Advances and challenges in global sensitivity analysis

- Pierre Gremaud (NC State)

joint work with:

- Joey Hart (Sandia NL)
- Tim David (U. of Canterbury, New Zealand)
- Alen Alexanderian, Mike Merritt,
  Robin Morillo, Ralph Smith (NC State)

August 1, 2019

but first: Happy 728th Birthday, Switzerland (Aug. 1, 1291)

what I usually do on Aug. 1

setup

model: $$\underbrace{q}_{\text{response}} = g(\underbrace{\theta_1, \dots, \theta_p}_{\text{parameters}})$$

typically

- g is a computer code
- parameters are uncertain
- p is large

p large?

- in stat speak: p and n (sample size/data) matters
  - classic stat: p is fixed, $n \to \infty$,
    law of large numbers, central limit theory
  - high dim: p and n large, often with $p \gg n$
- in math speak: standard approaches do not work
  - objects can't be handled $\Rightarrow$ sampling
  - low precision arithmetics(!)
  - see the OTHER RTG: randomized numerical analysis

sensitivity analysis ($q = g(\theta)$)

we want to:

> *quantify how uncertainties in the model response can be apportioned to uncertainties in model inputs*

the larger the contribution, the more important the input

rationale for SA (inspired by Saltelli)

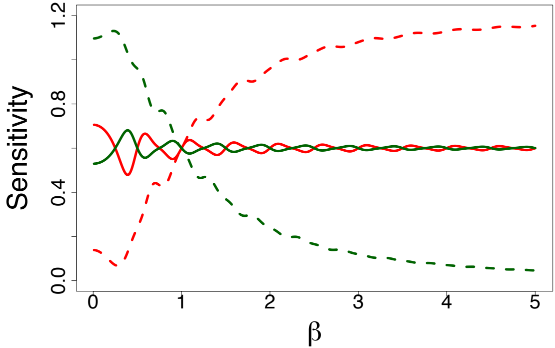
- model corroboration: is the inference robust?
- research prioritization: which factor most deserves further analysis/measurement?
- model simplification: can factors/compartments be fixed or simplified?
- model reliability: identify factors which interact and may lead to extreme values

GSA challenges

- no agreement on the meaning of important
- one SA method $\Leftrightarrow$ one definition of "importance"
- inputs can be correlated
- GSA results depend on how parameter uncertainty is modeled; robustness?
- meaning of GSA for evolution pbs; causality?
- practical considerations $\Rightarrow$ use of surrogates (often):
  - surrogate $\approx$ model $\overset{?}{\Rightarrow}$ GSA(surrogate) $\approx$ GSA(model)

importance?

- $g(\theta_1, \theta_2) = \sin^2 \beta\theta_1 \sin^2 \theta_2 \qquad \theta_i \sim U(0, 2\pi), i = 1, 2$



- dashed lines: partial derivative importance
- solid lines: total Sobol' indices
- only agree for $\beta = 1$!

GSA: lots of choices

- regression based
- variance based (Sobol' indices)
- derivative based (Morris screening, ...)
- game theoretic (Shapley values/effects)
- and others...

this talk: (mostly) variance based

variance based GSA

Ilya Sobol'

- considers θ_i's as random variables
- apportion to them their relative contribution to the variance of the response

- trivial example: $q = a\,\theta_1 + b\,\theta_2$, $\theta_i \sim N(0, \sigma_i^2)$, $a, b > 0$
- $q \sim N(0, \sigma_q^2)$ with $\sigma_q^2 = a^2\sigma_1^2 + b^2\sigma_2^2$

$$\Rightarrow 1 = \underbrace{\frac{a^2\sigma_1^2}{a^2\sigma_1^2 + b^2\sigma_2^2}}_{S_1} + \underbrace{\frac{b^2\sigma_2^2}{a^2\sigma_1^2 + b^2\sigma_2^2}}_{S_2}$$

- note the importance of the σ_i's!

total Sobol' indices

law of total variance

$$\mathrm{var}(\mathbb{E}[q|\theta_{\sim i}]) + \mathbb{E}[\mathrm{var}(q|\theta_{\sim i})] = \mathrm{var}(q)$$

and thus

$$\underbrace{\mathrm{var}(q) - \mathrm{var}(\mathbb{E}[q|\theta_{\sim i}])}_{\text{remaining variance if } \theta_{\sim i} \text{ were known}} = \mathbb{E}[\mathrm{var}(q|\theta_{\sim i})]$$

