

Nonlinear Modeling: Sensitivity Analysis

Introduction to Statistical Modelling

Prof. Joris Vankerschaver

Sensitivity Analysis

Why sensitivity analysis?

- Verify what *sources of uncertainty* contribute most to variance (uncertainty) of model output.
- Sources of uncertainty in model can be
 - Model parameters, initial conditions, inputs
 - Model structure
- Better understand changes in model predictions due to the above

Why sensitivity analysis?

- Detect what *model parameters* contribute most to model output uncertainty
- Want to reduce model uncertainty, so best to focus on most influential parameters
- Gives idea of correlation between parameters
- Helps in choice of what parameters to estimate (in parameter estimation)

Why sensitivity analysis?

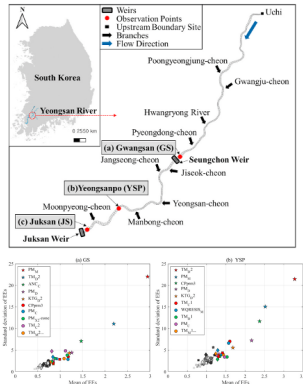
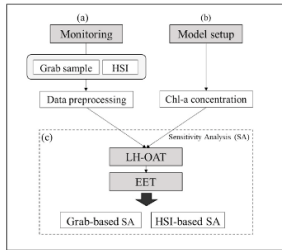
- Gives information about interesting location, time, ... to collect experimental data
- Basis for experimental design
- Gives information on insensitive model parameters
- Useful in model reduction of overparametrized models

Local vs global

- ① Local sensitivity analysis
 - Determine sensitivity at **one certain point** in parameter space
 - Not very computationally intensive
 - ② Global sensitivity analysis
 - Determine sensitivity in **delimited area** of parameter space
 - Usually gives a mean sensitivity
 - Can become extremely computationally intensive
- Each technique has advantages and disadvantages
 - Each technique gives different type of information

Examples of sensitivity analysis: water quality model

- Hundreds of parameters
- Each model simulation takes days to run
- Identifying highly sensitive parameters is critical

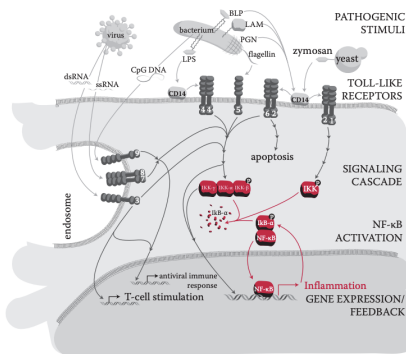


Source: *Developing a cloud-based toolbox for sensitivity analysis of a water quality model* (S. Kim et al, Environmental Modeling and Software, 2021)

Examples of sensitivity analysis: cell signaling

Toll-like signaling pathway:

- Cellular response to external stimuli (e.g. infection)
- Central role for NF- κ B transcription factor
- Shuttles back and forth between cytoplasm and nucleus



Source: Images from *Fundamentals of Systems Biology*, M. Covert, CRC Press, 2014.

Examples of sensitivity analysis: cell signaling

Hoffmann-Levchenko (2005): Computational model for NF- κ B

- 25 ODEs, 36 parameters
- Models protein production, degradation, transport
- Important role for **parameter estimation** and **sensitivity analysis**

$$\begin{aligned}
 \frac{d[NF\kappa B]}{dt} = & \underbrace{-a4 \cdot [IkBa] \cdot [NF\kappa B] - a4 \cdot [IKK_IkBa] \cdot [NF\kappa B] - a5 \cdot [IkBb] \cdot [NF\kappa B]}_{\text{association}} \\
 & \underbrace{-a5 \cdot [IKK_IkBb] \cdot [NF\kappa B] - a6 \cdot [IkBe] \cdot [NF\kappa B] - a6 \cdot [IKK_IkBe] \cdot [NF\kappa B]}_{\text{association}} \\
 & \underbrace{+d4 \cdot [IkBa_NF\kappa B] + d4 \cdot [IKK_IkBa_NF\kappa B] + d5 \cdot [IkBb_NF\kappa B]}_{\text{dissociation}} \\
 & \underbrace{+d5 \cdot [IKK_IkBb_NF\kappa B] + d6 \cdot [IkBe_NF\kappa B] + d6 \cdot [IKK_IkBe_NF\kappa B]}_{\text{dissociation}} \\
 & \underbrace{+r4 \cdot [IKK_IkBa_NF\kappa B] + r5 \cdot [IKK_IkBb_NF\kappa B] + r6 \cdot [IKK_IkBe_NF\kappa B]}_{\text{IKK reaction degradation}} \\
 & \underbrace{+deg4 \cdot [IkBa_NF\kappa B] + deg4 \cdot [IkBb_NF\kappa B] + deg4 \cdot [IkBe_NF\kappa B]}_{\text{spontaneous degradation}} \\
 & \underbrace{-k1 \cdot [NF\kappa B] + k01 \cdot [NF\kappa Bn]}_{\text{nuclear translocation}}
 \end{aligned}$$

