2018-08-03

A reverse predictive model towards design automation of microfluidic droplet generators

Lashkaripour, Ali

International Workshop on Bio-Design Automation (IWBDA)

A Reverse Predictive Model Towards Design Automation of Microfluidic Droplet Generators

Ali Lashkaripour[1], Christopher Rodriguez[2] and Douglas Densmore[1,3]

[1]Department of Biomedical Engineering, Boston University, Boston, MA
[2]Department of Cyber Engineering, Louisiana Tech University, Ruston, LA
[3]Department of Electrical & Computer Engineering, Boston University, Boston, MA
lashkari@bu.edu, cwr023@latech.edu , and dougd@bu.edu

1. INTRODUCTION

Droplet-based microfluidic devices in comparison to test tubes can reduce reaction volumes 10^9 times and more due to the encapsulation of reactions in micro-scale droplets [4]. This volume reduction, alongside higher accuracy, higher sensitivity and faster reaction time made droplet microfluidics a superior platform particularly in biology, biomedical, and chemical engineering. However, a high barrier of entry prevents most of life science laboratories to exploit the advantages of microfluidics. There are two main obstacles to the widespread adoption of microfluidics, high fabrication costs, and lack of design automation tools. Recently, low-cost fabrication methods have reduced the cost of fabrication significantly [7]. Still, even with a low-cost fabrication method, due to lack of automation tools, life science research groups are still reliant on a microfluidic expert to develop any new microfluidic device [3, 5]. In this work, we report a framework to develop reverse predictive models that can accurately automate the design process of microfluidic droplet generators. This model takes prescribed performance metrics of droplet generators as the input and provides the geometry of the microfluidic device and the fluid and flow settings that result in the desired performance. We hope this automation tool makes droplet-based microfluidics more accessible, by reducing the time, cost, and knowledge needed for developing a microfluidic droplet generator that meets certain performance requirement.

2. DROPLET GENERATION

As shown in Fig. 1, by flowing an aqueous and a non-aqueous phase through a narrow opening, called orifice, microfluidic droplets are generated. The two major performance metrics of a droplet generator are droplet size and generation rate, which we call "dependent variables". These parameters are dictated by device geometry (i.e., orifice size, aspect ratio, oil width ratio, water with ratio, orifice length, and expansion ratio) and flow rates of oil and water (for a given geometry, these flow rates are determined by Capillary number and flow rate ratio), which we call "independent variables". To have a design automation tool for droplet generation, for a prescribed droplet size and generation rate, we need to provide geometry and flow conditions. Therefore, the goal is to take the dependent variables as input, and output the independent variables, that would result in the given dependent variables.

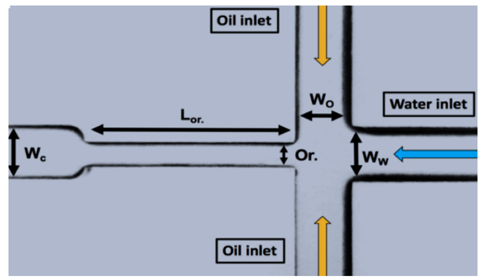

Figure 1: Droplet generation is achieved by flowing oil and water through a flow-focusing geometry. This process has eight inputs (six geometry, and two flow variables) and two outputs (droplet size and generation rate).

3. REVERSE PREDICTIVE MODELS

The first step in building a reverse predictive model is to construct a dataset of inputs-outputs over the range of expected values. in lieu of experimental data, we generated data points based on a formulaic relationship between the independent and dependent variables roughly derived from real world observations [6]. We scaled these equations to remain in reasonable ranges. These formulas are shown in Eqs. (1) & (2).

$$Generation\,rate = \frac{(OW + AR + ER + OL + WW + FR) * Ca * 5000}{Or} \tag{1}$$

$$Droplet\,size = \frac{Or * AR * ER * OL * WW}{OW * Ca * FR * 10} \tag{2}$$

where the parameters and their ranges are given in Table 1. Our models relied on min-max normalization to reduce biases in input magnitudes. For each column in our dataset (representing a type of parameter), we scaled each entry to be in the range of zero to one according to the minimum and maximum of that parameter's range.

We explored three methods for our reverse predictive models: nearest data point, M5P trees, and radial basis function (RBF) interpolation. Nearest data point is one of the simplest strategies, requiring no model to be fit to the data [1]. For any input of desired dependent variables, we simply search our data set for a point with dependent variable values that are the closest to the input. Then, we return the independent variables associated with that data point.

Table 1: The range of inputs (independent variables) used to build the input-output dataset using Eq. (1) and Eq. (2).

Symbol	Parameter	Range
OW	Normalized oil input width*	2 - 4
AR	Aspect ratio	1 - 3
OL	Normalized orifice length*	1 - 9
WW	Normalized water input width*	2 - 4
Or	Orifice width	50 - 300 μm
ER	Expansion ratio*	2 - 6
Ca	Capillary number	0.02 - 0.2
FR	Flow rate ratio	2 - 20

*The normalized values are divided by orifice width.

