7. Anexo I: Artículos de revista y de congresos derivados de este trabajo

Part characterization workflow using calibrated measurements, guaranteed simulations and a multiobjective optimization approach

Iván Alarcon-Ruiz, Raquel Andreu, Alejandro Vignoni, Yadira Boada, Jesús Picó
ivalrui1@upv.es,raanvi@upv.es,alvig2@upv.es,yaboa@upv.es,jpico@upv.es
Synthetic Biology and Biosystems Control Lab
Instituto de Automática e Informática Industrial
Universitat Politécnica de València
València, Spain

1 BACKGROUND

Synthetic biology aims at the targeted design or redesign and construction of new biological and bio-based parts, devices, and systems to perform desired functions [11, 13]. Not only is going up in this hierarchy (DNA, part, device, and system) the final objective of synthetic biology but also its main challenge [3]. To successfully accomplish it, engineering principles and methodologies are to be used. The design-build-test-learn (DBTL) cycle is the common paradigm used in any engineering discipline where the design is made from the bottom by combining basic biological parts into devices and these into systems [8]. Essential for the success of this inherently modular approach of bottom-up synthetic biology is the need of starting from well-characterized parts [9]. Currently, there is still a gap between the possibility of designing a system and its real implementation in the lab. This gap can be partially attributed to both the lack of repeatable and standardized measurements, and the absence of well-characterized biological parts [17]. On the one hand, calibration of equipment and standardization of units is a challenge. We cannot compare two systems in terms of measurements if they have not been created with the same methodology or they can not be taken into the same domain [2]. Luckily, there are many works from this community dealing with this issue [4, 6, 14, 16].

On the other hand, the characterization of biological parts understood as estimation of its model parameters is still a bottleneck in Synthetic Biology. Though the isolated characterization of the steady-state response of basic parts (e.g. promoters) can be performed by means of relatively simple experiments, their behavior when integrated within more complex circuits will be affected by the circuit context [15]. Therefore, *in situ* estimation of the parameters of each basic part in the circuit is carried out to account for context. However, estimating the parameters associated to biological parts embedded in a nonlinear dynamic model of a synthetic gene circuit remains a challenging inverse problem. Non-convexity, ill-conditioning caused by over-parametrization, experimental measurement errors, data scarcity, and uncertainty are the main difficulties [12]. The multiobjective optimization design (MOOD) framework to perform model parameter estimation has been successfully used for the optimal design of gene networks [5], or closed-loop genetic circuit identification [7], as it allows to address problems often found in synthetic gene networks that are difficult to tackle using single or weighted-objective optimization approaches.

2 METHODS

In [6] we propose a methodology combining measurement calibration with MOOD that is used to characterize the RBS strength of several genetic constructs. In this work, we propose an extended approach (see Figure 1) to solve the problem of biological parts characterization and fill the gap existing in the DBTL cycle based on multiobjective objective optimization that allows us to include uncertainty in the estimation and also to simultaneously include different scenarios and a set or library of genetic circuits or devices (which may have common parts and different parts).

The starting point (Figure 1.-1) is comprised of both the mathematical model and the set of biological parts to characterize (the model has to be one of those parts). From there we need to perform experiments to measure the desired magnitude (fluorescence of the reporter) and to calibrate the measurements (Figure 1.0) to obtain MEFL/Particle [1, 6]. Experimental data in this unit can be compared with mathematical models in molecules/cell. Also, it is necessary to define initial intervals for the model parameters to be estimated (and values for the parameters we don't want to estimate).

Once the prerequisites and the data preprocessing is performed, the actual methodology can be applied to characterize the selected parts. In the first step (Figure 1.1), we need to define the cost function for the optimization problem. In this case, we use a model for guaranteed simulations [10]

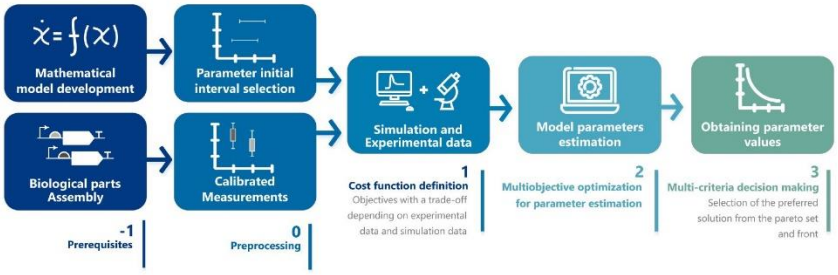

Figure 1: Schematic of the methodology. Prerequisites, data preprocesing, cost function definition, multiobjective optimization and multi-criteria decision making steps.

