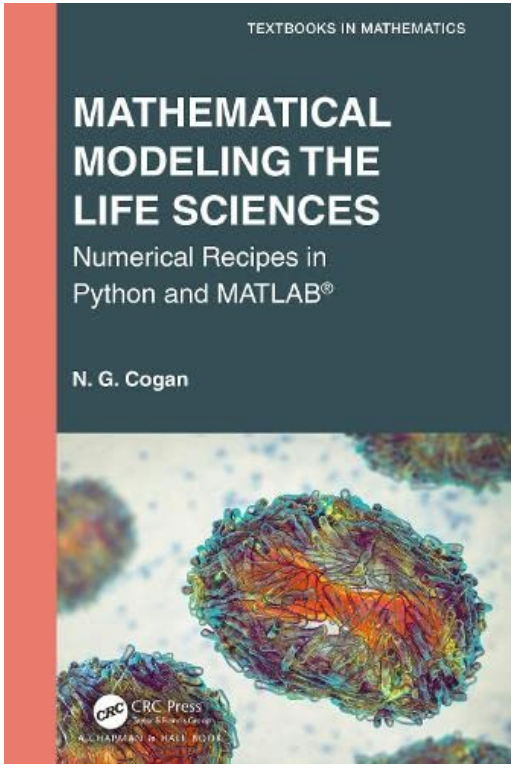
Sensitivity Analysis in the Life Sciences

Nick Cogan
Florida State University

June 11, 2024

Purpose:

- **Introduce** the dominant concepts of SA
- **Motivate** the move from trivial SA to better methods
- **Demonstrate** the flexibility of SA and different interpretations
- **Provide** codes naive implementations and some more sophisticated packages
  https://www.math.fsu.edu/∼cogan/LS2024/MT.html



Sensitivity Analysis:

Given an input/output model:

$$M(\vec{p}) \quad = \quad \vec{F}$$

given parameters, p_i.

SA quantifies how variations of parameters impact $\vec{F}$.

This can be relative, ranked or absolute.

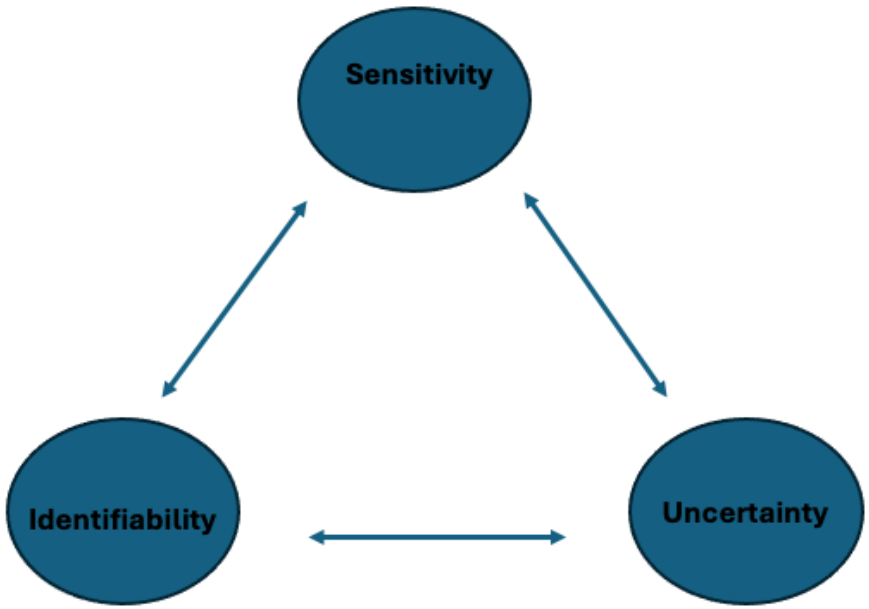
Why do we care about SA?

- "Model Reduction": Insensitive parameters can be frozen
- Sensitive parameters play an important role: Potential control targets, needed data etc.
- Provide insight into robustness of predictions given parameter variations
- Informs model structure, identifiability, identifiability

Parameter values vary due to

Aleatoric Uncertainty: Intrinsic, irreducible uncertainty (roll the dice)

Epistemic Uncertainty: Lack of knowledge, imprecise measurements, data issues



Classifications/terminology:

- Local vs. Global: ~~One-at-a-time~~ or all parameters
  Note that all estimates are local relative to nominal parameters
- Variance/Correlation: Not necessarily exhaustive list
- Screening methods: Fast/imprecise.
- Other measures that are specific for topics (Correlated parameters, surrogate models, ...)