- total index: $T_i = \frac{\mathbb{E}[\mathrm{var}(q|\theta_{\sim i})]}{\mathrm{var}(q)} = 1 - \frac{\mathrm{var}(\mathbb{E}[q|\theta_{\sim i}])}{\mathrm{var}(q)}$

$T_i = 0 \Leftrightarrow \theta_i$ non-important

$\Leftarrow$:

$$\theta_i \text{ non-import.} \Rightarrow \mathrm{var}(q|\theta_{\sim i}) = 0 \Rightarrow \mathbb{E}[\mathrm{var}(q|\theta_{\sim i})] = 0 \Rightarrow T_i = 0$$

$\Rightarrow$:

$$T_i = 0 \Rightarrow \mathbb{E}[\mathrm{var}(q|\theta_{\sim i})] = 0 \underset{\mathrm{var} \geq 0}{\Rightarrow} \mathrm{var}(q|\theta_{\sim i}) = 0 \Rightarrow \theta_i \text{ not import.}$$

Unimportance is important! (Art Owen)

- allows focus on key inputs
- potential for faster codes

ANOVA (Reader's Digest version)

- assume θ_i, iid, $\theta_i \sim U(0,1)$
- split $\theta = (\theta_i, \theta_{\sim i})$ and decompose g as

$$g(\theta) = g_0 + g_1(\theta_i) + g_2(\theta_{\sim i}) + g_{12}(\theta_i, \theta_{\sim i})$$

  where
  - $g_0 = \int g(\theta)\, d\theta$,
  - $g_1(\theta_i) = \int (g - g_0)\, d\theta_{\sim i}, \quad g_2(\theta_{\sim i}) = \int (g - g_0)\, d\theta_i$
  - $g_{12} =$ remainder
- above functions have zero average $\Rightarrow \perp \Rightarrow$

$$\begin{aligned} \mathrm{var}(q) &= \int (g(\theta) - g_0)^2\, d\theta = \int g(\theta)^2\, d\theta - g_0^2 \\ &= \underbrace{\int g_1^2\, d\theta}_{\mathrm{var}(g_1)} + \underbrace{\int g_2^2\, d\theta}_{\mathrm{var}(g_2)} + \underbrace{\int g_{12}^2\, d\theta}_{\mathrm{var}(g_{12})} \end{aligned}$$

another way to look at things

equivalent definition:

$$\underbrace{T_i = \frac{\text{var}_i}{\text{var}(q)}}_{\text{total index}}; \qquad \underbrace{S_i = \frac{\text{var}(g_i)}{\text{var}(q)}}_{\text{1st order index}}$$

where

$$\text{var}_i = \text{var}(g_1) + \text{var}(g_{12}) = \text{ total variance corresponding to } \theta_i$$

exercise:

$$\text{var}_i = \frac{1}{2} \int\int (\overbrace{g(\theta) - g(\theta')}^{\frac{\partial g}{\partial \theta_i}(\hat{\theta})(\theta_i - \theta_i')})^2 \, d\theta \, d\theta_i'$$

where $\theta' = (\theta_1, \ldots, \theta_{i-1}, \theta_i', \theta_{i+1}, \ldots, \theta_p)$.

independent variables: summary

notation: $u \subset \{1, \ldots, p\}$

- $\displaystyle\sum_{k=1}^{p} \sum_{|u|=k} S_u = 1$ **conservation**
- $\forall u, 0 \leq S_u \leq 1$ **boundedness**
- $\displaystyle\min_{k \in u} T_k \leq T_u \leq \sum_{k \in u} T_k$
- S_u close to 1 $\Rightarrow \theta_u$ is important
- T_u close to 0 $\Leftrightarrow \theta_u$ is non-important

dependent variables?[1]

$$S_u = \frac{\mathrm{cov}(g_u, g)}{\mathrm{var}(q)}, \quad T_u = \sum_{v \cap u \neq \emptyset} S_v$$

[1] Chastaing et al., J. Stat. Comput. Sim., 2012

however...

