Sensitivity of inducible gene expression

Satyajeet Bhonsale*, Yadira Boada, Alejandro Vignoni**, Jesus Pico**, Jan Van Impe***

**BioTeC+, Chemical Engineering Department, KU Leuven, Technology Campus Ghent, Belgium*
*** SB2CL, Instituto de Automática e Informática Industrial, Universitat Polit´ecnica de Valencia, Valencia, Spain*

1. INTRODUCTION

Synthetic Biology is the engineering study of biology to enable (re)construction of cells to influence and control cellular behaviour. To avoid laborious trial and error experiments, synthetic biologists rely on mathematical models to design gene circuits. This belongs to the Design phase of the Design-Build-Test-Learn (DBTL) cycle. The use of modular modelling approaches allows synthetic biologists to compile various components in a variety of combinations and study cellular behaviour in silico. However, such use of mathematical models requires the bioparts to be appropriately characterized, i.e., the parameters which represent a particular biopart in the model must be uniquely and accurately identified.

Given the stochasticity of gene expression and measurement errors, these parameters are highly uncertain. Thus, in the design stage, it is imperative to understand the influence of this uncertainty. Moreover, it is important to know uncertainty in which of the bioparts has the largest influence.

In this paper, we demonstrate the utility of Sobol Global Sensitivity Analysis (GSA) in understanding how bioparts (i.e., model parameters) affect the engineered gene expression. As a case study, a synthetic gene circuit in which the expression of green fluorescence protein (GFP) by a transcription factor is considered. This circuit is illustrated in Figure 1. In the next sections, the gene circuit and the model used is described. The approach used to perform the GSA is then described briefly. Finally, the results of the sensitivity analysis are presented.

2. CASE STUDY

The gene circuit described here produces two proteins: LuxR (x_3) and GFP (x_8). LuxR is a constitutive protein produced in the cell (*Escherichia coli*). GFP on the other hand is an induced protein. The expression of GFP is triggered by addition of N-Acyl homoserine lactone (AHL) in the liquid medium (x_{10}). AHL diffuses through the cell membrane (x_9) and forms an AHL·LuxR complex is the transcription factor necessary to start GFP expression. Based on all the biochemical reactions involved in the above circuit, a detailed mathematical model is obtained. Then a reduced order model is obtained using quasi-steady state assumptions (Pushkareva et al. 2023). For the cellular growth, the standard Baranyi-Roberts growth model is used. The model however contains a variety of parameters. Some of these parameters are listed in Table 1.

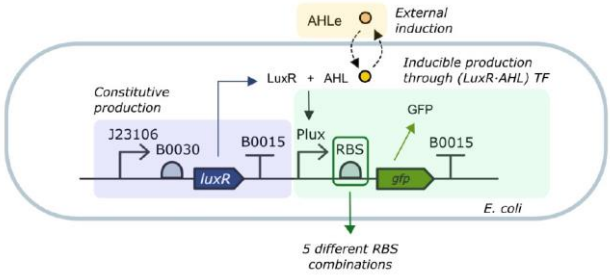

Figure 1. Gene circuit for expression of LuxR and GFP

$$\dot{x_3} = C_N \cdot \frac{k_1 \cdot k_2}{d_1 + \mu} - (d_2 + \mu) \cdot x_3$$

$$\dot{x_8} = C_N \cdot \frac{k_3 \cdot k_4}{d_3 + \mu} \cdot (\alpha + (1 - \alpha) \cdot \frac{(x_9)^n}{(\frac{k_{dlux} \cdot k_i \cdot C_N}{x_3})^n + (x_9)^n}) - (d_4 + \mu) \cdot x_8$$

$$\dot{x_9} = D \cdot \frac{V_{cell}}{V_{ext}} \cdot x_{10} - (D + d_A + \mu) \cdot x_9$$

$$\dot{x_{10}} = -D \cdot x_{11} \cdot \frac{V_{cell}}{V_{ext}} \cdot x_{10} + D \cdot x_9$$

Figure 2. Reduced model for expression system

3. GLOBAL SENSITIVITY ANALYSIS

The Sobol indices based GSA is used in this study. These indices are determined via a variance decomposition where in the total variance in the model response can be decomposed into variations due to individual parameters, and their higher order interactions as

$$\mathbb{V}[y] = \sum_{i}^{n} V_i + \sum_{i<j}^{n} V_{ij} + \cdots$$

Where V_i is the variance in the output due to parameter i, V_{ij} is the variance due to parameters i and j. With this

