Versatility and stability optimization of flow-focusing droplet generators *via* quality metric-driven design automation†

David McIntyre, [ID] [ab] Ali Lashkaripour, [ID] [cd] Diana Arguijo, [ID] [ab]
Polly Fordyce [ID] [cde] and Douglas Densmore [ID] *[bf]

Droplet generation is a fundamental component of droplet microfluidics, compartmentalizing biological or chemical systems within a water-in-oil emulsion. As adoption of droplet microfluidics expands beyond expert labs or integrated devices, quality metrics are needed to contextualize the performance capabilities, improving the reproducibility and efficiency of operation. Here, we present two quality metrics for droplet generation: performance versatility, the operating range of a single device, and stability, the distance of a single operating point from a regime change. Both metrics were characterized *in silico* and validated experimentally using machine learning and rapid prototyping. These metrics were integrated into a design automation workflow, DAFD 2.0, which provides users with droplet generators of a desired performance that are versatile or flow stable. Versatile droplet generators with stable operating points accelerate the development of sophisticated devices by facilitating integration of other microfluidic components and improving the accuracy of design automation tools.

Introduction

Droplet microfluidics is a core aspect of many high-throughput platforms in biotechnology, including functional antibody discovery, drug screening, metabolic pathway optimization, and single-cell genomics.[1–4] Typically, resource and time-intensive design iterations are necessary to achieve a desired performance, specifically when using poorly characterized biological samples[5–7] or multiple component devices (*e.g.*, generators, sorters, mergers).[8–10] Knowledge from previously successful implementation needs to be captured through standardized designs, well-annotated fabrication protocols, automated computer-aided design tools, and quality metrics to reduce the need for design iterations and make the design process more robust.[11,12]

Quality metrics provide essential insight into system performance (*e.g.*, sensitivity and specificity of a diagnostic, sampling rate of an integrated circuit, or fuel economy of a car). Quality metrics for key microfluidic components are needed as small errors from bespoke device design, fabrication, and operation can accumulate between researchers, resulting in large performance deviations across different groups.[13] Adoption of these metrics in microfluidic devices would improve the ease of implementation by non-experts, reduce batch variability, and provide important context on their stable and feasible performance range.

In droplet generation, the droplet size and generation rate are dictated by the geometric design of the device, flow conditions, fluid properties, and surface chemistry.[14,15] Monodisperse droplet generation at a single size and rate is essential for integration with other components and for encapsulation of cells, beads, or other reagents. T-junction, co-flow, flow-focusing, and step-emulsification geometries alongside pressure-driven or flow-rate-driven fluidic systems are commonly used for droplet generation with varying degrees of performance range (droplet sizes and generation rates) and parameter sensitivity.[16–18] Flow-focusing devices can deliver a wide range of diameters and generation rates while maintaining high droplet monodispersity, often making them more desirable in comparison to other geometries.[16,19]

Flow-focusing geometries are traditionally designed by choosing an orifice width close to the desired droplet diameter. However, experimental evidence suggests that other design parameters including channel depth, oil inlet width, water inlet width, and outlet channel width play a significant role in determining the characteristics of droplet generation.[20,21] This design strategy arises from the complex

[a] Biomedical Engineering Department, Boston University, MA, USA.
E-mail: dougd@bu.edu
[b] Biological Design Center, Boston University, Boston, MA, USA
[c] Department of Bioengineering, Stanford University, Stanford, CA, USA
[d] Department of Genetics, Stanford University, Stanford, CA, USA
[e] Chan-Zuckerberg Biohub, San Francisco, CA, USA
[f] Electrical & Computer Engineering Department, Boston University, Boston, MA, USA
† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3lc00189j

fluid velocity fields and large design space in flow-focusing geometries, which prevents the introduction of generalizable and accurate scaling laws.

To this end, machine learning-based design automation tools have been introduced for flow-focusing droplet generators that suggest a design that delivers a user-specified diameter and generation rate.[22] However, microfluidic design automation tools have primarily focused on predicting performance as accurately as possible, taking little to no account for the performance range, stability of the operating point, and the difficulties in fabricating or operating the device.[12]

Here, we define versatility and stability, two quality metrics that streamline flow-focusing droplet generator design and operation. Versatile devices are ideal for on-chip component integration, resource-constrained settings, novice microfluidic operators, or early-stage discovery, where surveying a wide range of droplet sizes and generation rates is needed (Fig. 1a). The broader operating range of these devices can also facilitate the integration of droplet generators with other microfluidic components. Stable droplet generators can be used to ensure that the operating point of the device is not near a regime change boundary (*e.g.*, from dripping to jetting) (Fig. 1b). These designs improve the robustness of droplet generation by avoiding large jumps in the observed performance from a regime

change caused by small errors in fabrication, operation, or predictive models.