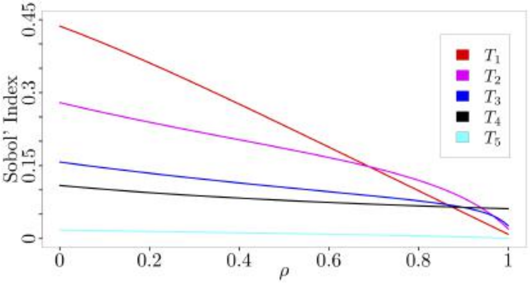
- $\sum_{k=1}^{p} \sum_{|u|=k} S_u = 1$ **conservation** YES
- $\forall u, 0 \leq S_u \leq 1$ **boundedness** NO
- $\min_{k \in u} T_k \leq T_u \leq \sum_{k \in u} T_k$ NO
- S_u close to 1 $\Rightarrow \theta_u$ is important NO
- T_u close to 0 $\Leftrightarrow \theta_u$ is non-important NO

both interpretation and computations are challenging!

illustration

$$g(\theta) = 20\theta_1 + 16\theta_2 + 12\theta_3 + 10\theta_4 + 4\theta_5, \qquad \text{with } \theta \sim N(\mu, \Sigma)$$

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 1 & .5\rho & .5\rho & 0 & .8\rho \\ .5\rho & 1 & 0 & 0 & 0 \\ .5\rho & 0 & 1 & 0 & .3\rho \\ 0 & 0 & 0 & 1 & 0 \\ .8\rho & 0 & .3\rho & 0 & 1 \end{bmatrix}, \qquad 0 \leq \rho \leq 1.$$



for $\rho = 1$:

$T_1 \quad = \quad 0.0087,$

$T_2 \quad = \quad 0.0196,$

$T_{1,2} = 0.4228$

approximation theoretic perspective

- g square integrable; $g_0 = \mathbb{E}[g(\theta)]$
- a key question:

how accurately can $g(\theta) - g_0$ be approximated WITHOUT θ_u?
or
what is the error associated with the approximation

$$g(\theta) - g_0 \approx \mathcal{P}_{\sim u} g(\theta_{\sim u})?$$

$\mathcal{P}_{\sim u} g$: optimal L^2 approximation which does not depend on θ_u

approximation theoretic perspective

Theorem (Hart, G., 2018)

$$T_u = \frac{\|(g(\theta) - g_0) - \mathcal{P}_{\sim u} g(\theta_{\sim u})\|_2^2}{\|g(\theta) - g_0\|_2^2}.$$

dimension reduction

ways to get a "smaller" model:

- project onto subspace of fcts of fewer var. (expensive)
- train a surrogate model using fewer variables
- fix some of the θ_i's to "nominal" values (common)

- above result helps: natural dimension reduction for Sobol is the above projection:

$$T_u = 1 - \frac{\|\mathcal{P}_{\sim u} g\|_2^2}{\|g - g_0\|_2^2}$$

- ongoing work: efficient ways to compute $\mathcal{P}_{\sim u} g$ "in passing"

dimension reduction II

- $\theta = (\theta_u, \theta_{\sim u})$; assume T_u small
- we want:

$$g(\theta) \approx g(h(\theta_{\sim u}), \theta_{\sim u}) \qquad \text{for instance: } h(\theta_{\sim u}) = \mathbb{E}[\theta_u]$$

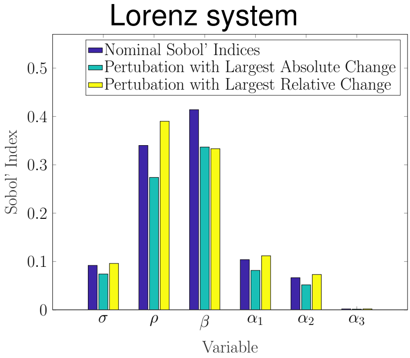
key quantity: $\delta_u = \dfrac{\|g(\theta) - g(h(\theta_{\sim u}), \theta_{\sim u})\|_2^2}{\|g(\theta) - g_0\|_2^2}$

