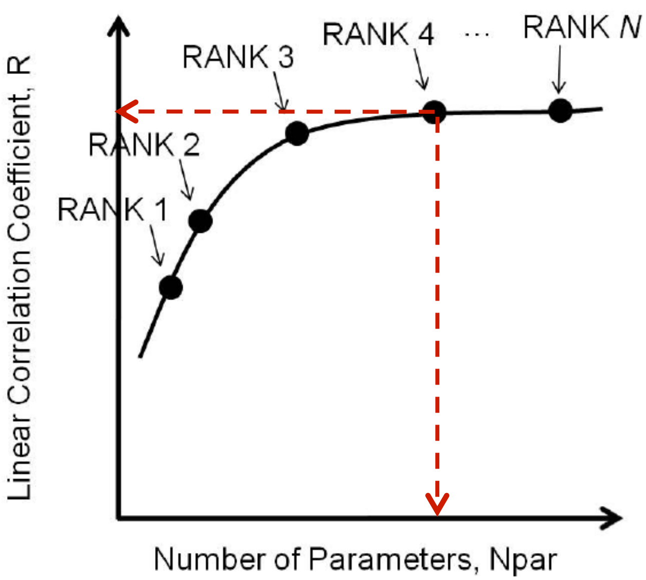


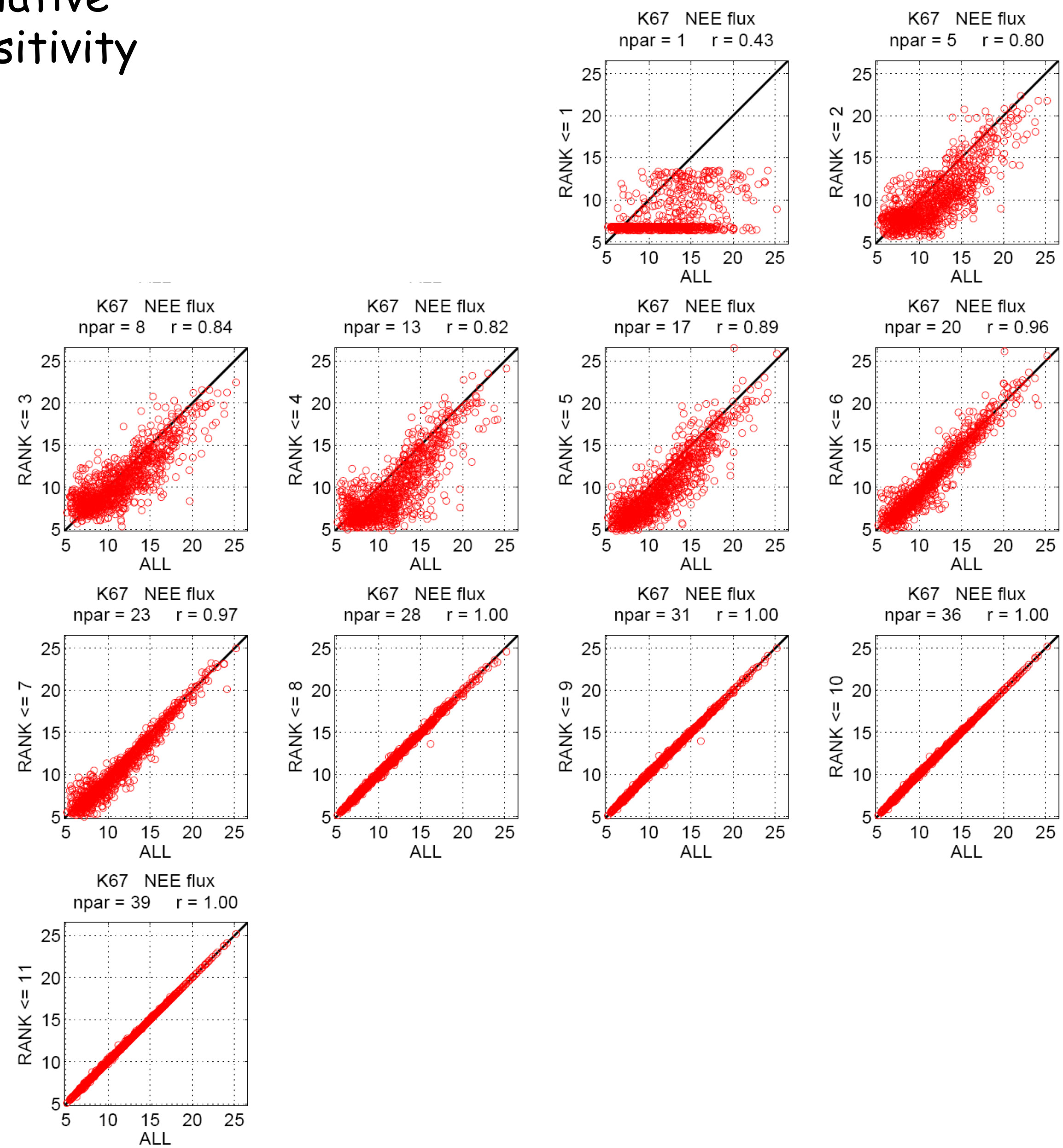




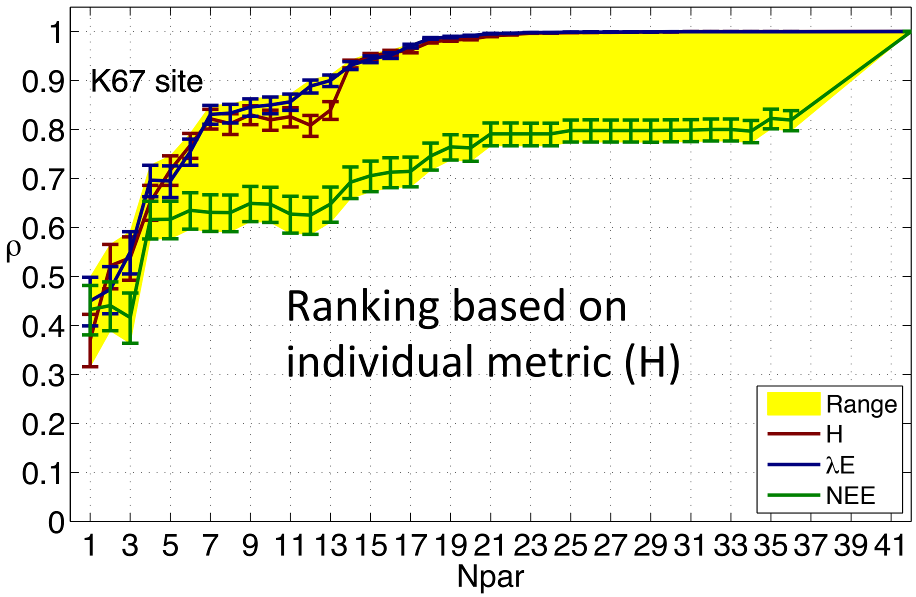
Sobol application: New multi-objective approach