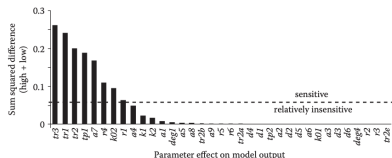
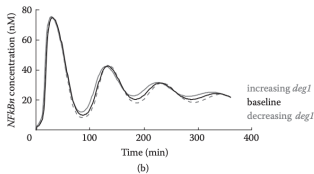
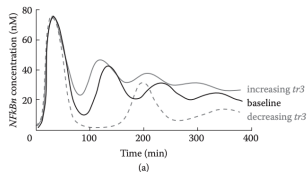
Source: Images from *Fundamentals of Systems Biology*, M. Covert, CRC Press, 2014.

Examples of sensitivity analysis: cell signaling

Sensitivity analysis: which parameters affect the model the most?

- Transcription rate: affects output a lot (**sensitive**)
- Degradation rate: relatively **insensitive**

Gives rough idea, needs to be corroborated with full model.



Source: Images from *Fundamentals of Systems Biology*, M. Covert, CRC Press, 2014.

Local sensitivity analysis

Local sensitivity analysis

How sensitive is model output (y) to changes of model parameter (θ) *at one single point* in parameter space?

- **(Absolute) local sensitivity:** partial derivative of variable with respect to parameter at single point in parameter space

$$S(\theta, x) = \frac{\partial y}{\partial \theta}(\theta, x)$$

- If k parameters, then also k sensitivity functions:

$$S_i(\theta, x) = \frac{\partial y}{\partial \theta_i}(\theta, x), \quad i = 1, \dots, k.$$

Local sensitivity analysis: absolute sensitivity

Problem: often very hard to compute partial derivative analytically.

Solution: compute derivative **numerically** through finite difference method:

- Forward difference:

$$\left. \frac{\Delta y}{\Delta \theta_j} \right|_+ = \frac{y(x, \theta_j + \Delta \theta_j) - y(x, \theta_j)}{\Delta \theta_j}$$

- Backward difference:

$$\left. \frac{\Delta y}{\Delta \theta_j} \right|_- = \frac{y(x, \theta_j) - y(x, \theta_j - \Delta \theta_j)}{\Delta \theta_j}$$

Local sensitivity analysis: absolute sensitivity

- How to choose perturbation $\Delta\theta_j$?
 - Too large: approximation is not good
 - Too small: numerical instabilities.
- In practice, choose $\Delta\theta_j$ **small** and **fixed**, e.g.

$$\Delta\theta_j = 10^{-6}.$$

Convergence

Both the forward and the backward difference agree with the derivative up to **first order** in $\Delta\theta_j$:

$$\frac{\partial y(x)}{\partial \theta_j} = \left. \frac{\Delta y(x)}{\Delta \theta_j} \right|_+ + \mathcal{O}(\Delta \theta_j), \quad \frac{\partial y(x)}{\partial \theta_j} = \left. \frac{\Delta y(x)}{\Delta \theta_j} \right|_- + \mathcal{O}(\Delta \theta_j).$$

Local sensitivity analysis: absolute sensitivity

- Third option: central difference

$$\frac{\Delta y(x)}{\Delta \theta_j} = \frac{y(x, \theta_j + \Delta \theta_j) - y(x, \theta_j - \Delta \theta_j)}{2\Delta \theta_j}$$

Convergence

The central difference agrees with the derivative up to **second order** in $\Delta \theta_j$:

$$\frac{\partial y(x)}{\partial \theta_j} = \frac{\Delta y(x)}{\Delta \theta_j} + \mathcal{O}((\Delta \theta_j)^2).$$

Local sensitivity analysis: relative sensitivity

Absolute sensitivity is influenced by magnitude of variable and parameter.