Example 1: Local methods

Regression: Competitive exclusion

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1 \frac{\kappa_1 - N_1 - \alpha_{12} N_2}{\kappa_1}, \\ \frac{dN_2}{dt} &= r_2 N_2 \frac{\kappa_2 - N_2 - \alpha_{21} N_1}{\kappa_1}.\end{aligned}$$

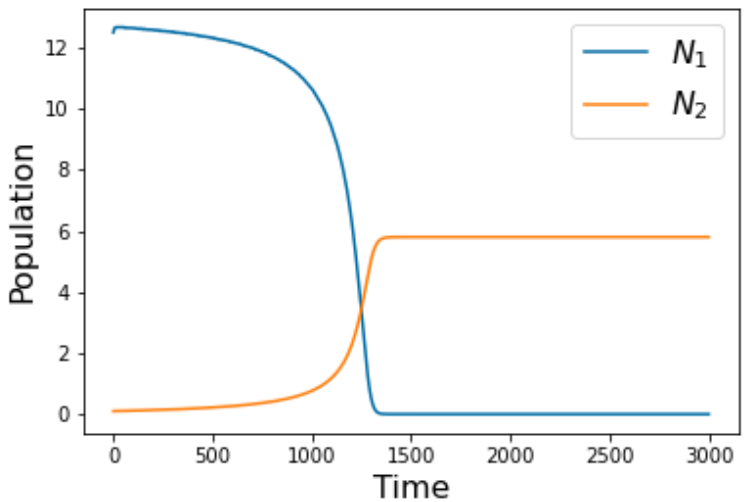


Figure: An example of exclusion using parameters from Gauses paper. N_2 exerts more pressure on N_1 than the other way around.

Example 1: QoI

QoI: Dominance

$$
\begin{aligned}
s_1 &= \frac{N_1}{N_1 + N_2} \\
s_2 &= \frac{N_2}{N_1 + N_2}
\end{aligned}
$$

Example 2: Local methods

Relative Change: Tuberculosis

Free bacteria, B

Dormant bacteria, Q

Immune response, X

$$\begin{aligned}
\frac{dB}{dt} &= rB + gQ - (hBX + fB), \\
\frac{dQ}{dt} &= fB - gQ, \\
\frac{dX}{dt} &= a + sX\left(\frac{B}{k+B}\right) - dX.
\end{aligned}$$

$$S = \frac{\frac{\Delta QoI}{QoI}}{\frac{\Delta p}{p}}.$$

This comes from approximating $S = \frac{\partial QoI}{\partial p}$ which is essentially the definition of local sensitivity.

Example 2: QoI=Time of maximum $B + Q$

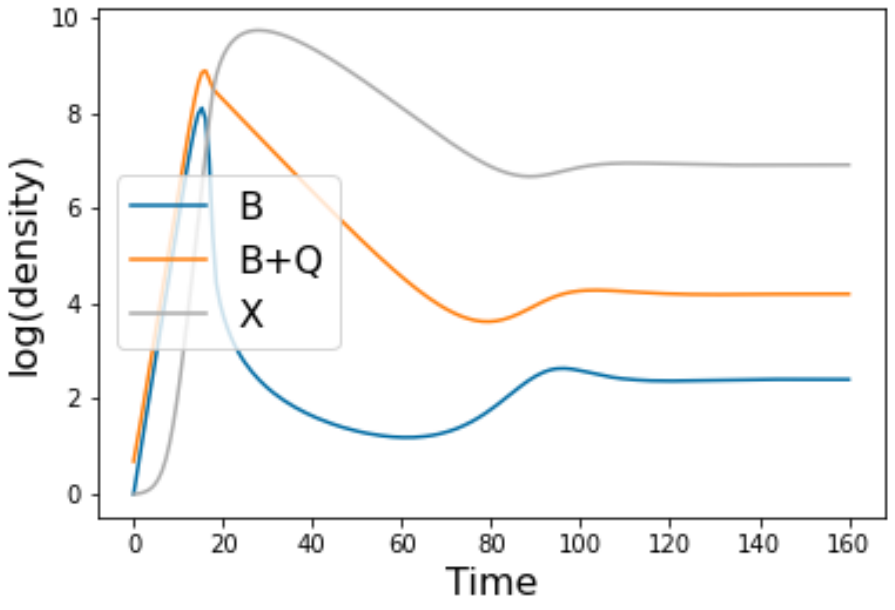
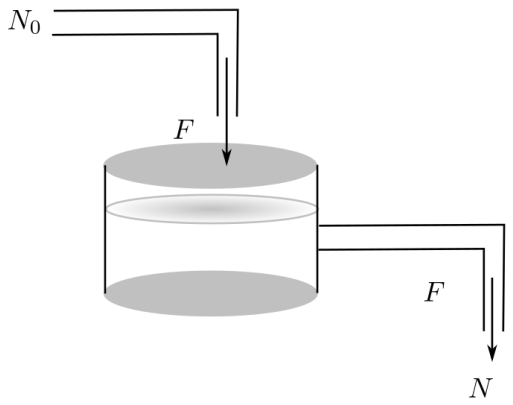


