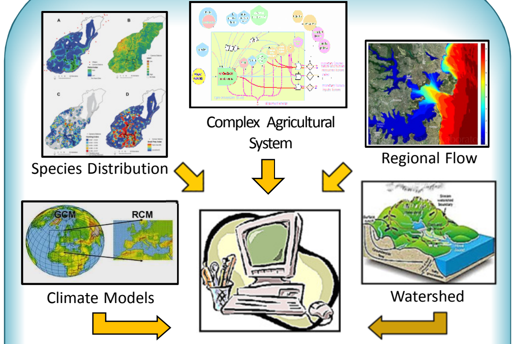
# New Multi-Criteria Parameter Screening Sampling Strategy For High Dimensional Models

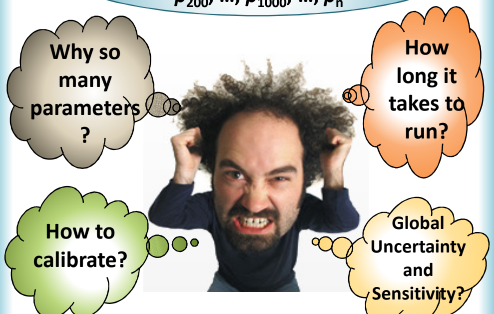
**Yogesh P. Khare***, Christopher J. Martinez, Rafael Muñoz-Carpena, and Robert Rooney
Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, 32611
Contact: khareyogesh1@ufl.edu, (352) 392 1864 ext. 290

## Background



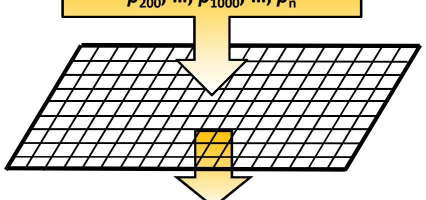

**Parameters**

$p_1, p_2, p_3, p_4, p_5, p_6, \ldots, p_{100}, \ldots, p_{200}, \ldots, p_{1000}, \ldots, p_n$




**AT LOW COMPUTATIONAL COST, CAN WE SCREEN IMPORTANT MODEL PARAMETERS???**

$p_1, p_2, p_3, p_4, p_5, p_6, \ldots, p_{100}, \ldots, p_{200}, \ldots, p_{1000}, \ldots, p_n$



$p_1, p_5, p_{10}, p_{21}, p_{67}, p_{99}$

## Motivation

**Ideal sampling strategy should result in**

- **Effective tracing of parameter space**→ evenly and widely spread points
- **Computational efficiency** → low CPU time for sample generation
- **Accurate parameter screening**

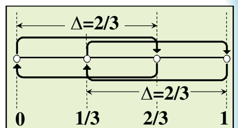
**Alternate sampling methods**

- Over 5 sampling strategies exists
- Optimized Trajectories **[OT]** (Campolongo et al. 2007)
- Modified Optimized Trajectories **[MOT]** (Ruano et al. 2012)
- Sampling for Uniformity **[SU]** (Khare et al. 2014)

## Screening Method

**Method of Elementary Effects (Morris 1991)**

- q-level, k-dimensional parameter hyperspace to generate r trajectories
- (k+1) points per trajectory, and two consecutive points differ exactly in one parameter coordinate




Schematic of q = 4 levels and complimentary coordinates/ jump size for an arbitrary parameter in a unit parameter space

$$EE_i = \frac{[y(p_1, p_2 \ldots p_i, p_{i+1}, \ldots p_k) - y(p_1, p_2 \ldots p_i + \Delta, p_{i+1}, \ldots p_k)]}{\Delta}$$