- Problematic if we want to compare sensitivities of different combinations of outputs and parameters
- Use **relative sensitivity**.

Local sensitivity analysis: relative sensitivity

Different definitions, depending on what's important:

- 1 Relative sensitivity w.r.t. parameter:

$$\frac{\partial y(t)}{\partial \theta_j} \cdot \theta_j$$

Compare sensitivity of same variable w.r.t. *different parameters*

- 2 Relative sensitivity w.r.t. variable

$$\frac{\partial y_i(t)}{\partial \theta} \cdot \frac{1}{y_i}$$

Compare sensitivity of *different variables* w.r.t. same parameter

Local sensitivity analysis: relative sensitivity

③ Total relative sensitivity

$$\frac{\partial y_i(t)}{\partial \theta_j} \cdot \frac{\theta_j}{y_i}$$

Compare all sensitivities (of *different variables* w.r.t. *different parameters*)

Local sensitivity analysis

- Relative sensitivities allow to **rank sensitivities**. Important for:
 - Choice parameters for parameter estimation
 - Choice parameters for model reduction
 - Choice for additional measurement or experimental determination of parameter (reduce sources of uncertainty)
- Ranking **depends on value of parameter**, can be different at different position in parameter space
- How to compare continuous sensitivity functions?
- Interest in specific values of independent variable
 - Where measurements are available
 - Where measurements will be collected

Local sensitivity analysis

- Create generic model with
 - Time t as independent variable
 - Outputs y_i , $i = 1, \dots, v$
 - Parameters θ_j , $j = 1, \dots, p$
 - Moments of measurements t_k , $k = 1, \dots, N$
- Total relative sensitivity of variable y_i w.r.t. parameter θ_j at moment t_k

$$S_{i,j,k} = \frac{\partial y_i(t_k)}{\partial \theta_j} \cdot \frac{\theta_j}{y_i}$$

Local sensitivity analysis

Importance parameter is determined by its impact on *all* variables

→ sum and average over all variables

→ take sign into account (square and root)

root mean square sensitivity for parameter θ_j

$$\delta_{j,k}^{rmsq} = \sqrt{\frac{\sum_{i=1}^v S_{i,j,k}^2}{v}}$$

$\delta_{j,k}^{rmsq}$ can be very variable from moment to moment

→ sum and average over all time points

time mean root mean square sensitivity for parameter θ_j

$$\delta_j^{rmsq} = \frac{1}{N} \sum_{k=1}^N \delta_{j,k}^{rmsq}$$

Local sensitivity analysis

- Gives one single measure for sensitivity of parameter
- Use this measure to determine importance of parameter
- Obtained value depends on
 - nominal parameter value: nonlinear models give different values at different location in parameter space (see also global sensitivity analysis)
 - choice of time points is arbitrary: this can lead to different set of parameters that are best estimated using dataset (see also identifiability)
- Modifications can be defined based on application/goal

Side track: Monte Carlo simulation

Monte Carlo simulation technique

- Computational algorithm, based on
 - Large number of simulations (thousands to millions)
 - Repeated random sampling
- Used for modeling of
 - Climate and environment
 - Nuclear reactions
 - Financial systems
 - Molecular dynamics
 - ...
- In this course:
 - Global sensitivity analysis
 - Uncertainty analysis



Principles behind Monte Carlo

Algorithm:

- 1 Draw random parameter samples
- 2 For each sample, run simulation
- 3 Aggregate results of simulations into quantity of interest (e.g. mean)

Advantages:

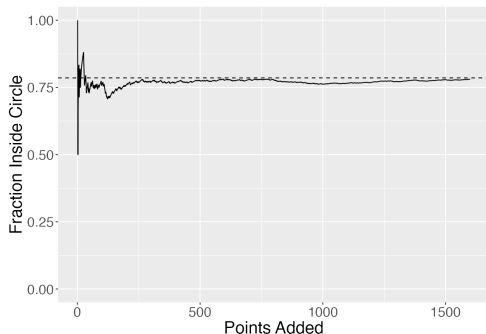
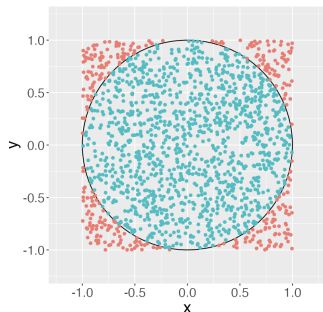
- Easy to implement/understand
- Applicable to wide range of problems (model-agnostic)