To characterize the effect of device geometry on these metrics, we capitalized on a previously developed machine learning tool, DAFD,[22] to fully simulate the droplet generator design space. We established the influence of each parameter on versatility and stability with main effect analysis. Select devices were fabricated and used to experimentally validate each metric. Next, both metrics were integrated into the design automation algorithm to create quality metric-driven design automation of flow-focusing droplet generators, DAFD 2.0 (Fig. 1c). These metrics can be implemented to tailor the performance range of a device, improve the robustness of operation, or simplify the development of multi-component microfluidic devices (Fig. 1d).

Results

Rapid modeling of droplet generation

To determine if geometric parameters of a flow-focuser other than orifice width impact droplet generation, we analyzed previously published experimental data from 5 orthogonal flow-focusing devices with different design parameters while keeping the same orifice width (Table S1†).[21] These devices were tested at the same range of capillary number (0.066–1.06) and flow rate ratio (10–22) and their observed diameters
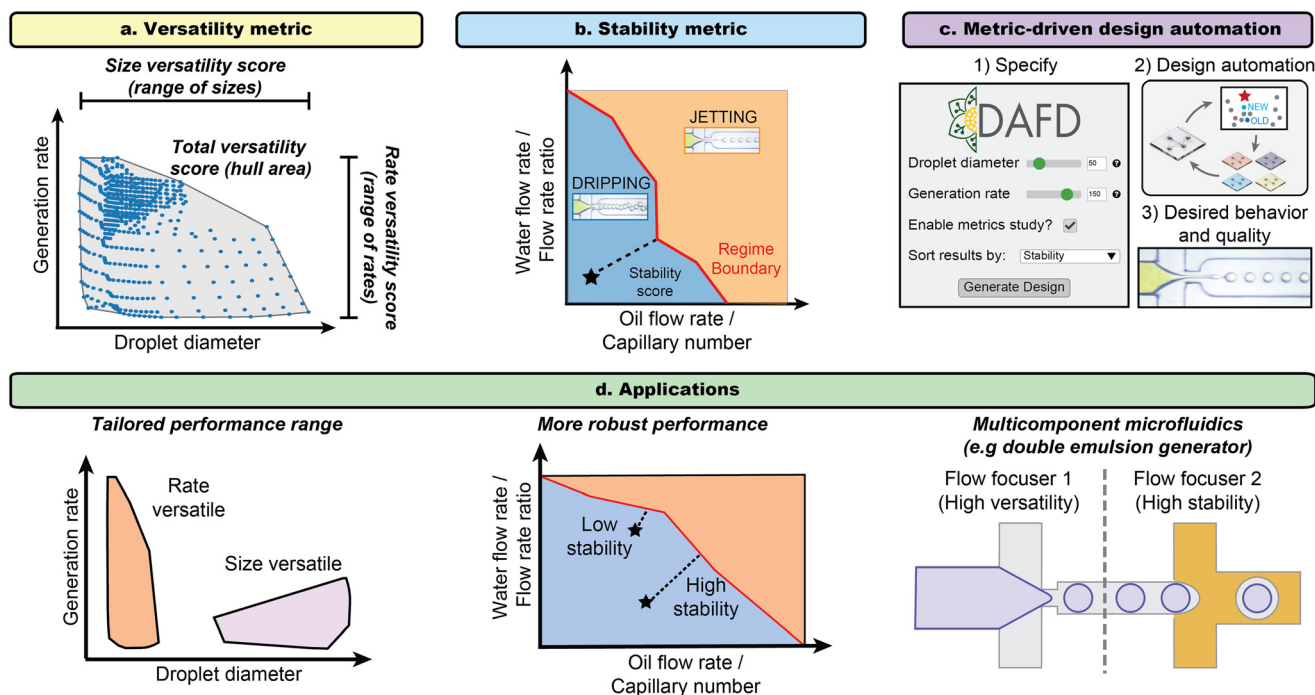
Fig. 1 Versatility and stability are established and characterized as new quality metrics for flow-focusing droplet generators. (a.) Size, rate, and total versatility are defined using the droplet diameter range, generation rate range, and the convex hull of observed performance while testing the device over the flow condition design space, respectively. (b.) Stability is defined as the 2-D Euclidean distance of oil and water flow rates from a specified point to a regime change boundary. (c.) These quality metrics were integrated into the DAFD design automation workflow to develop the next generation of the online tool that can achieve user-specified performance while maximizing desired quality metrics. (d.) Versatility and stability metrics can be used in applications such as tailoring the performance range of a device, improving the robustness of droplet generation, or simplifying multicomponent microfluidic development, such as a double emulsion generator.