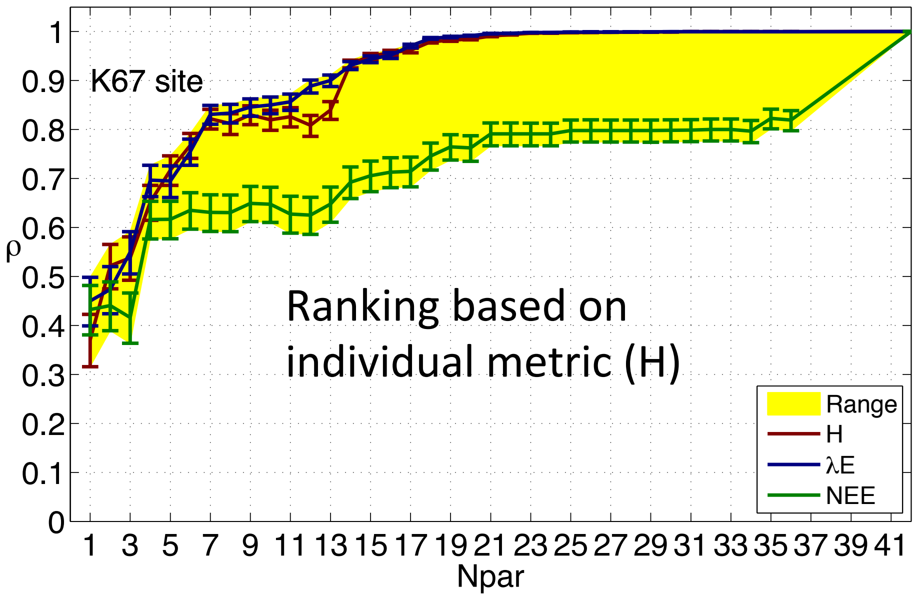


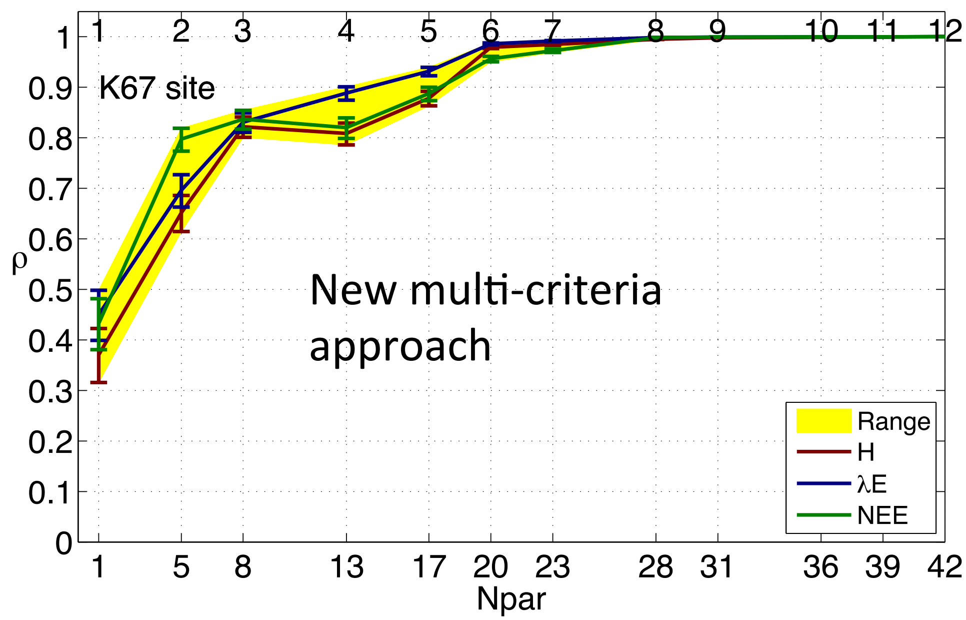


Single-criterion versus Multi-criteria Ranking



Single-criterion versus Multi-criteria Ranking





A Multi-Algorithm Genetically Adaptive Multiobjective method - AMALGAM (Vrugt and Robinson, 2007)

PNAS

Improved evolutionary optimization from genetically adaptive multimethod search

Jasper A. Vrugt* and Bruce A. Robinson

Los Alamos National Laboratory, EES-6, Mail Stop T003, Los Alamos, NM 87545

Communicated by Stirling A. Colgate, Los Alamos National Laboratory, Los Alamos, NM, November 27, 2006 (received for review March 29, 2006)

In the last few decades, evolutionary algorithms have emerged as a revolutionary approach for solving search and optimization problems involving multiple conflicting objectives. Beyond their ability to search intractably large spaces for multiple solutions, these algorithms are able to maintain a diverse population of solutions and exploit similarities of solutions by recombination. However, existing theory and numerical experiments have demonstrated that it is impossible to develop a single algorithm for population evolution that is always efficient for a diverse set of optimization problems. Here we show that significant improvements in the efficiency of evolutionary search can be achieved by running multiple optimization