

Metamodel based Global sensitivity analysis

The variance-based method of Sobol' sensitivity indices (SI) is popular due to its easiness of interpretation. Computation of Sobol' indices using direct Monte Carlo method requires a large number of function evaluations to achieve convergence. Metamodel based methods proved to be more efficient especially for complex practical problems. In this work we used two different metamodeling techniques, namely Quasi Random Sampling-High Dimensional Model Representation (QRS-HDMR) and Bayesian Sparse Polynomial Chaos Expansion (BSPCE) methods implemented in SobolGSA software to compute Sobol SI. In many areas practitioners are also interested in minimizing/maximizing quantities of interests. Developed approach was applied to Membrane Enhanced Peptide Synthesis model (MEPS).

A unified approach to Global Sensitivity Analysis (GSA) and Global Optimization

SobolGSA is an integrating GSA and global optimization software. Built metamodel(s) can be used for both GSA and Global Optimization. Interpretation of inputs variables and objectives are different as shown in this table:

$$Y = f(\vec{X}), \vec{X} = (X_1, X_2, \dots, X_n) \in \Omega$$

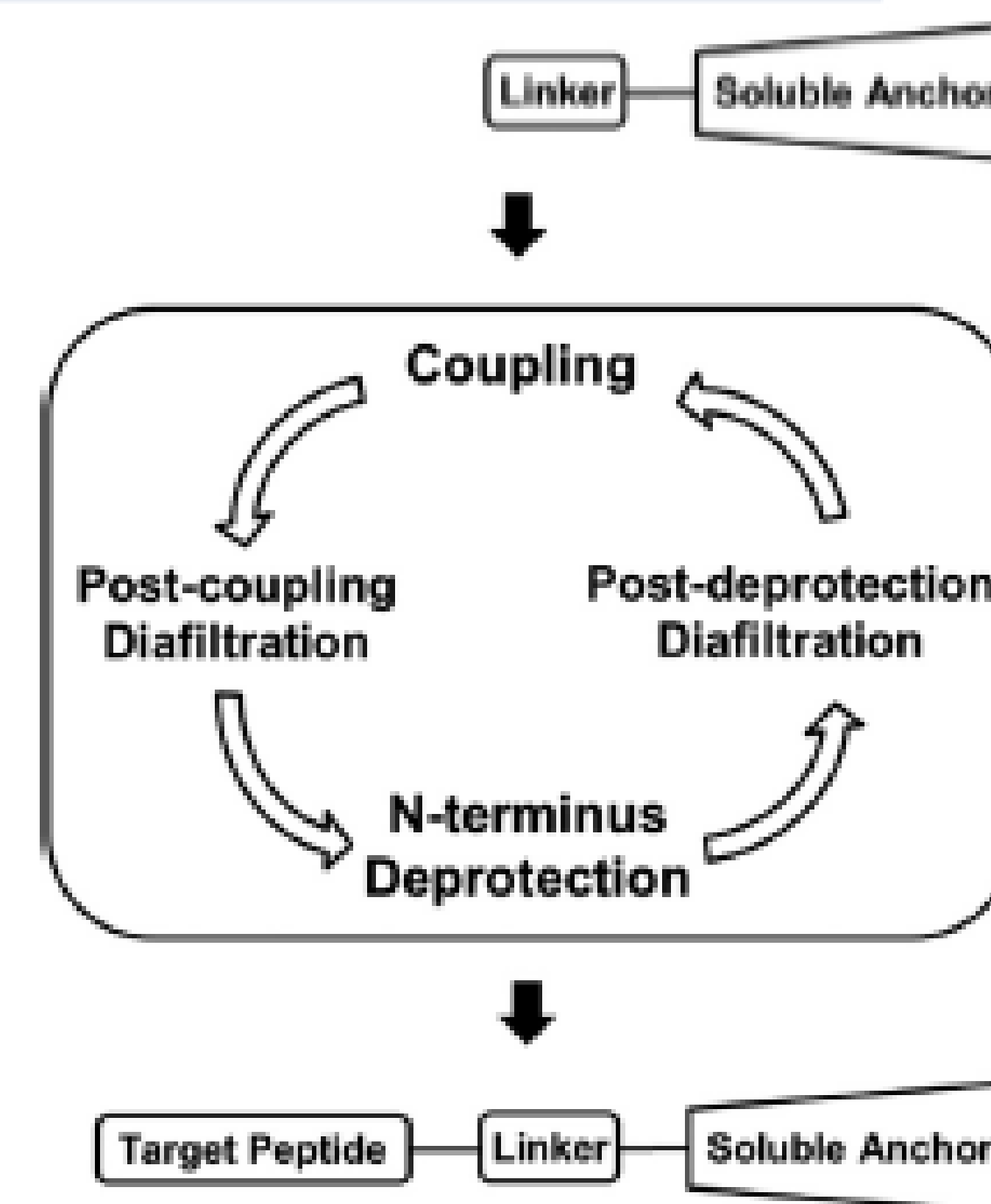
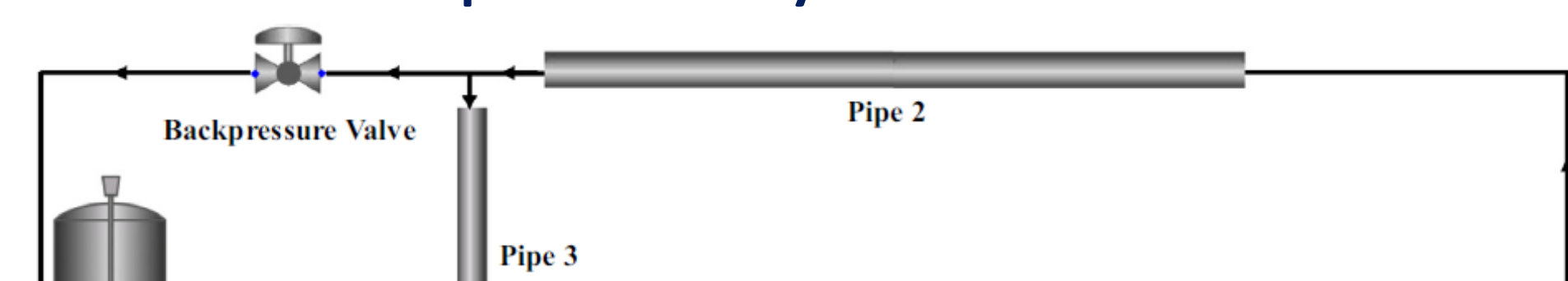
Global Sensitivity Analysis	Global Optimization
Interpretation (random setting): X is a vector of uncertain variables	Interpretation (deterministic setting): X is a vector of inputs
Objective: study the effect of input uncertainty on the variability in the model output. Sobol' indices – normalised partial variances of the total variance of Y	Objective: Search the global minima of f and the set of global minimizers X^*

Membrane-enhanced Peptide Synthesis (MEPS)

Peptides are comprised of short chains of amino acids linked together by peptide bonds. Peptides are short fragments of proteins with the main difference between a peptide and protein being size. Peptides are used in a variety of applications such as drug therapies, vaccines and cosmetics. Coupled to DNA, peptide cargo can be transported easily into cells (DNA-peptide conjugates).

MEPS is a new technology that efficiently combines organic solvent nanofiltration using membranes with solution phase peptide synthesis.