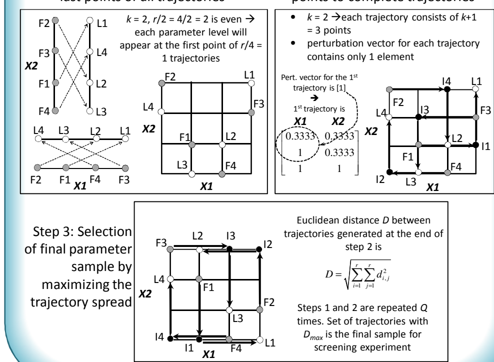
$y$ = model, $EE_i$ = elementary effect associated with $i^{th}$ parameter, $k$ = number of parameters, $\Delta$ = jump size = $q/[2(q-1)]$

$$\mu_i^* = \frac{1}{r}\sum_j^r |EE_i|_j$$

$\mu_i^*$ → total effect sensitivity

## SU : New Strategy

**UNIFORMITY and SPREAD**




## Experiments

**Sampling Strategies**

- OT, MOT, and SU (with Q = 300)

**Other Characteristics**

- k = {15, 20, 35, 50, 80, 100}, r = {10}, q = {4}
- 5 standard test functions [B, G, G*, M and O], Each combination simulated 500 times (UF HPC)

**Evaluation Criteria**

- **Trajectory spread** → Euclidean distance
- **Uniformity** → $\chi^2$ goodness of fit test
- **CPU time**
- **Screening efficiency** → g and R skill scores

## Results

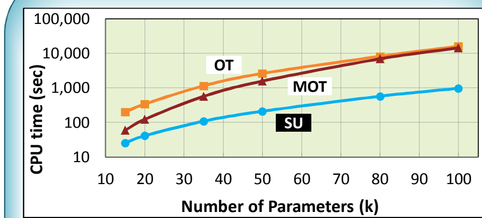



**Figure 1: CPU time required for a single sample generation by sampling strategies**

**Table 1: The average number of parameter failing the $\chi^2$ test for discrete uniform distribution at α = 10%**

| k | OT | MOT | SU |
|---|---|---|---|
| 15 | 0.57 | 0.66 | 0.00 |
| 20 | 0.59 | 0.67 | 0.00 |
| 35 | 0.66 | 0.70 | 0.00 |
| 50 | 0.65 | 0.72 | 0.00 |
| 80 | 0.69 | 0.72 | 0.00 |
| 100 | 0.69 | 0.72 | 0.00 |

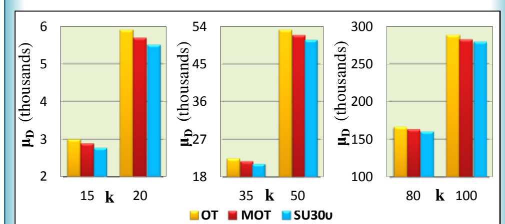



**Figure 2: Median Euclidean distance ($\mu_D$) between trajectories generated using sampling strategies**

**Table 2: Parameter screening efficiencies of sampling strategies based on two skill scores**

| OT | MOT | SU |
|---|---|---|
| avg_g across all five functions | | |
| 0.8868 | 0.8732 | 0.8905 |
| avg_R across all five functions | | |
| 0.2117 | 0.2042 | 0.2142 |

## Summary

- **Sampling for Uniformity (SU) performed better than OT and MOT across range of performance criteria**

**Future Work:**

- **Ideal oversampling size (Q) for SU**
- **Other combined criteria**
- **Application of SU to a watershed model GSA/UA**

| Strategy | Uniformity | Time | Spread | Screening | Total |
|---|---|---|---|---|---|
| OT | ** | * | *** | ** | 8 |
| MOT | * | ** | ** | * | 6 |
| SU | *** | *** | * | *** | 10 |

Computer code for SU is available for free download from http://abe.ufl.edu/carpena/software/SUMorris.shtml

**References:**

Campolongo, F, Cariboni, J., and Saltelli, A. 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software* 22: 1509-1518.

Morris, M.D. 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics,* 33(2): 161-174.

Ruano, M.V, Ribes, J., Seco, A., and Ferrer, J. 2012. An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. *Environmental Modelling and Software* 37: 103-109.

Khare, Y.P., Muñoz-Carpena, R., Rooney, R., and Martinez, C. J. A multi-criteria trajectory-based parameter sampling strategy for the screening method of elementary effects. Submitted to *Environmental Modelling and Software* (ENVSOFT-D-14-00144)