algorithms simultaneously using new concepts of global information sharing and genetically adaptive offspring creation. We call this approach a multialgorithm, genetically adaptive multiobjective, or AMALGAM, method, to evoke the image of a procedure that merges the strengths of different optimization algorithms. Benchmark results using a set of well known multiobjective test problems show that AMALGAM approaches a factor of 10 improvement over current optimization algorithms for the more complex, higher dimensional problems. The AMALGAM method provides new opportunities for solving previously intractable optimization problems.

evolutionary search | multiple objectives | optimization problems | Pareto front

solutions by recombination. These attributes lead to efficient convergence to the Pareto-optimal front in a single optimization run (13). Of these, the nondominated sorted genetic algorithm II (NSGA-II) (14) has received the most attention because of its simplicity and demonstrated superiority over other methods.

Although the multiobjective optimization problem has been studied quite extensively, current available evolutionary algorithms typically implement a single algorithm for population evolution. Reliance on a single biological model of natural

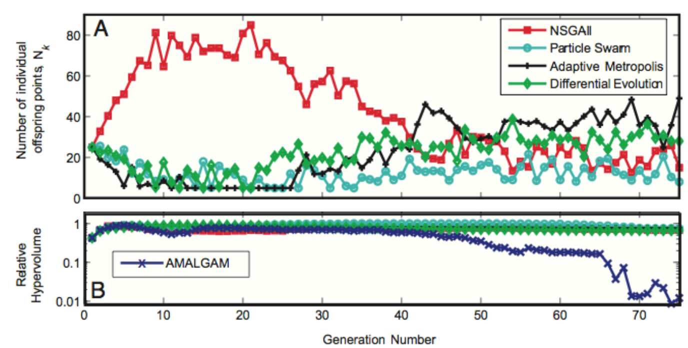
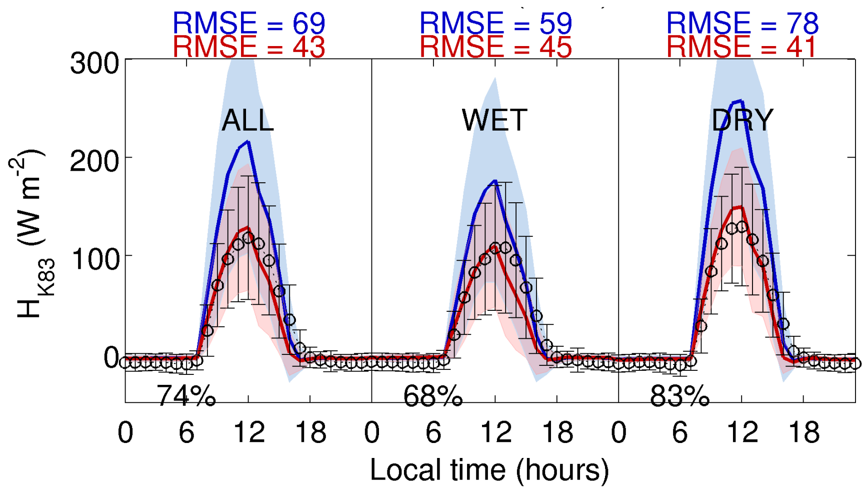