It involves repeated cycles of reaction and diafiltration

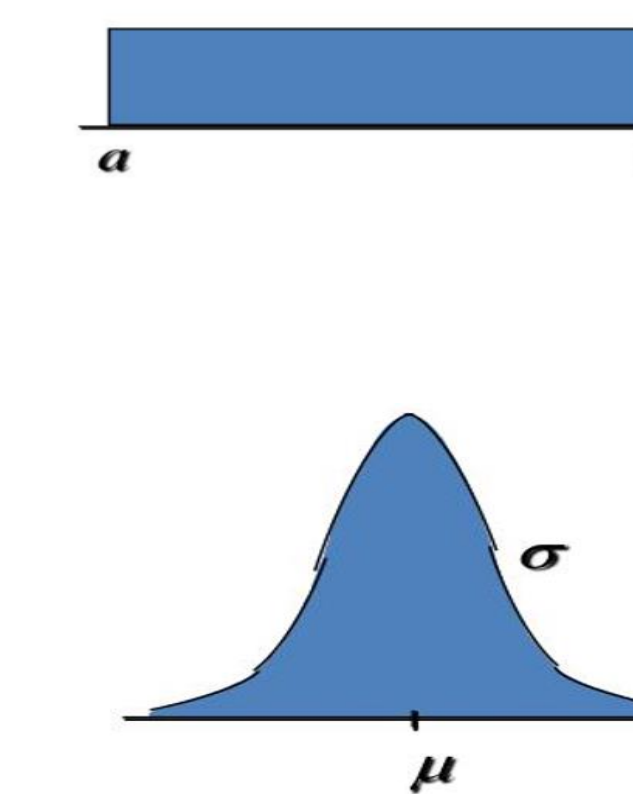


SobolGSA Software

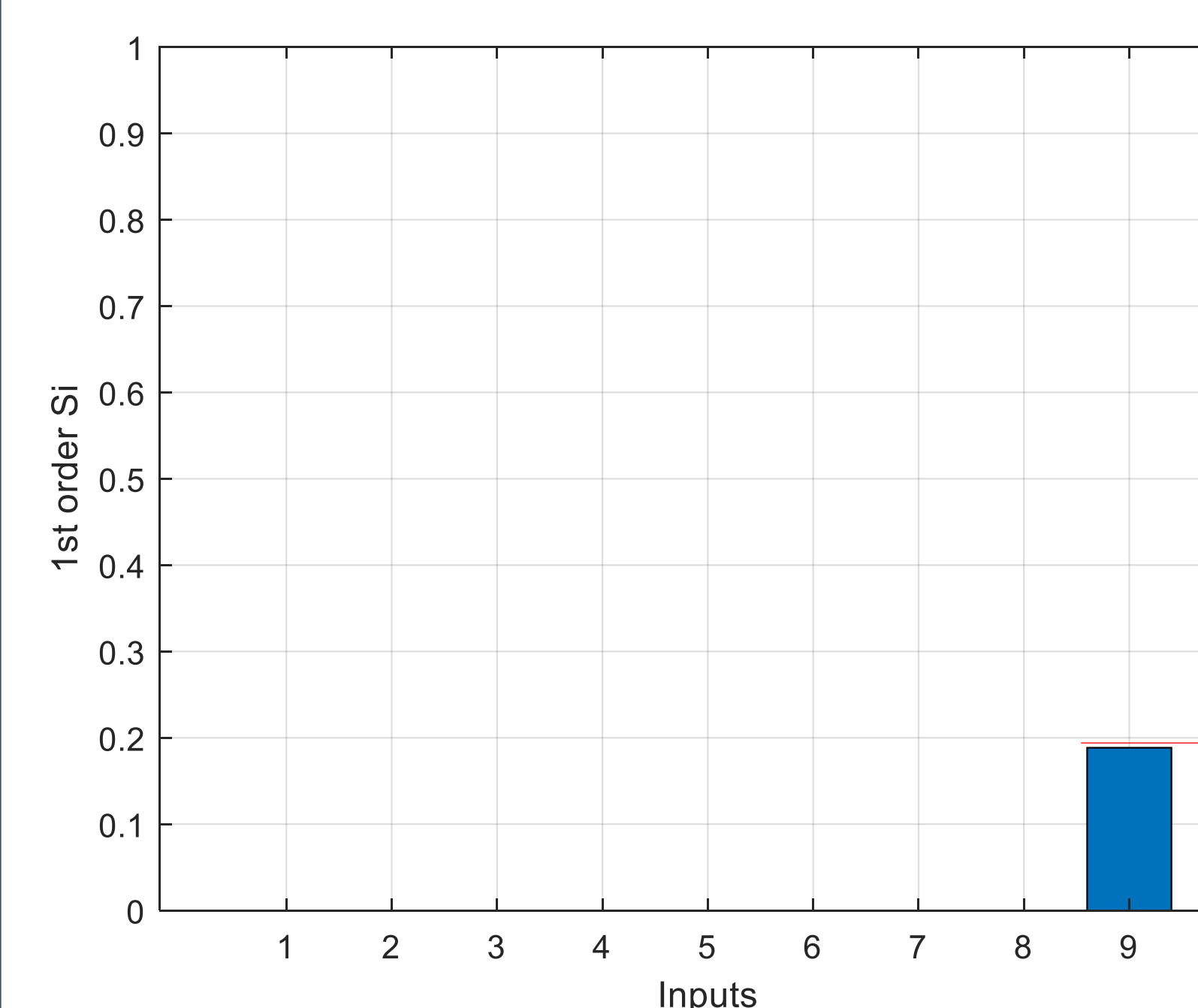
- Easy to use GUI
- Extensive set of cases studies, documentation and manuals
- Can be linked to MATLAB, gPROMS, Python and other packages
- Can handle explicitly defined models and models given as input-output data (black box models)
- Several GSA measures (Sobol' indices, derivative based measures (DGSM), FAST, Morris)
- Different metamodeling techniques (QRS-HDMR, BSPCE, RBF)
- Produced metamodels given as self-contained C# / MATLAB/ Python files
- Global Optimization with SobolOpt