and generation rates were used to compare their performance range. Capillary number and flow rate ratio were used as dimensionless representations of flow rates to maintain the generalizability of our findings to a broader range of channel dimensions. As shown in Fig. 2, the performance of these devices varied significantly despite having the same orifice width. Interestingly, some devices produced a broader range of droplet diameters or generation rates than others. Next, to establish the accuracy of previously validated neural networks in predicting the performance of aqueous-in-oil droplet generation,[22] we predicted the droplet diameters and generation rates of the 5 orthogonal devices at the same range of capillary numbers and flow rate ratios. The neural network models closely recapitulated experimental observations (Fig. 2).

To guide experimental design and characterize the effect of device geometry on versatility and stability, we simulated the droplet generator performance space using these predictive models. This newly generated dataset had 4.2 million entries with an approximately equal representation in both dripping and jetting regimes (45% dripping and 55% jetting). The full-factorial parameter space and range details of the dataset are given in Table 1. The distribution of the predicted droplet diameters and generation rates are shown in Fig. S1.†

Versatility

Versatile droplet generators can produce a broad range of diameters and generation rates for a given range of capillary numbers and flow rate ratios. For each design, we used the predicted performance range for all flow conditions (Table 1) to establish three performance metrics: size versatility, the total range of observed droplet diameters; rate versatility, the total range of observed generation rates; and total versatility, the convex hull area of the observed diameters and generation rates (Fig. 1a). The same definitions were also used for analyzing regime-specific performance versatility in both dripping and jetting regimes, where data points were grouped according to their predicted regime and analyzed separately. A wide range of versatility scores was observed in both generation regimes (Fig. S2†), suggesting that the design parameters of a droplet generator affect its versatility.

Main effect analysis. To characterize the effect of each geometric parameter on versatility, we performed main effect analysis on the size, rate, and total versatility metric scores separately for data points in dripping, jetting, and both regimes.[23] The total versatility scores spanned 4 orders of magnitude across all devices, with the majority of designs having relatively low scores (Fig. 3a). Geometric parameters, therefore, determine the total versatility (Fig. 3b), as well as diameter and rate versatility as given in Fig. S3 and S4.† These results indicate that optimization of geometric parameters is necessary to achieve a versatile device.

In the dripping regime, increasing the orifice length increased total versatility, likely because longer orifices delay the regime change from dripping to jetting to a higher capillary number.[21] Increasing the normalized channel depth significantly reduced total versatility, potentially because deeper channels limit the maximum possible generation rate and cause the regime transition from dripping to jetting to occur at lower capillary numbers.[21] Orifice width, normalized oil inlet, and outlet widths had only minor effects on total versatility, while normalized water inlet width had a negligible effect. The effects of all geometric parameters on diameter, rate, and total versatility are established and provided in Fig. S3 and S4† and are quantified using the coefficient of determination (R^2) in Tables S2 and S3.† These results suggest that normalized channel depth and orifice size can be changed to generate droplets with a large range of diameters and a small range of generation rates, or *vice versa*. In contrast, normalized oil inlet width and normalized orifice length can tune size versatility or rate versatility without affecting the other, respectively.

