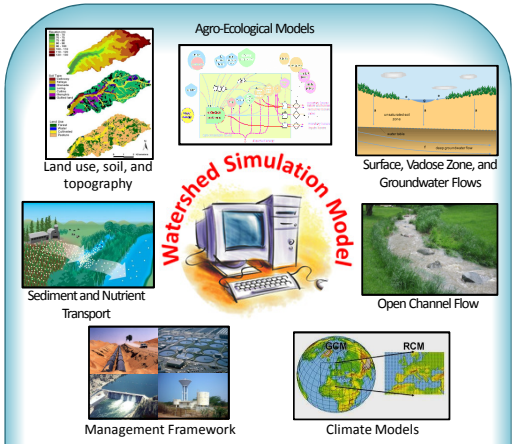
Quantifying the Efficiency of the Morris Method for High Dimensional Hydrologic Modeling Applications

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

Background



Sensitivity and Uncertainty Analysis

- Identification of important model factors
- Research prioritization
- Estimation of output uncertainty
- Reliable model applications
- Model calibration and validation

SA TECHNIQUES

Global

- Robust but computationally demanding

eFAST – N ≈ M k
Sobol' – N = M(2k+2)
Saltelli-Sobol' – N = M(k+2)

→ 50000-100000 runs for k = 100

Local

- Low cost but unreliable

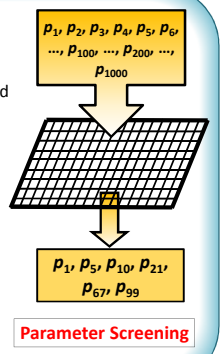
N = total model runs; M = model runs per factor, usually set to 500-1000; k = number of factors

Motivation

- Good correlation between Morris and GSA sensitivity indices
- Proposal to replace variance based GSA with Morris
- Need for efficient sampling strategies for the Morris method

GOALS

- Compare existing (OT, MOT) and newly developed (SU) sampling strategies for the Morris method
- Quantify efficiency of the Morris method



Experiments

Sampling Strategies

- **O**ptimized **T**rajectories **[OT], M**odified **OT [MOT]** (Campolongo et al., 2007; Ruano et al., 2012) → Spread
- **S**ampling for **U**niformity **[SU]** (Khare et al., 2014) → Spread and Uniformity

Other Characteristics

- k = {15, 20, 35, 50, 80, 100}, r = {10}, q = {4}
- 5 standard test functions [B, G, G*, M and O]

Evaluation Criteria

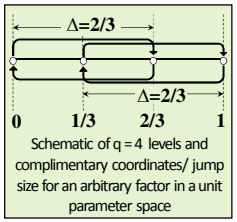
- **CPU time**
- **Uniformity** → χ^2 goodness of fit test
- **Trajectory spread** → Euclidean distance
- **Screening efficiency** → g and R skill scores

$$g = \frac{(TI \cap MI)}{TI}, R = \frac{(TI \cap MI)_{Rank}}{TI}$$

TI → set of true important parameter
MI → set of important parameters based on the Morris method

Screening Method of Morris (Morris, 1991)

- q-level, k-dimensional hyperspace to generate r trajectories
- (k+1) points per trajectory, and two consecutive points differ exactly in one coordinate
- Total runs N = r (k+1); with r = 10 and k = 100, N = 1010



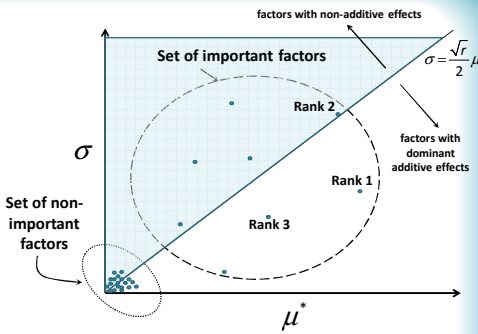
Schematic of q = 4 levels and complimentary coordinates/ jump size for an arbitrary factor in a unit parameter space

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

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

$$\mu_i^* = \frac{1}{r}\sum_j^r |EE_i|_j \qquad \sigma_i = std(EE_i)$$

μ_i^* → Overall influence
σ_i → Non-additive effects



Schematic showing segregation of model factors into important and non-important ones

Summary

- **SU (newly developed strategy) performed better than OT and MOT across range of performance criteria**
- **g = 89% implies that the Morris method is useful for input screening . However, R ~ 21% indicates that it is qualitative is nature and shall not be used as an alternative to variance based GSA.**

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

Results

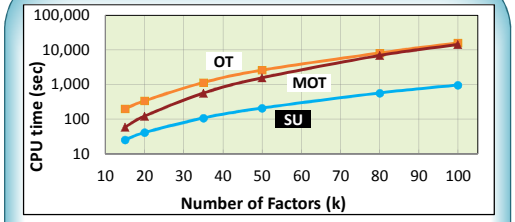


Figure 1: CPU time required for a single sample generation by sampling strategies

Table 1: The average number of factors failing the χ^2 test for generated distributions 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**

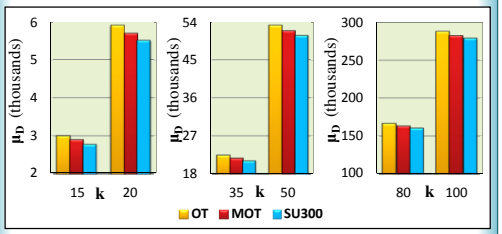


Figure 2: Median Euclidean distance (μ_D) between trajectories generated using sampling strategies

Table 2: 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**

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)