DOI: 10.34726/9016

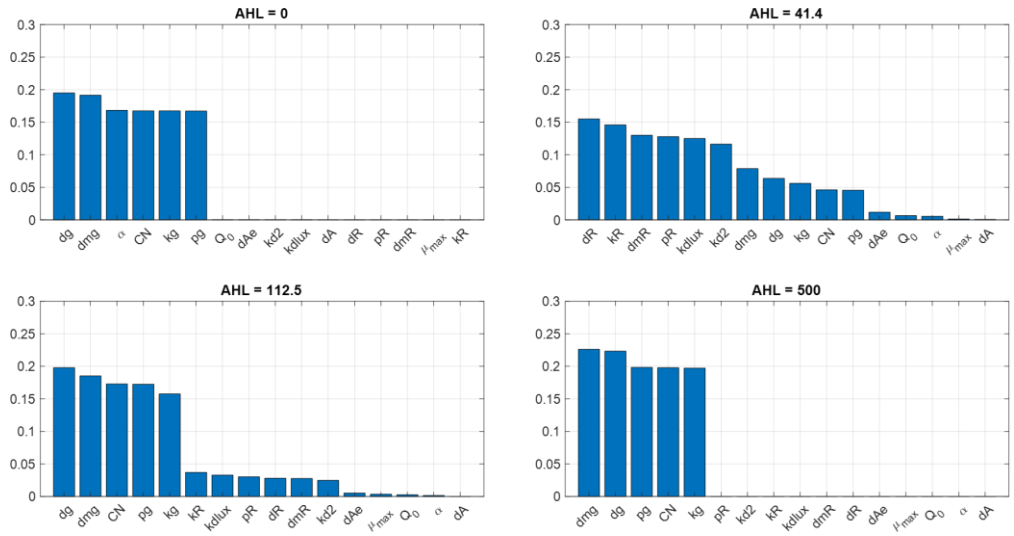

Figure 3. Sensitivity of GFP concentration at 48 h to model parameters

decomposition, the first order sensitivity index is defined as

$$S_i = \frac{V_i}{\mathbb{V}[y]}$$

This index accounts for the variance in the model response *only* due to parameter i. S_i answers to the question "which parameter should be fixed first to reduce the variance of the output?". In other words, a better estimation of this parameter (low parametric uncertainty) will reduce the output uncertainty. Higher order sensitivity indices can also be defined using the decomposed variance. In this study, the total and decomposed variances are computed using the polynomial chaos expansion approach (Bhonsale et al. 2019).

Table 1. Model Parameters

Parameter	Description	Parameter	Description
dg	GFP degradation	dmg	mGFP degradation
α	Basal GFP production	CN	Copy Number
kg	mGFP transcription	pg	GFP translation
dR	LuxR degradation	kR	mLuxR transcription
dmR	mLuxR degradation	pR	LuxR translation
Kdlux	Half-life	μ	Growth rate

4. RESULTS

Figure 2 depicts the sensitivity of GFP concentration to all the parameters at end of 48 h. The sensitivity is reported for cases with induction by different 4 AHL concentrations. It can be observed that at AHL induction concentrations of 0 nM and 500 nM, the GFP concentration is sensitivity to parameters related to GFP production, but insensitive to LuxR production. At AHL inductions of 41.4 nM and 112.5 nM, the GFP concentration is sensitive to parameters related to both LuxR and GFP production. In all cases, the GFP concentration is insensitive to the growth rate parameters.

These results highlight the nature of the model. The relationship between AHL induction and synthesis rate is captured by a Hill function. At no induction or very small induction (i.e. the Hill function is close to 0), the basal production rate (α) is important and thus parameters related to GFP production (translation, transcription, base production rate of GFP as well as degradation rates of GFP and GFP associated mRNA) are highlighted as the important. When the AHL concentration is high, the Hill function saturates, and the maximum synthesis rate is achieved. Here, the influential parameters are related to GFP production but not the basal production rate. For intermediate concentrations, the parameters involved in LuxR production (i.e., transcription and translation rates for LuxR as well as degradation rates for LuxR and associated mRNA) become important.

4. CONCLUSION

The sensitivity indices show certain bioparts become important under certain environmental conditions (in this case AHL). Stochasticity in insensitive parts will not affect the final output. For example, if the circuit must produce GFP under high AHL concentrations, LuxR parts don't quantitatively influence the production (they are still necessary). Such information can be used to design robust gene circuits.

REFERENCES

Bhonsale S., Munoz Lopez C. A., Van Impe J. (2019). Global Sensitivity Analysis of a Spray Drying Process. *Processes,* 7, 562.

Pushkareva A., Beltran J., Diaz-Iza H., Arboleda-Garcia A., Boada Y., Vignoni A., Picó J. (2023). Towards automation of the Design-Build-Test-Learn (DBTL) bioengineering cycle: Application to the testing and characterization of standard bioparts. XLIV Jornadas de Automática 2023

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni, J., Gatelli D., Saisana M., Tarantola S. (2008). *Global Sensitivity Analysis: The Primer*; John Wiley & Sons Ltd