In the jetting regime, geometric parameters also significantly affected versatility despite having notably different dynamics compared to the dripping regime,[24] as shown in Fig. S4.† Increasing normalized water inlet and
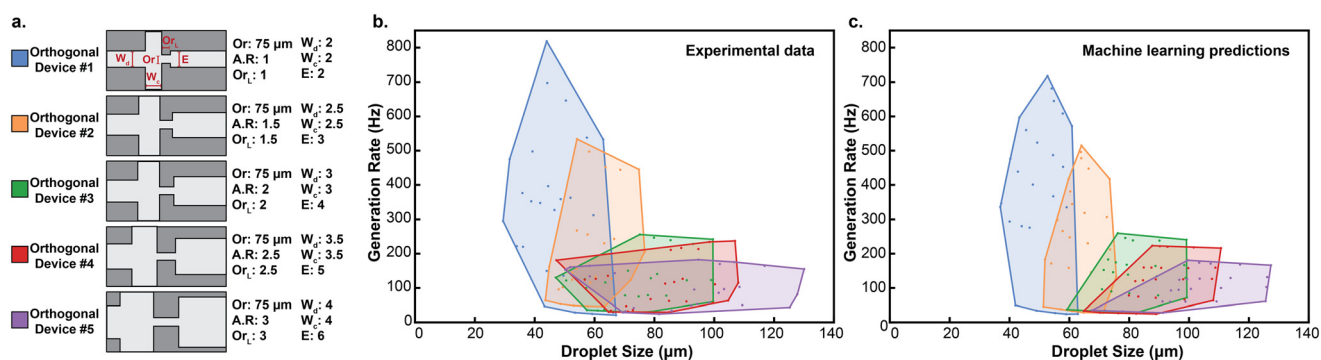

Fig. 2 (a.) To understand the effect of geometric parameters other than orifice width on droplet generation, five devices were designed with a 75 μm wide orifice and orthogonal other geometric parameters. (b.) Taking data under the same flow conditions from a previous publication,[21] drastically different performance ranges are observed from these orthogonal devices, indicating that design parameters other than orifice width influence behavior. (c.) These results can be accurately recapitulated with a machine learning-based predictive model.

Table 1 Range of design parameters, flow conditions, and the number of variations used in this study

Parameters		Range		
Name	Unit	Lower bound	Upper bound	Number of variations
Geometry				28 125
Orifice width	μm	75	175	5
Normalized[a] orifice length	N.A.	1	3	5
Normalized[a] water inlet width	N.A.	2	4	5
Normalized[a] oil inlet width	N.A.	2	4	5
Normalized[a] channel depth	N.A.	1	3	5
Normalized[a] outlet width	N.A.	2	6	9
Flow condition				150
Flow rate ratio[b]	N.A.	2	22	10
Capillary number	N.A.	0.05	1.05	15

[a] Parameters were normalized by dividing their value by the orifice width. [b] Ratio of oil flow rate to water flow rate.

outlet widths resulted in a notable decrease in total versatility. In contrast, increasing normalized oil inlet width produced a notable increase in total versatility. The influence of these parameters on versatility is likely from the resultant change in oil flow rate (for a given capillary number) and fluid acceleration through the orifice, both of which dictate droplet generation rate. This is supported by the high correlation between total versatility and generation rate versatility in the jetting regime (Fig. S4†). Medium orifice widths yielded slightly higher total versatility scores in comparison to the extremes, which can be attributed to the smaller orifices delivering a broader range of generation rates, while larger orifices deliver a wider range of diameters. Normalized channel depth and orifice length were observed to have a negligible effect on total versatility in the jetting regime. These results indicate that the design of microfluidic droplet generators can be tailored to meet user requirements in delivering a versatile performance, or delivering a wide range of generation rates while only producing a narrow range of diameters or *vice versa*.

Experimental validation. To experimentally validate the versatility quality metric, the design parameters that led to the highest and lowest mean versatility in the dripping regime (according to the main effect analysis) were used to design two flow-focusing droplet generators (see Table S1†). The two devices were fabricated and tested at the capillary number and flow rate ratio combinations at the edge of the convex hull of the simulations. If the observed droplet generation regime was in the jetting regime, the capillary number (*i.e.*, oil flow rate) was reduced until droplet generation reached the dripping regime. Excitingly, the experimentally observed performance range between the two devices was notably different, with approximately a 4-fold difference in total versatility scores (Fig. 3c–f). The more versatile device delivered a droplet diameter in the range of 27.1 to 77.2 μm and rates of 67 to 515.9 Hz. In contrast, the less versatile device generated droplets 114.3 to 329.5 μm in diameter and rates between 9 to 36.4 Hz. The observed difference for these designs results partly from a delayed regime change from dripping to jetting while increasing the capillary number, thus enabling droplet generation at higher

capillary numbers and therefore higher generation rates. Despite this larger capillary number range, the range of flow rates in the more versatile device was significantly smaller than the less versatile device. The smaller flow rate range indicates that versatility stems from the geometric parameters of the device and not just the range of flow rates that result in dripping droplet formation (Fig. 3d). Additionally, the smaller water and oil inlet widths of the versatile design suggest that geometries that further accelerate the fluid at the orifice result in a higher generation rate for a given flow rate. The small deviation of the predictive models from experimental data is expected given the small errors in our models for performance prediction and regime classification.[22] Nonetheless, the main versatility characteristics are upheld (Fig. 3c, right three panels), instilling confidence that our models are capturing the high-level behavior of the device.