Disadvantages:

- Relatively slow convergence (many simulations needed)
- Requires information about parameter distributions

Example: computing π

- Drop points uniformly at random in the square $[-1, 1] \times [-1, 1]$
- Count fraction for which $x^2 + y^2 < 1$
- As $N \rightarrow \infty$, gives approximation for $\pi/4$.



App available at <https://shiny-stats.fly.dev/monte-carlo/>.

Global sensitivity analysis (GSA)

Global sensitivity analysis (GSA)

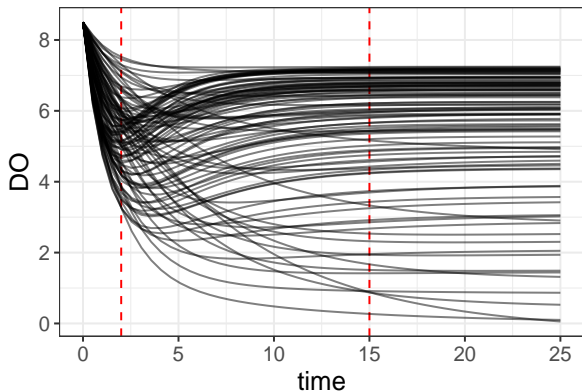
- Measure for sensitivity in delimited area in parameter space
- PDFs for parameters need to be chosen/found (same as for uncertainty analysis)

3 techniques will be discussed:

- Standardized regression coefficients
- Screening techniques
- Variance decomposition

GSA: Standardized regression coefficients

- Linear regression of Monte Carlo simulations
- Each line is simulation of variable y for different parameter set Θ , i.e., other point in parameter space



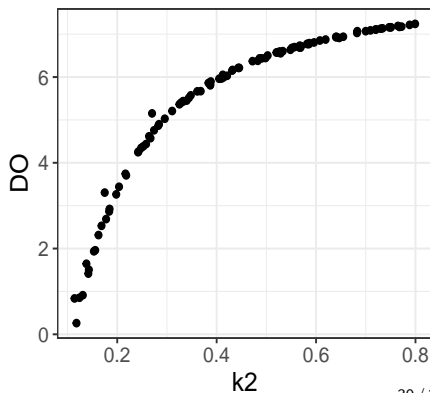
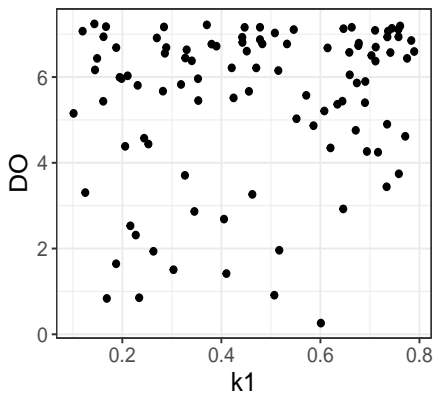
- Figure: 100 simulations of dissolved oxygen, with k_1, k_2 sampled uniformly between 0.1 and 0.8.

GSA: Standardized regression coefficients

- Consider outcomes at fixed time T
- Quantify effect of parameters $\theta_1, \dots, \theta_p$ through linear model

$$y_{t=T} = b_1\theta_1 + \dots + b_p\theta_p + \epsilon$$

- Regression coefficient b_i gives contribution of parameter θ_i in explaining variance of $y_{t=T}$



GSA: Standardized regression coefficients

- Correct for spread on both parameter and output
- Recalculate coefficients b_i to $SRCs$

$$SRC_{\theta_i} = b_i \cdot \frac{\sigma_{\theta_i}}{\sigma_y}$$

- Sample standard deviations from
 - vector $y_{t=T}$ for output
 - parameter samples for parameter θ_i .