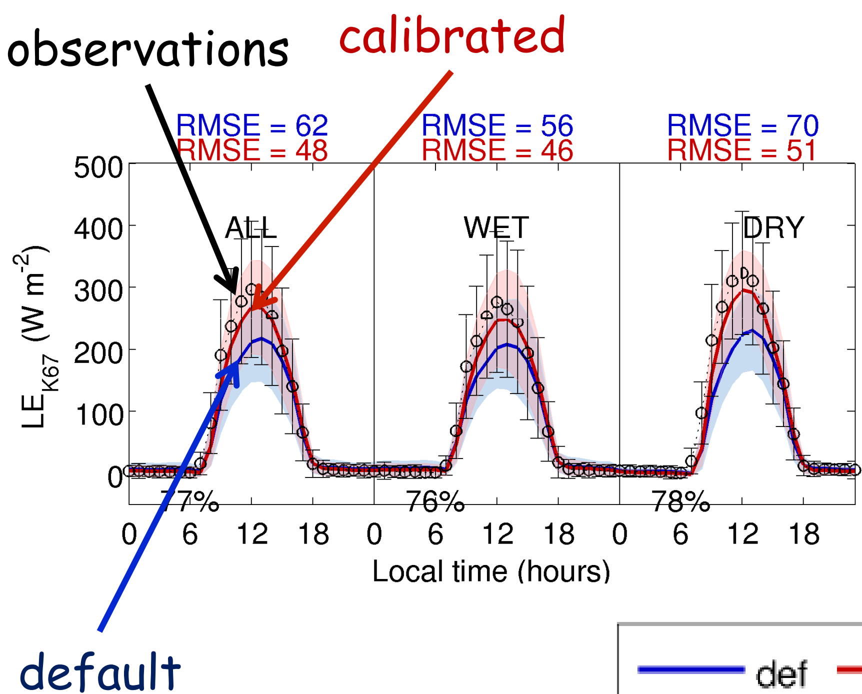
Fig. 2. Illustration of the concept of self-adaptive offspring creation. (*A*) Evolution of the number of offspring points generated with the NSGA-II (squares), PSO (circles), AMS (+), and DE (diamonds) algorithms within AMALGAM's multimethod search as function of generation number for test problem ZDT4. (*B*) The hypervolume convergence metric for AMALGAM and each algorithm used individually. These results illustrate the utility of individual search algorithms during different stages of the optimization, and provide numerical evidence of Wolpert and Macready's "No Free Lunch" theorem, showing that it is impossible to develop a single search algorithm that will always be superior to any other algorithm (16).

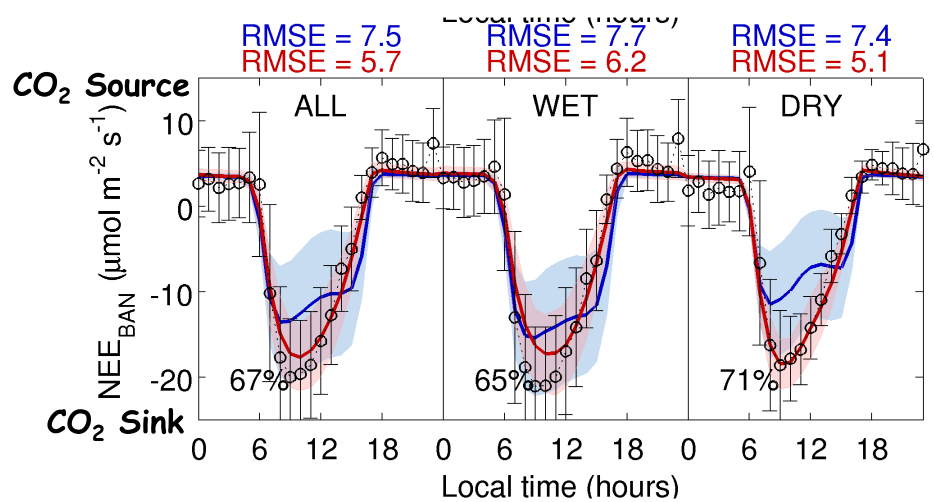
Improvement RMSE for simulated fluxes at diurnal-scale

We select a **'compromise' solution**, for which average normalized RMSE of errors in matching all three fluxes is minimum

This solution provides a balanced (**equally weighted**) reproduction of all three **fluxes**.