To demonstrate the utility of the versatility metric in practice, we fabricated two devices capable of producing 100 μm diameter droplets at 100 Hz with predicted performances such that the versatile device's performance range completely encompassed that of the less versatile device (Fig. 4; see Table S1† for design parameters). As predicted, the versatile device had a significantly higher versatility compared to the less versatile device and encompassed 98.2% of its deliverable performance area, while the less versatile design only encompassed 49.2% of the deliverable performance area of the more versatile design (Fig. 4). The more versatile device could operate in the dripping regime within a wider range of flow rates; nonetheless, limiting the observed performance to similar flow rates (*i.e.*, same as the less versatile device) still resulted in a much larger performance range in the more versatile device (Fig. S5†).

Stability

During droplet generation, different fluidic regimes occur depending on the device geometry and flow rates of the operating point. A previously validated machine learning model was used to predict the generation regime for the 4.2 million data points.[22] These data points were then labeled to
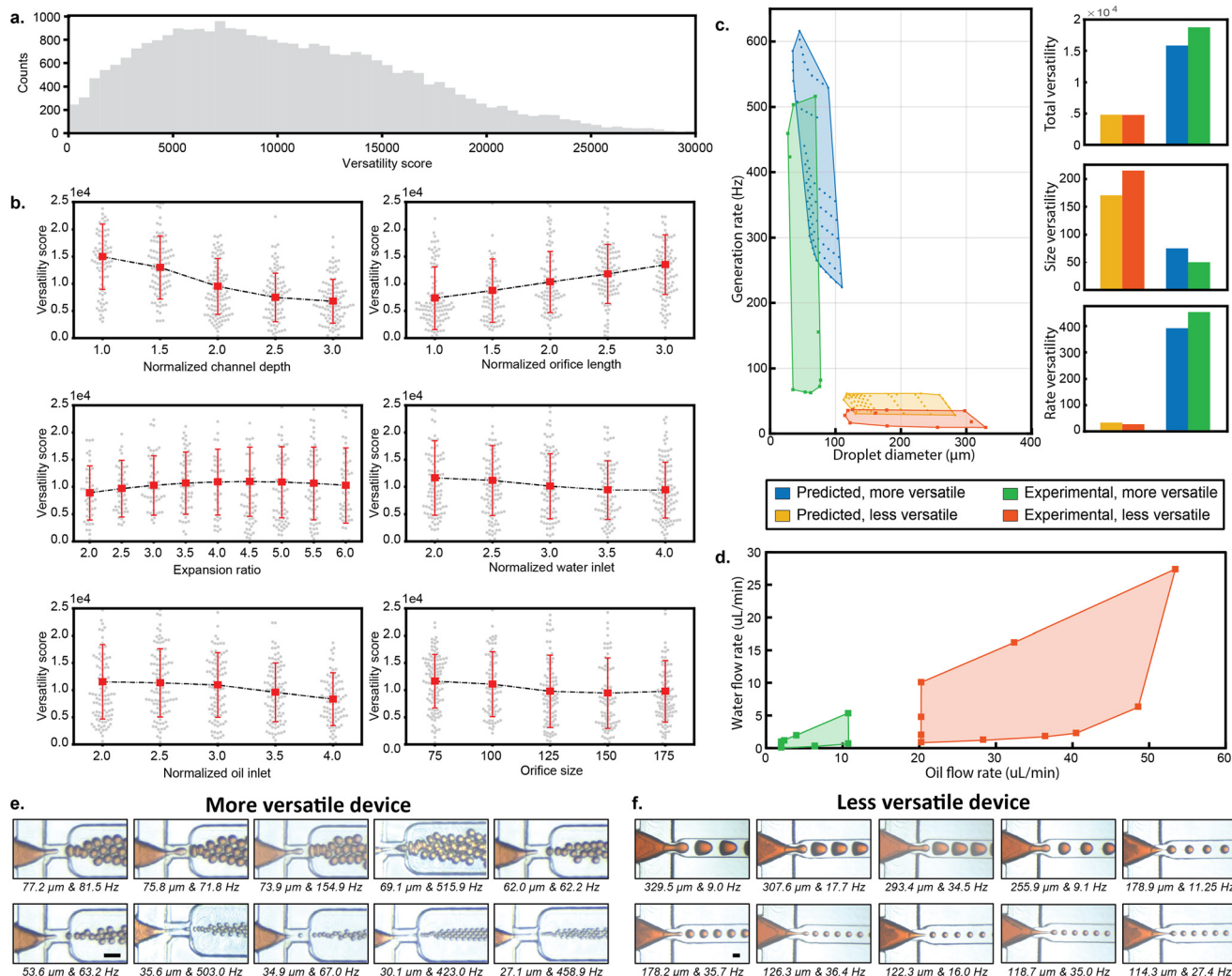
Fig. 3 The geometric parameters of a flow-focusing device can be adjusted to tune device versatility. (a.) The versatility score of a large-scale simulated dataset was calculated (28 125 devices) to produce a wide range of scores. (b.) Main effect analysis is used to quantify the effect of variations in geometric parameters on performance versatility for both dripping and jetting regimes. The effect of geometric parameters on the total versatility (the convex hull area of possible droplet diameters and rates) for the dripping regime is shown. The effect of geometric parameters on droplet diameter versatility, generation rate versatility, and total versatility in both regimes are provided in Fig. S3 and S4.† The coefficient of determination (R^2) values for all parameters are provided in Tables S2 and S3.† (c.) Based on the main effect analysis, two droplet generators were designed using the parameters that resulted in the highest or lowest versatility. The performance of these devices was tested experimentally and predicted within the dripping regime and was shown to behave as expected both experimentally and *in silico*. (d.) The more versatile device exhibited a larger deliverable performance space in the dripping regime with a smaller range of flow rates. (e. and f.) Images of experimental results. Scale bars are 100 μm.