Theorem (Hart, G., 2018)

$$T_u \leq \delta_u \leq 4T_u$$

robustness

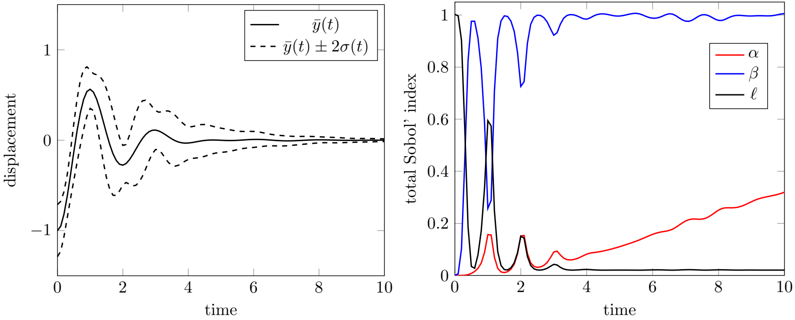
- it is possible to evaluate robustness of T_u at "no" extra cost
- involves regarding T_u as $\{\text{PDF}\} \to \{\text{Sobol indices}\}$ and taking derivatives (Fréchet differentiable) and... evaluating carefully
- limitations…



time dependent GSA

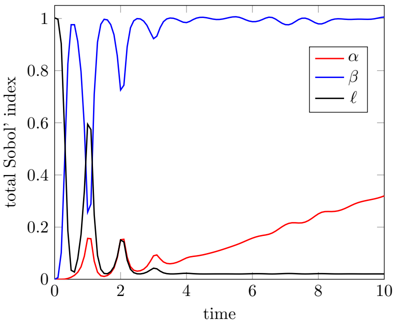
$$y'' + 2\alpha y' + (\alpha^2 + \beta^2)y = 0,$$
$$y(0) = \ell, \quad y'(0) = 0$$

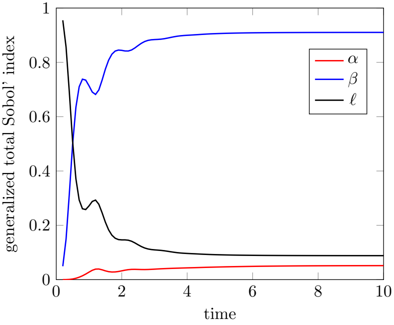
underdamped mechanical oscil: $y = \ell e^{-\alpha t}(\cos \beta t + \frac{\alpha}{\beta} \sin \beta t)$



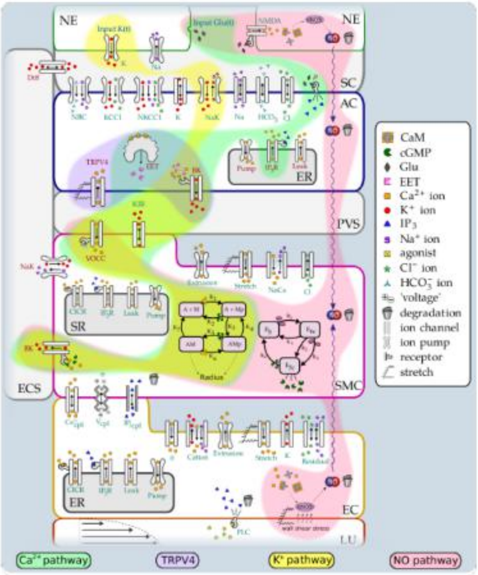
time dependent GSA II

- yardstick! $\Rightarrow$ $T_i(t)$ not very informative
- better way: $\mathbb{T}_i(t) = \dfrac{\int_0^t \text{var}_i \; d\tau}{\int_0^t \text{var}\, q \, d\tau}$. (Gamboa et al. 2014)





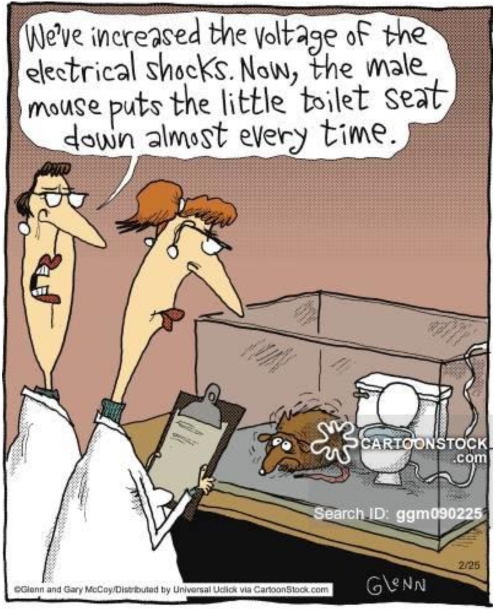
an application (with T. David and J. Hart)



picture from T. David et al.