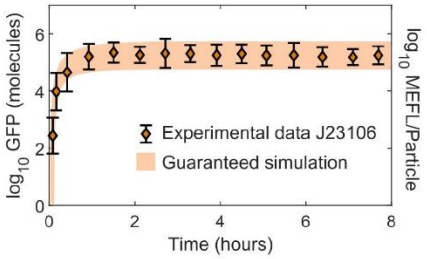

Figure 2: Example of simulation and experimental data to be used with the proposed methodology.

to obtain an interval solution (Figure 2) and in the objective function, we compare this interval solution with the calibrated measurements (including the measurement uncertainty). Then, we perform the optimization in the second step (Figure 1.2) to minimize the discrepancy between the experimental data and the simulation. The solution of the optimization is a pareto front (with the values of the objectives functions for the solutions) and the pareto set with the values of the parameters corresponding to those solutions. The last step (Figure 1.3) is to select a desired solution from the pareto front and set corresponding to the part we wanted to characterize.

REFERENCES

[1] Beal, J., Farny, N. G., Haddock-Angelli, T., Selvarajah, V., Baldwin, G. S., Buckley-Taylor, R., Gershater, M., Kiga, D., Marken, J., Sanchania, V., Sison, A., Workman, C. T., and . Robust estimation of bacterial cell count from optical density. *bioRxiv* (2019).

[2] Beal, J., Haddock-Angelli, T., Baldwin, G., Gershater, M., Dwijayanti, A., Storch, M., De Mora, K., Lizarazo, M., Rettberg, R., et al. Quantification of bacterial fluorescence using independent calibrants. *PloS one 13*, 6 (2018), e0199432.

[3] Beal, J., Haddock-Angelli, T., Farny, N., and Rettberg, R. Time to get serious about measurement in synthetic biology. *Trends in biotechnology 36*, 9 (2018), 869–871.

[4] Beal, J., Overney, C., Adler, A., Yaman, F., Tiberio, L., and Samineni, M. Tasbe flow analytics: A package for calibrated flow cytometry analysis. *ACS synthetic biology 8*, 7 (2019), 1524–1529.

[5] Boada, Y., Reynoso-Meza, G., Vignoni, A., and Picó, J. Multiobjective optimization framework to obtain model-based guidelines for tuning biological synthetic devices: an adaptive network case. *BMC Systems Biology 10* (2016), 27.

[6] Boada, Y., Vignoni, A., Alarcon-Ruiz, I., Andreu-Vilarroig, C., Monfort-Llorens, R., Requena, A., and Picó, J. Characterization of gene circuit parts based on multiobjective optimization by using standard calibrated measurements. *ChemBioChem 20*, 20 (2019), 2653–2665.

[7] Boada, Y., Vignoni, A., and Picó, J. Multiobjective identification of a feedback synthetic gene circuit. *IEEE Transactions on Control Systems Technology, (early access)* , (2019).

[8] Canton, B., Labno, A., and Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nature biotechnology 26*, 7 (2008), 787–793.

[9] Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A., and Weiss, R. Realizing the potential of synthetic biology. *Nature Reviews. Molecular Cell Biology 15*, 4 (2014), 289–294.

[10] De Pereda, D., Romero-Vivo, S., Ricarte, B., and Bondia, J. Guaranteed computation methods for compartmental in-series models under uncertainty. *Computers & Mathematics with Applications 66*, 9 (2013), 1595–1605.

[11] ERASynBio. Next steps for european synthetic biology: a strategic vision from erasynbio. Report, ERASynBio, 2014.

[12] Gábor, A., and Banga, J. R. Robust and efficient parameter estimation in dynamic models of biological systems. *BMC systems biology 9*, 1 (2015), 74.

[13] Kojima, R., and Fussenegger, M. Synthetic biology: Engineering mammalian cells to control cell-to-cell communication at will. *ChemBioChem 20*, 8 (2019), 994–1002.

[14] Martin, Y., Page, M., Blanchet, C., and de Jong, H. Wellinverter: a web application for the analysis of fluorescent reporter gene data. *BMC bioinformatics 20*, 1 (2019), 309.

[15] McBride, C., Shah, R., and Del Vecchio, D. The effect of loads in molecular communications. *Proceedings of the IEEE* (2019), 1–18.

[16] Robinson, C. M., Fedorec, A. J. H., Yan Wen, K., and Barnes, C. P. Flopr: An open source software package for calibration and normalization of plate reader and flow cytometry data. *bioRxiv* (2020).

[17] Way, J. C., Collins, J. J., Keasling, J. D., and Silver, P. A. Integrating biological redesign: where synthetic biology came from and where it needs to go. *Cell 157*, 1 (2014), 151–161.