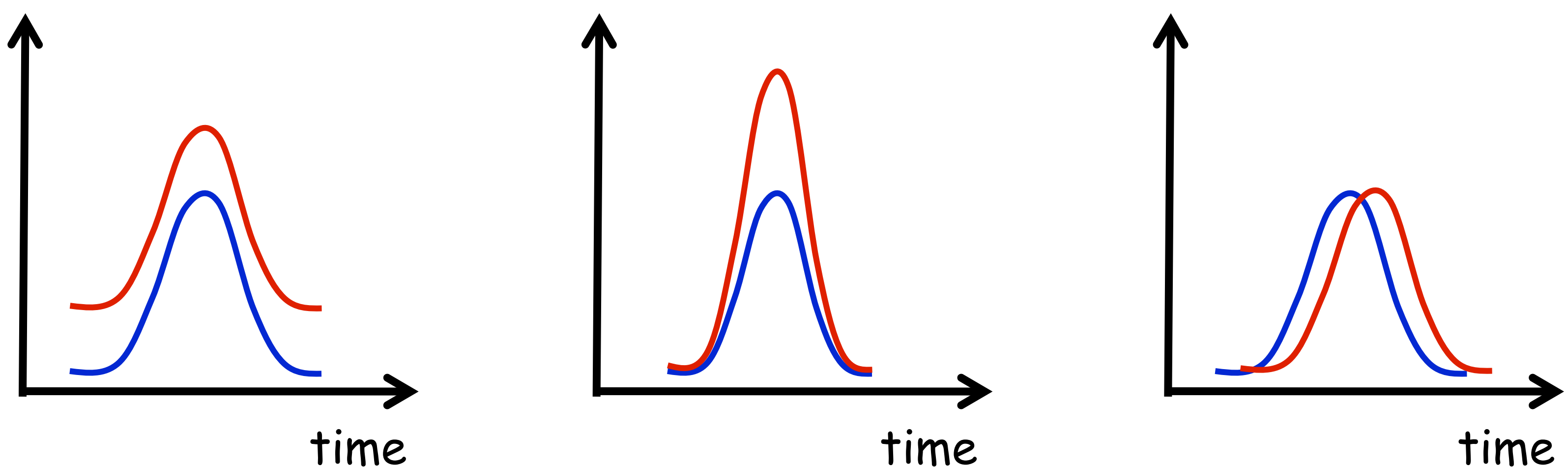






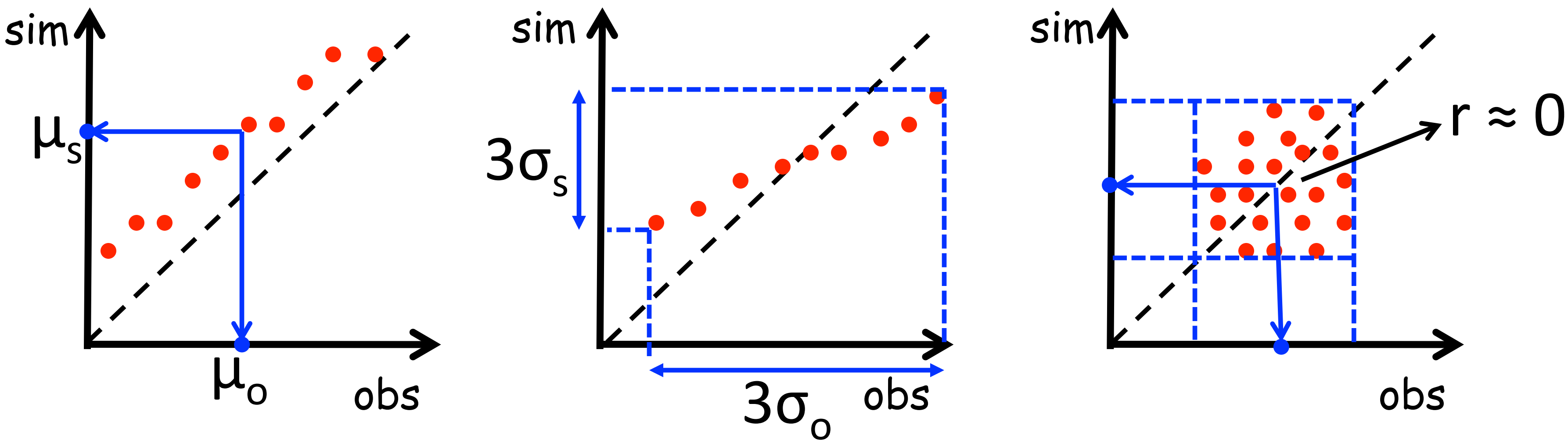
Mean-Squared-Error (MSE) Decomposition (Gupta et al. 2009)

$$MSE = \underbrace{\left(\mu_s - \mu_o\right)^2}_{\substack{\text{Error in} \\ \text{Signal Mean}}} + \underbrace{\left(\sigma_s - \sigma_o\right)^2}_{\substack{\text{Error in} \\ \text{Signal Variability}}} + \underbrace{2 \cdot \sigma_s \cdot \sigma_o \cdot \left(1 - r\right)}_{\substack{\text{Error in} \\ \text{Timing and shape}}}$$