be a regime boundary or not. A regime boundary was assigned if an adjacent point (± one step in flow rate ratio or capillary number) had a different predicted regime. Then, a "stability score" was assigned to each point as the Euclidean distance in flow rate values to a boundary point. Points on the boundary line were assigned a score of zero (see Methods, stability scoring for more detail).

For a given device geometry, a change in the capillary number or flow rate ratio would result in a 1-to-1 ratio change in oil flow rate or water flow rate according to eqn (2). Therefore, for a given device geometry, using either flow rates or capillary number and flow rate ratio would result in the same relative Euclidean distance. Additionally, since

operating errors typically occur in units of flow rates (*e.g.*, ±1 μL per hour), the Euclidean distance in flow rates was used instead of capillary number and flow rate ratio to develop an unbiased quality metric for different devices operating at either high or low flow rates.

Main effect analysis. Within the created dataset, a wide range of stability scores that spanned an order of magnitude was calculated (Fig. S6†). The majority of data points were observed to have a relatively low stability score, further emphasizing the need to characterize and optimize stability. Main effect analysis was performed on the stability scores of the 4.2 million data points to estimate the geometric design parameters that had the most influence over stability.
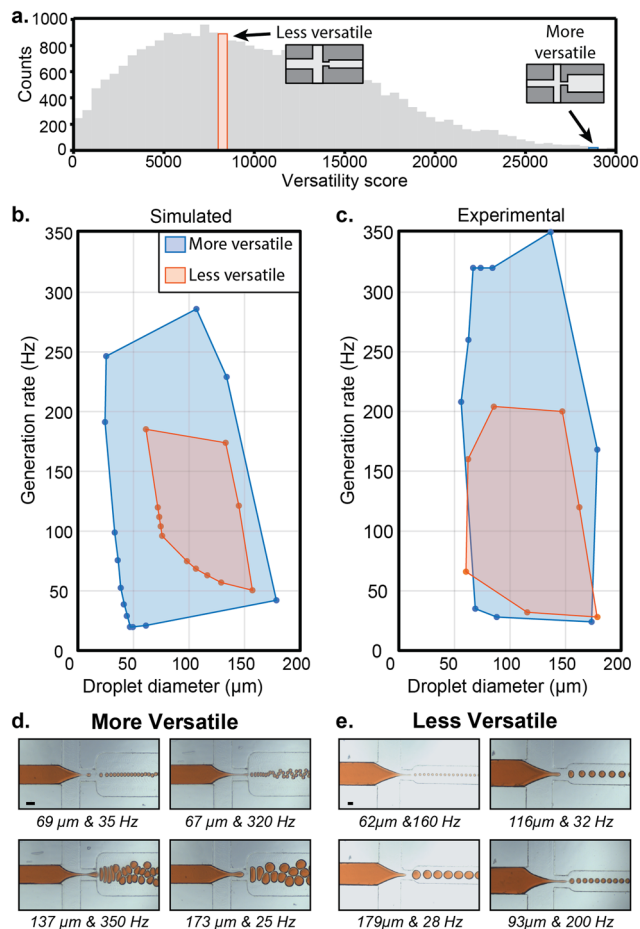
Fig. 4 Direct comparison of versatile and non-versatile droplet generators. (a.) Two droplet generators capable of generating 100 μm diameter droplets at 100 Hz but drastically different versatility scores in the dripping regime and device geometry were manually selected. (b.) The predicted performance range of the more versatile design fully encompassed the deliverable performance range of the less versatile design. (c.) The more versatile design delivered an experimentally observed performance convex hull area approximately two times larger than the less versatile design while almost completely encompassing its performance space in the dripping regime. (d. and e.) Images of experimental results.