M5P trees are a more advanced version of linear regression where model trees branch out based on the value of the independent variables and data points that are close are put together in a same leaf. Each leaf contains an equation that represents a linear regression on the grouped data points [8]. We grow two M5P trees (one to optimize on each dependent variable) from our data set. Next, we search our training data set for a data point P which is closest to our desired input, much in the same way as nearest data point. We input the independent variable values from P into both M5P trees to obtain two linear equations Eq. (3) and Eq. (4). We require our solutions to satisfy both of these equations with no error. Therefore $f(x)$ must equal our desired droplet generation rate and $g(x)$ must equal our desired droplet size. There are an infinite number of points which we can accept, as we have only 2 constraints and 8 degrees of freedom. Therefore, we attempt to find a solution that deviates the least amount possible from our original closest data point P.

$$f(x) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \quad (3)$$

$$g(x) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (4)$$

RBF interpolation is a fast way to form regression models in high dimensions. In this study, we used a multiquadric function to build RBF regression models [2]. Much in the same way as the M5P trees, we fit two models to the data, one for each dependent variable. In order to generate suggestions using RBF interpolation, a nearest data point P is found (again, using the same method as nearest data point). P is used as the starting point for our optimization algorithm. We seek to find a point S that minimizes the error for all M models against all Y desired dependent variable values (performance metrics) as shown in Eq. (5). We use a form of gradient descent called SLSQP (Sequential Least Squares Programming) as our cost-minimization function.

$$\sum_{i=0}^{N} |M_i(S) - Y_i| \quad (5)$$

4. RESULTS

We created a dataset of 2500 points for training and another dataset of 2500 for accuracy verification. These datapoints are produced using Eqs. (1) & (2), while parameter values are taken randomly from the range given in Table 1. We tested the accuracy for both single and combined optimizations. Single optimization attempts to find a perfect solution on a single performance metric. Combined optimization attempts to find the best compromise, considering both performance metrics. The error is calculated as given in Eq. (6). Where x is the desired value, $M(x)$ is the model suggestion to get that desired value, and $f(M(x))$ is the "real" value of that suggestion calculated from Eqs. (1) & (2). The results are shown in Fig. 2., a) for single and b) for combined optimization.

$$Error = \frac{|f(M(x)) - x|}{x} \quad (6)$$

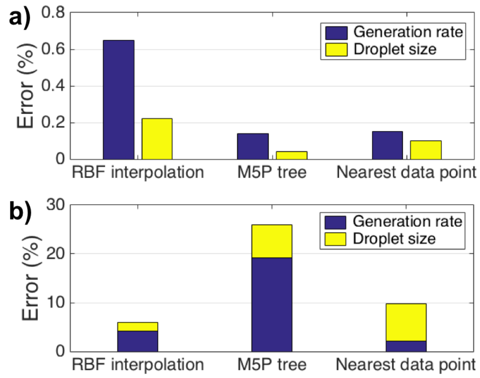

Figure 2: Accuracy comparison of different reverse predictive models. **a)** Only considering one performance metric. **b)** Considering both performance metrics, simultaneously.

5. CONCLUSION AND FUTURE WORK

We proposed a framework for design automation of microfluidic droplet generators, using a reverse predictive model. This model, takes the prescribed performance metrics as the input (droplet size and generation rate). Then, outputs geometry and flow conditions required to achieve this desired performance. The dataset of this study can be replaced by experimental data to accurately capture the real world behavior of microfluidic droplet generators.

6. REFERENCES

[1] D. M. Bates and D. G. Watts. *Nonlinear regression analysis and its applications*, volume 2. Wiley Online Library, 1988.

[2] D. S. Broomhead and D. Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks. Technical report, Royal Signals and Radar Establishment Malvern (United Kingdom), 1988.

[3] K. Chakrabarty and F. Su. Design automation challenges for microfluidics-based biochips. *DTIP of MEMS & MOEMS, Montreux, Switzerland*, pages 01–03, 2005.

[4] A. Lashkaripour, A. Abouei Mehrizi, M. Rasouli, and M. Goharimanesh. Numerical study of droplet generation process in a microfluidic flow focusing. *Journal of Computational Applied Mechanics*, 46(2):167–175, 2015.

[5] A. Lashkaripour, M. Goharimanesh, A. A. Mehrizi, and D. Densmore. An adaptive neural-fuzzy approach for microfluidic droplet size prediction. *Microelectronics Journal*, 78:73–80, 2018.

[6] A. Lashkaripour, A. A. Mehrizi, M. Goharimanesh, M. Rasouli, and S. R. Bazaz. Size-controlled droplet generation in a microfluidic device for rare dna amplification by optimizing its effective parameters. *Journal of Mechanics in Medicine and Biology*, 18(01):1850002, 2018.

[7] A. Lashkaripour, R. Silva, and D. Densmore. Desktop micromilled microfluidics. *Microfluidics and Nanofluidics*, 22(3):31, 2018.

[8] J. R. Quinlan et al. Learning with continuous classes. In *5th Australian joint conference on artificial intelligence*, volume 92, pages 343–348. Singapore, 1992.