Figure: Numerical solution of the model of TB. The two bacterial densities are added together to compare the total with the active load.

Example 3: Correlation Coefficients

Correlation Coefficient: Chemostat



$$\frac{dN}{dt} = N_0 F - \frac{1}{Y}\frac{\mu N}{K_N + N}B - FN \quad (1)$$

$$\frac{dB}{dt} = \frac{\mu N}{K_N + N}B - FB. \quad (2)$$

Example 3: Correlation Coefficients

Spearman Correlation Coefficient: Freter Sampling matters (?): LHS, Monte Carlo, QMC,

$$r_{p_j,y} = \frac{\sum_{i=1}^{n}(p_{ij} - \bar{p}_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(p_{ij} - \bar{p}_j)^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \quad (3)$$

define

$$y = QoI = \int_{t=t_1}^{t=t_2} B(t)\ dt.$$

Example 3: Spearman (Pearson on ranked data)

$$\frac{dN}{dt} = N_0 F - \frac{1}{Y}\frac{\mu N}{K_N + N}(B_u + B_b) - FN, \tag{4}$$

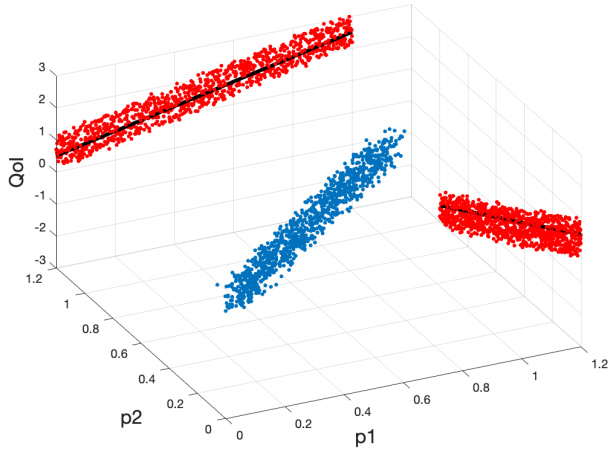
$$\frac{dB_u}{dt} = B_0 F - \frac{\alpha_{max}}{K_\alpha + B_b} B_u + \frac{V}{A}\beta B_b - FB_u$$

$$+(1 - \frac{B_b}{K_b + B_b})\frac{1}{Y}\frac{\mu N}{K_N + N} B_b, \tag{5}$$

$$\frac{dB_b}{dt} = \frac{A}{V}\frac{\alpha_{max}}{(K_\alpha + B_b)} B_u - \beta B_b$$

$$+\frac{A}{V}\frac{B_b}{K_b + B_b}\frac{1}{Y}\frac{\mu N}{K_N + N} B_b. \tag{6}$$

Example 3: PRCC (discount interactions)



The partial rank correlation coefficient for a single parameter, p_1 and for a QoI that depends on parameters (p_1, p_2) is,

$$r_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}.$$

$$prcc_{p_1} = \frac{r_{p_1,QoI} - r_{p_1,p_2} r_{p_2,QoI}}{\sqrt{(1 - r^2_{p_1,p_2})(1 - r^2 p_2, QoI)}}.$$

Example 4: Global – Hodgkin-Huxley

$$
\begin{aligned}
C_m \frac{dV}{dt} &= -\bar{g}_K n^4 (V - V_k) - \bar{g}_{Na} m^3 h (V - V_{Na}) \\
&\quad -\bar{g}_l (V - V_l) + I_{app}, \\
\frac{dm}{dt} &= \alpha_m (1 - m) - \beta_m m, \\
\frac{dn}{dt} &= \alpha_n (1 - n) - \beta_n n, \\
\frac{dh}{dt} &= \alpha_h (1 - h) - \beta_h h,
\end{aligned}
$$