Mean-Squared-Error (MSE) Decomposition (Gupta et al. 2009)

$$MSE = \underbrace{(\mu_s - \mu_o)^2}_{\text{Error in Signal Mean}} + \underbrace{(\sigma_s - \sigma_o)^2}_{\text{Error in Signal Variability}} + \underbrace{2 \cdot \sigma_s \cdot \sigma_o \cdot (1 - r)}_{\text{Error in Timing and shape}}$$



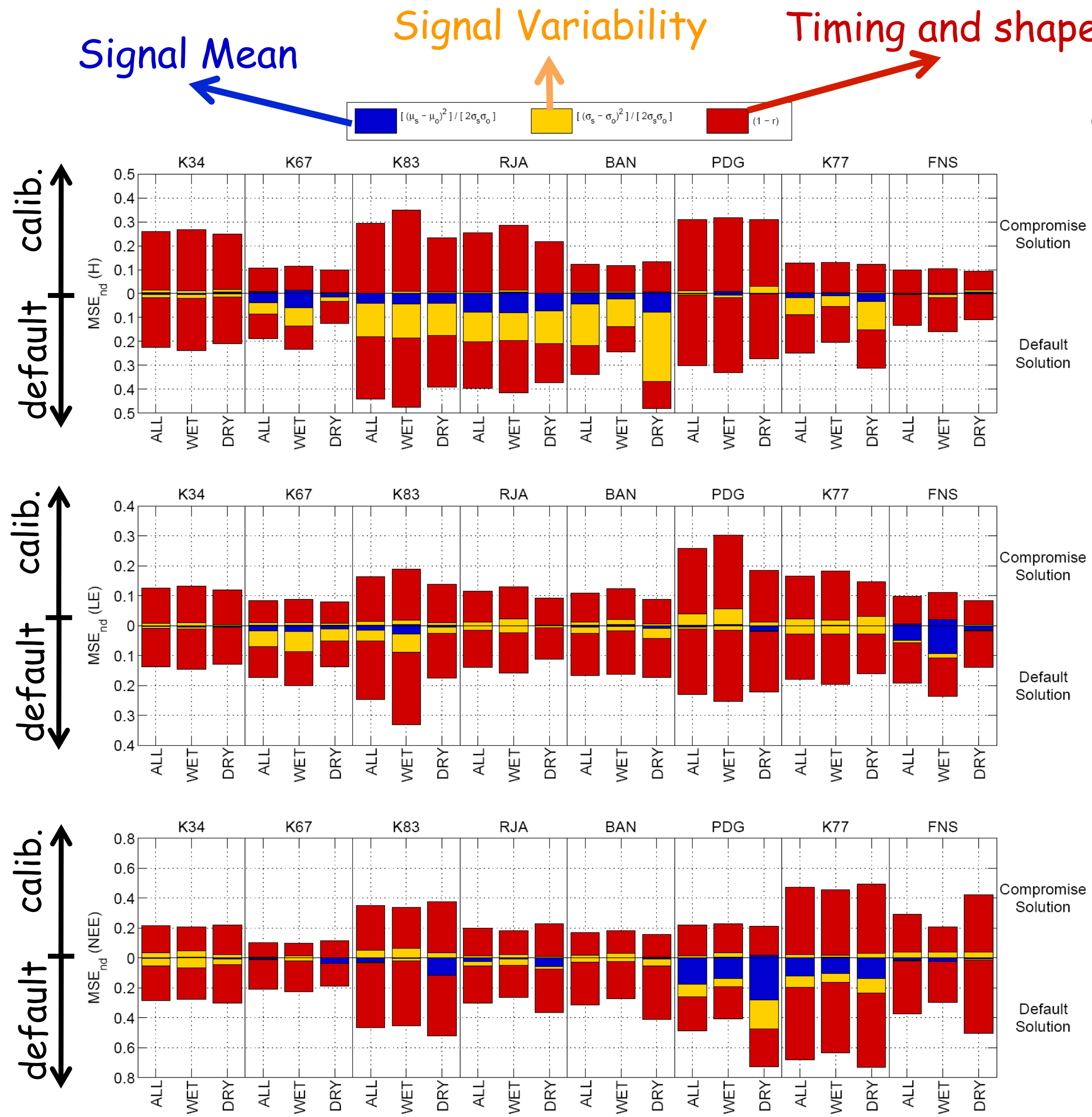
Mean-Squared-Error (MSE) Decomposition (Gupta et al. 2009)

$$MSE = \underbrace{(\mu_s - \mu_o)^2}_{\text{Error in Signal Mean}} + \underbrace{(\sigma_s - \sigma_o)^2}_{\text{Error in Signal Variability}} + \underbrace{2\cdot\sigma_s\cdot\sigma_o\cdot(1-r)}_{\text{Error in Timing and shape}}$$