Types of uncertainty

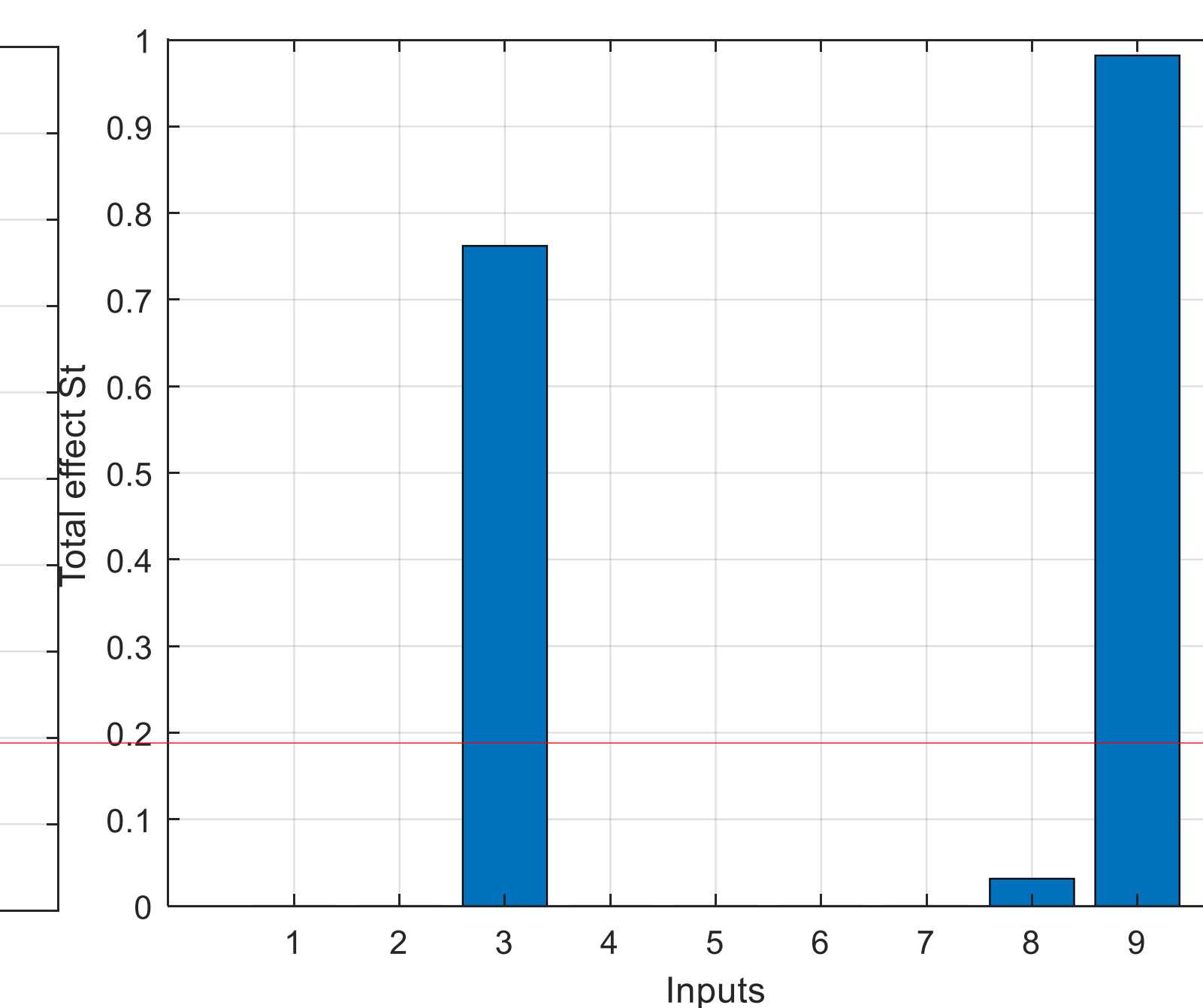
uniform and normal uncertainty distributions for nine process parameters including membrane properties and reaction rate constants. Model Output - Total yield



Permeance	:= 2.71E-7;	# Unit: m3 m-2 s-1 bar-1
Rejection_Anchor	:= 0.9965 ;	# Dimensionless
Rate_Constant_Coupling(i)	:=0.257;	# Time in second
Rate_Constant_deFmoc(i)	:=0.00184;	# Time in second
Rate_Constant_Side_Reaction (i)	:=0.257;	# Time in second
Rejection_AA(1:Number_AA)	:= 0.33;	# Dimensionless
Rejection_Impurity(1:Number_Impurity)	:= 0.9965 ;	# Dimensionless
Rejection_Product(1:Number_Product)	:= 0.9965 ;	# Dimensionless
Rejection_Piperidine	:= 0.33;	# Dimensionless



Main effect Sobol SI



Total Sobol SI

Sobol' indices were computed using SobolGSA in which the input-output tables were exported from gPROMS. Parameter distributions affect the results of GSA (Table 1). However, the main driver of uncertainty in the model outputs remains the same for both distributions. There are strong interactions between parameters for the all distributions

Parameters	$S_i(U)$	$S_i(N)$	$S_i^{tot}(U)$	$S_i^{tot}(N)$
Rejection Piperidine	0.0	0.0	0.06	0.75
Rejection Amino	0.0	0.0	0.32	0.0
Rejection Product	0.6	0.2	0.96	0.98

Optimization. Metamodel built to perform GSA was used to maximize total peptide yield using two different methods for finding global maximum.

Conclusions. Although there is some discrepancy between predictions based on SobolGSA metamodel and validation results obtained using the full model with the maximizer values in gPROMS, proposed approach gives a good guidance with regards to further directions of modelling and experimental work.

Acknowledgements. We acknowledge the financial support of Eli Lilly and Company.