- 4 pathways
- 67 state variables
- 160 parameters
- 4 "objects"/models
  - NE
  - AC
  - SMC, EC
  - WM
- share 13 variables (coupling)

experimental setup



Goals

- physiology: understand dominant cellular mechanisms resulting in cerebral tissue perfusion after neuronal stimulation
- diagnostics (understanding) rather than prognostics (predictions)
- mathematics: develop GSA methods for large systems of the type

$$\begin{aligned} \frac{d\mathbf{y}}{dt} &= \mathbf{f}(\theta, \mathbf{y}) + \text{ICs} \\ \text{QoI: } q &= G(\theta, \mathbf{y}(\theta)) \equiv g(\theta) \end{aligned}$$

Challenges

- large systems of ODEs: state variables $\mathbf{y} = (y_1, \ldots, y_{67})$
- large number of uncertain parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{160}) \Rightarrow$ high-dimensionality

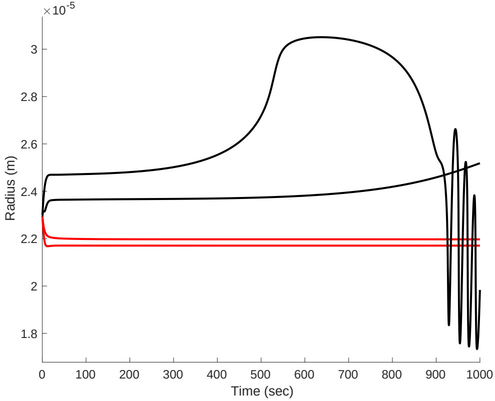
$$q = g(\theta_1, \ldots, \theta_{160})$$

- standard GSA methods are too expensive out of the box
- multiple time scales $\Rightarrow$ stiffness
- "fuzzy" goal $\Rightarrow$ several QoIs

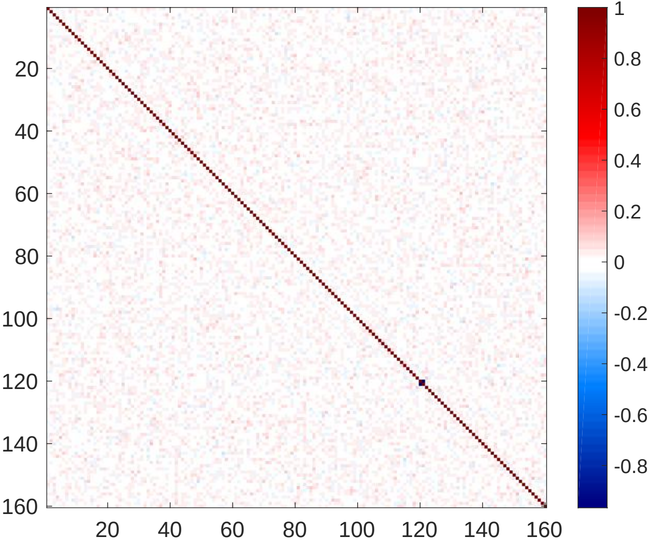
hunting for feasible points

919 initial samples (θ_k) without stimulus; of these

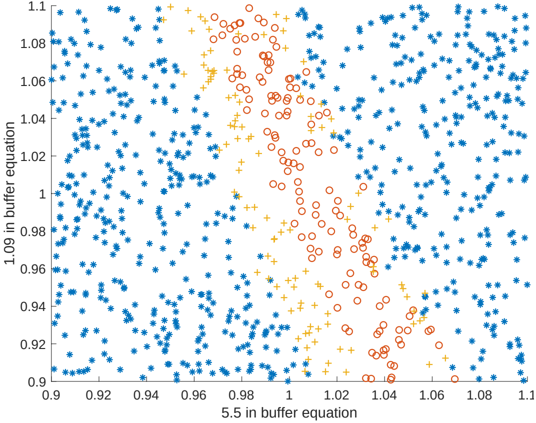
- 670 fail (the MATLAB ODE solver ode15s can't take time step small enough)
- 110 reach an **unstable** steady-state and then go nuts
- 139 reach a **stable** steady-state



Pearson ⇒ two parameters are strongly correlated



correlated pair



$*$: premature term.
$+$: unstable steady state
$\circ$: stable steady state

let's sample "harder"

updated distribution:

- $\sim$correlated pair: as before (iid. U. $\pm 10\%$)
- correlated pair: fitted Frank copula with beta marginals

$$CDF = c(F_i(\theta_i), F_j(x_j))$$

- with $F_i(x_i) = Pr(\theta_i \leq x_i) = \dfrac{B(x_i; \alpha_i, \beta_i)}{B(1, \alpha_i, \beta_j)}$ with

$$B(x; \alpha, \beta) = \int_0^x t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

- $c(u, v) = -\dfrac{1}{\mu} \log \left(1 + \dfrac{(\exp(-\mu u) - 1)(\exp(-\mu v) - 1)}{\exp(-\mu) - 1}\right)$

let's sample "harder" (II)

iterative process

- while $\sim$ tired or $\sim$ converged
  - fit above bivariate distribution to (θ_i, θ_j)
  - sample $\boldsymbol{\theta}$
  - solve
  - keep stable steady state solutions

in the end: 902 stable steady solutions

dimension reduction and GSA roadmap

for each QoI

- reduce dimension through linear regression
- fit polynomial chaos surrogate on reduced set of parameters
- compute total Sobol' indices from PCE

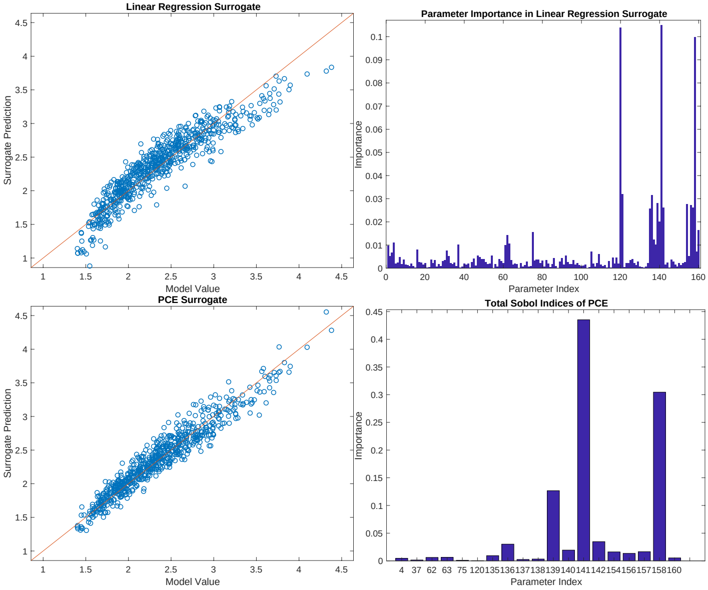
linear regression

- $QoI = g(\boldsymbol{\theta}) \approx \beta_0 + \sum_{k=1}^{160} \beta_k \theta_k$
- importance of k-th parameter

$$L_k = \frac{|\beta_k|}{\sum_{j=1}^{160} |\beta_j|}, \qquad k = 1, \dots, 160$$

- retain only θ_k's with $L_k > 0.01$ for nonlinear fitting

PCE+ Sobol’: flow



Perspectives and Conclusions

- lots of work to do in high dim approximation
- dimension reduction is key
- how to construct surrogate models
- to solve a specific problem, quantitative experts and field experts have to work together
- there may be "cultural issues" (is everyone happy with a linear model with 10 parameters?)

papers

- J.L. Hart, P. Gremaud, *Robustness of the Sobol' indices to distributional uncertainty*, Int. J. Uncertainty Quant., accepted.
- J.L. Hart, P. Gremaud, and T. David, *Global sensitivity analysis of high-dimensional neuroscience models: an example of neurovascular coupling*, Bull. Math. Biol., 2019
- J.L. Hart, P. Gremaud, *Robustness of the Sobol' indices to marginal distribution uncertainty*, SIAM/ASA J. UQ, accepted.
- J.L. Hart, P. Gremaud, *An Approximation Theoretic Perspective of Sobol' Indices with Dependent Variables*, Int. J. Uncertainty Quant., 2018
- A. Alexanderian, P. Gremaud, R.C. Smith, *Variance-based sensitivity analysis fo time-dependent processes*, under revision
- J.L. Hart, A. Alexanderian, P. Gremaud, *Efficient computation of Sobol indices for stochastic models*, SIAM J. Sc. Comp., 39 (2017), pp. A1514-A1530