Rearranging terms so individual signals are directly related to cross-correlation coefficient (r):

$$MSE_{non-\mathrm{dim}ensional} = \frac{MSE}{2\cdot\sigma_s\cdot\sigma_o} = \underbrace{\frac{(\mu_s - \mu_o)^2}{2\cdot\sigma_s\cdot\sigma_o}}_{\text{Signal Mean}} + \underbrace{\frac{(\sigma_s - \sigma_o)^2}{2\cdot\sigma_s\cdot\sigma_o}}_{\text{Signal Variability}} + \underbrace{(1-r)}_{\text{Timing and shape}}$$

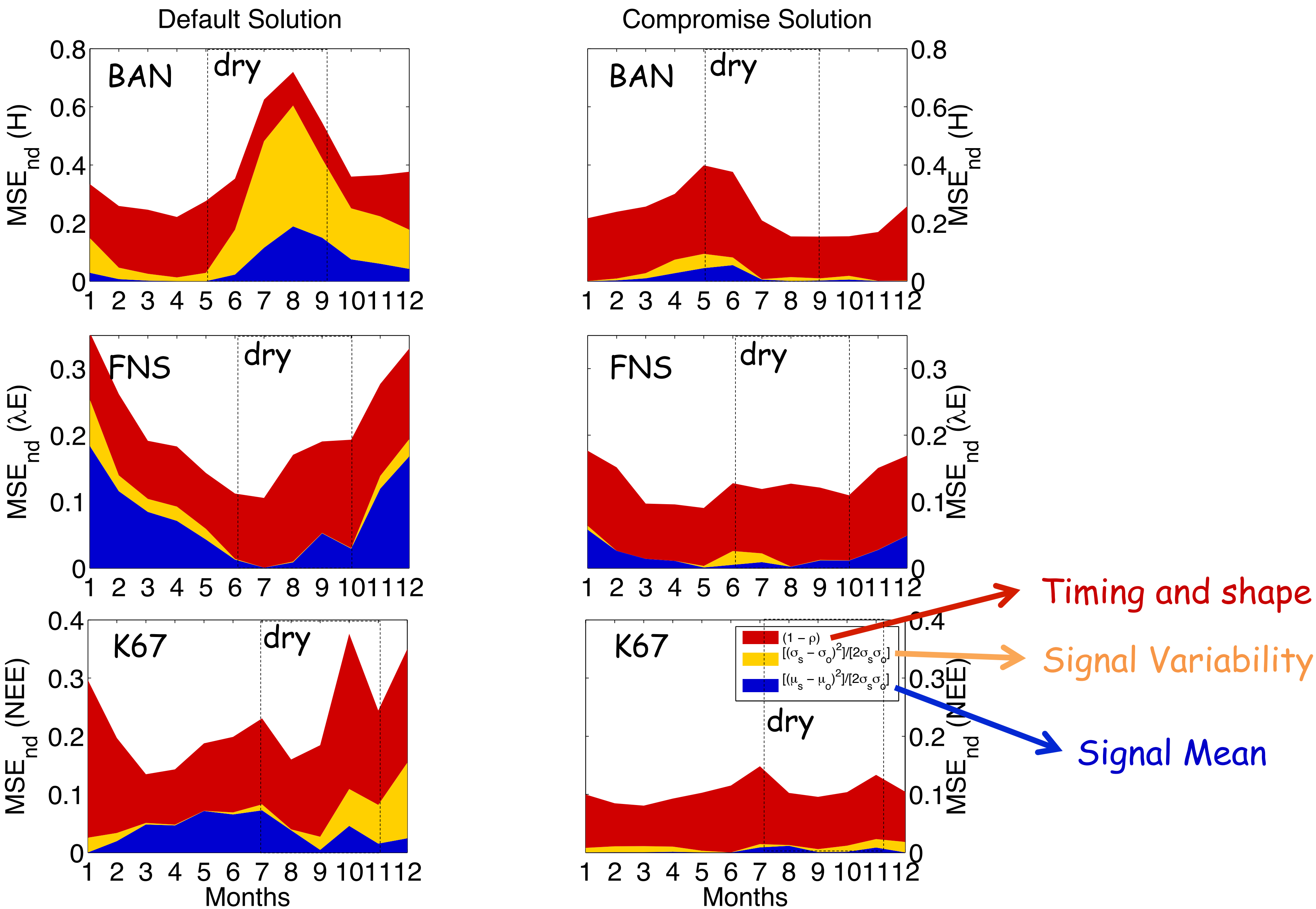
MSE Decomposition: Reduced error in signal mean and signal variability



Percent contribution to overall uncertainty from individual components

2%	3%	95%
9%	23%	68%
2%	9%	89%
6%	12%	82%
1%	10%	90%
11%	9%	80%

MSE Decomposition: Reduced error in signal mean and signal variability



Summary

1. ***New parameter sensitivity analysis approach:*** Variance-based multi-criteria extension of the **Sobol method** - accounts for full **multi-output** nature of LSMs.