Example 4: Global – Hodgkin-Huxley

$$
\begin{aligned}
\alpha_m &= 0.1\frac{25-v}{e^{\frac{25-v}{10}}-1}, \\
\beta_m &= 4e^{-\frac{v}{18}}, \\
\alpha_h &= 0.07e^{-\frac{v}{20}}, \\
\beta_h &= \frac{1}{e^{\frac{30-v}{10}}+1}, \\
\alpha_n &= 0.01\frac{10-v}{e^{\frac{10-v}{10}}-1}, \\
\beta_n &= 0.125e^{-\frac{v}{80}}.
\end{aligned}
$$

QoI equal to the maximum voltage

Example 4: Variance based (Sobol')

Sobol showed that for a very general functional relationship between Y and parameters X_i, $Y = f(X_1, X_2, ...X_p)$ there is a unique sequence of functions that can be used to construct Y that have two properties. First, the functions have specific interdependencies between the parameters and second that they are orthogonal.

Example 4: Variance based (Sobol')

Generalized ANOVA: Functions can be uniquely decomposed:

$$Y = f(X_1, X_2, ...X_p) = f_0 + \sum_{1}^{p} f_i(X_i) + \sum_{1 \le i < j \le p} f_{ij}(X_i, X_j) + ... + f_{1,...,p}(X_1, ..., X_p).$$

where 1.)

$$\int_{[0,1]^p} f_i(X_i) dx_j = 0$$

2.)

$$\int_{[0,1]^p} f_{i_1,...,i_p}(X_{i_1}, ..., X_{i_q}) f_{j_1,...,j_p}(X_{j_1}, ..., X_{j_q}) d\mathbf{x} = \mathbf{0}.$$

This is a notationally dense way to say that the integral of any of the functions that are used to construct Y against any other is zero.

Example 4: Variance based (Sobol')

By squaring both sides of Equation 7 and using the orthogonality, we find a relationship between the total variance and variance due to the parameters,

$$\int_{[0,1]^p} f^2 d\mathbf{X} = \mathbf{V} = \sum_{i=1}^{p} V_i + \sum_{i<j} V_{ij} + \sum_{i<j<k} V_{i,j,k} + \ldots + V_{1,2,3,\ldots p}\,, \tag{7}$$

where

$$V_{i_1,i_2,\ldots,i_s} = \int_{[0,1]^p} f^2_{i_1,i_2,\ldots,i_s} dX_{i_1}, \ldots, dX_{i_s},$$

is the variance due to the specific parameter combinations $(X_{i_1}, \ldots, X_{i_s})$.

Example 4: Variance based (Sobol')

It is more useful to consider the comparisons of the partial variances to the total variance, so that Equation 7 can be written as,

$$1 = \frac{\sum_{i=1}^{p} V_i}{V} + \frac{\sum_{i<j} V_{ij}}{V} \ldots + \frac{\sum_{i<j<k} V_{i,j,k} + \ldots + V_{1,2,3,\ldots p}}{V},$$

by dividing by V.
The main or first order effect of parameter X_i is just $\frac{V_i}{V}$.
total effect of parameter X_i each of the terms that include the parameters X_i ,

$$S_{T_i} = \frac{V_i}{V} + \frac{\sum_{i \neq j} V_{ij}}{V} \ldots + \frac{\sum_{i \neq j,k} V_{i,j,k} + \ldots + V_{1,2,3,\ldots p}}{V}.$$

Other Methods: Morris Screening

- Repeatedly estimate $\frac{\partial QoI}{\partial p_i}$ in a systematic method
- Use statistics of these estimates as ranking measures
  1. Start at the nominal parameter value: $(\bar{p}_1, \bar{p}_2, ..., \bar{p}_n)$
  2. Pick a direction randomly, use step in this direction for second value: $(\bar{p}_1, \bar{p}_2, ..., \bar{p}_i + \Delta p, ..., \bar{p}_n)$
  3. $$EE_i^1 = \frac{QoI(\bar{p}_1, \bar{p}_2, ..., \bar{p}_n) - QoI(\bar{p}_1, \bar{p}_2, ..., \bar{p}_i + \Delta p, ..., \bar{p}_n)}{\Delta p}.$$
  4. Repeat and generate:
  $$\mathbf{EE}_k = (EE_k^1, EE_k^2, ..., EE_k^m),$$
- Data for local sensitivity, for a region of parameter space (quasi global)

Other QoI's

Given a model and data, QoI $= ||data - model||$ is a carrier for identifiability

Things to take away:

- It is pretty simple to make impactful SA
- Creative QoI provides novel information
- I do not create SA methods, mostly just a consumer, but there are needs
- Don't do one-at-a-time