GSA: Standardized regression coefficients

- For linear model in parameters that were examined, the total variance is explained by *SRCs*

$$\sum_i SRC_{\theta_i}^2 = 1$$

- For nonlinear models (in the parameters) *not* all variance will be explained. The part that is explained is given by determination coefficient

$$R^2 = \sum_{i=1}^n \frac{(\hat{y}_i - \bar{y})^2}{(y_i - \bar{y})^2}$$

- \hat{y}_i : prediction by regression model
- Technique only valid if $R^2 > 0.7$

GSA: Screening techniques

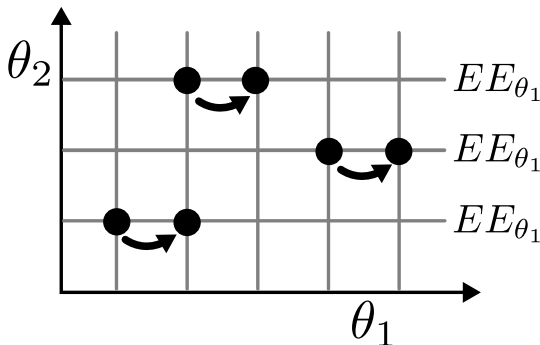
- Goal: obtain idea of importance of model parameters using only a limited number of simulations
- Example of technique: Morris screening
- Calculation of *elementary effect* for θ_i

$$EE_{\theta_i} = \frac{y(\theta_i + \Delta) - y(\theta)}{\Delta}$$

- Δ is a predetermined step size in parameter
- Remark analogy with local sensitivity, however, step size much larger

GSA: Morris screening

- Assume 2 parameters θ_1 and θ_2
- Choose regions in parameter space to compute elementary effects
- Summarize EE using mean and variance



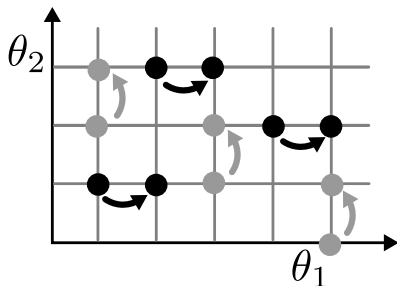
GSA: Morris screening

- Vector of EE (in this case 3)
- Statistical analysis of this vector
 - $\mu_{EE_{\theta_i}}$: indication on average effect of this parameter over entire parameter space; large value means important parameter and vice versa
 - $\sigma_{EE_{\theta_i}}$: information about linear behavior of parameter; large value means nonlinear parameter or parameter involved in interactions with other parameters
- Do same for θ_2

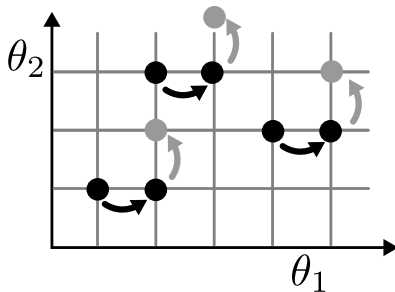
GSA: Morris screening

- Normally 2 simulations needed per EE
- 1991: Morris introduced more efficient way
- Number of simulations needed for p parameters:
 - Naive: $(2 \times \#EE)^p$
 - Morris: $(p + 1) \times \#EE$

Naive screening



Morris screening



GSA: Variance decomposition

- Goal: find share of each model parameter in variance of model output
- Used for models that are strongly nonlinear or nonmonotonous
- Example of model with 3 parameters:

$$\sigma_y^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2 + \sigma_{123}^2$$

- Normalisation (i.e., divide by σ_y^2) gives *sensitivity indices*

$$1 = S_1 + S_2 + S_3 + S_{12} + S_{13} + S_{23} + S_{123}$$

- Indicate which fraction of total variance is determined by certain parameter or parameter combination

GSA: Variance decomposition

- *Total sensitivity indices*

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123}$$

$$S_{T2} = S_2 + S_{12} + S_{23} + S_{123}$$

$$S_{T3} = S_3 + S_{13} + S_{23} + S_{123}$$

- Give total contribution of a certain parameter, including interaction effects
- Watch out: some contributions are counted multiple times, hence sum of all total sensitivity indices is no longer 1

GSA: Variance decomposition

- Two techniques:
 - FAST (Fourier Amplitude Sensitivity Test): uses Fourier decomposition of model output; can determine first order effects (total effects \rightarrow extendedFAST); computationally intensive ((ten) thousands of simulations)
 - Sobol indices: uses multiple integrals, both first order and higher order effects; computationally expensive; less efficient than FAST