2. **Method** is **model-independent:** Can be applied with other sensitivity analysis procedures (when a sensitivity index is computed for two or more functions).

3. **Substantial improvement** in *Energy*, *Water* and *CO_2 flux* simulations after calibration (10-30% reduction of RMSE).

4. **Magnitude of uncertainties** varies **seasonally**.

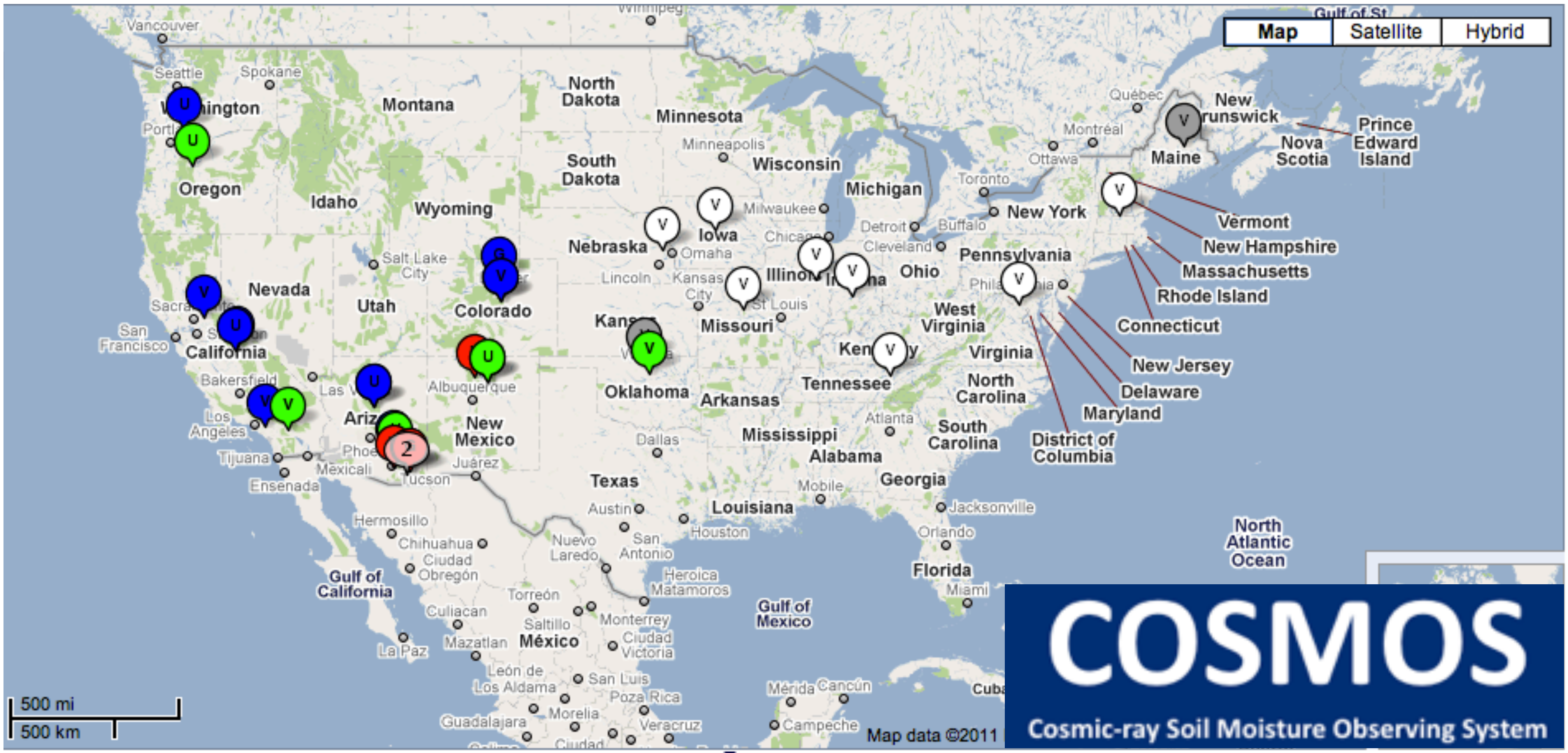
5. **Model performance is improved:**

   - Uncertainties in signal mean and variability components are reduced
   - Timing/shape is little affected (system dynamics).

6. **MSE decomposition** approach is **general**. Can also be used to **compare performance of different models** (or model versions) against a reference (e.g., *in situ* observations).

These ideas will be used to analyze CLM4 and DA methods with COSMOS-derived soil moisture (600m diameter footprint, 15 to 80 cm depth) across sites in the USA

http://cosmos.hwr.arizona.edu/



Soil Moisture (V=volumetric, G=gravimetric, U=uncalibrated)

0 - 05% | 05 - 15% | 15 - 25% | 25 - 35% | > 35% | mixed

rosolem@email.arizona.edu