In the dripping regime, increasing capillary number caused the largest decrease in stability values (Fig. 5a). This is expected, as the regime transition from dripping to jetting occurs as the capillary number increases. Increasing orifice size increased the stability score; this may be attributed to a reduced flow acceleration through larger orifices for a given change in flow rates that delays the regime change from dripping to jetting. Similar to orifice width, it is expected that larger normalized channel depths (*i.e.*, deeper channels) increase stability by reducing flow acceleration through the orifice. However, larger normalized depths also facilitate regime change from dripping to jetting at lower capillary numbers. Therefore, normalized depth minimally affects stability overall. A negligible effect on dripping regime stability was observed in other design parameters.

In the jetting regime, an inverse effect of capillary number on stability was observed, as the chance of a regime boundary (*i.e.*, a transition from jetting to dripping) decreases drastically as capillary number increases (see Fig. S7†). Increasing orifice width, normalized depth, and oil inlets increased stability in the jetting regime. This can be attributed both to a lower sensitivity to changes in flow rates due to a smaller flow acceleration at the orifice and to the positive correlation of these parameters with oil flow rate for a given capillary number and flow ratio.

Experimental validation. The stability metric was experimentally validated by fabricating two devices with design parameters that resulted in the highest and lowest average stability scores in the dripping regime according to the main effect analysis shown in Fig. 5a (see Table S1† for design parameters). Between capillary numbers of 0.05 and 0.27 and flow rate ratios of 2 and 22, predictions for each device showed a much larger dripping performance space in the more stable device (206 out of 230 points; see Fig. 5b, left) compared to the less stable device (29 out of 230 points; see Fig. 5c, left). Experimentally, the regime boundary was found by increasing the capillary number and flow rate ratio until a regime change was observed. As predicted, experimental regime boundaries showed a similar difference; although the regime boundary was not exactly the same, the dripping performance space of the more stable device was much larger (167 out of 230 points; Fig. 5b, right) compared to the less stable device (47 out of 230 points; Fig. 5c, right).

This discrepancy between experimental and simulated data is to be expected: the architecture of our predictive tool utilizes separate models for each regime, and therefore datapoints on the regime boundary are at the edge of the training set distribution. Furthermore, experimental data close to the regime boundary are inherently unstable, with droplet generation regularly changing between each regime due to small changes in flow rates, device fabrication, surface properties, or operation. Therefore, we would expect some discrepancy between the simulated and experimental data around the regime change boundary. For instance, the more stable device and less stable device were predicted to deliver similar performance at a capillary number of 0.05 and flow rate ratio of 15.3 (103 μm size, 45 Hz rate for the less stable device, 110 μm size, 38 Hz rate for the more stable device). When running these devices experimentally, a generation rate of 42 Hz is observed for both devices, however, a droplet size of 63.6 μm (38% experimental error) is observed for the less stable device, while the more stable device delivered a diameter of 92.57 μm (16.1% experimental error), as shown in Fig. 5d. This discrepancy is therefore mitigated by the introduction of a stability metric: by generating stable points, the user can be assured that the datapoint is far away from the regime boundary and thus errors from fabrication, testing, or predictive models are limited.
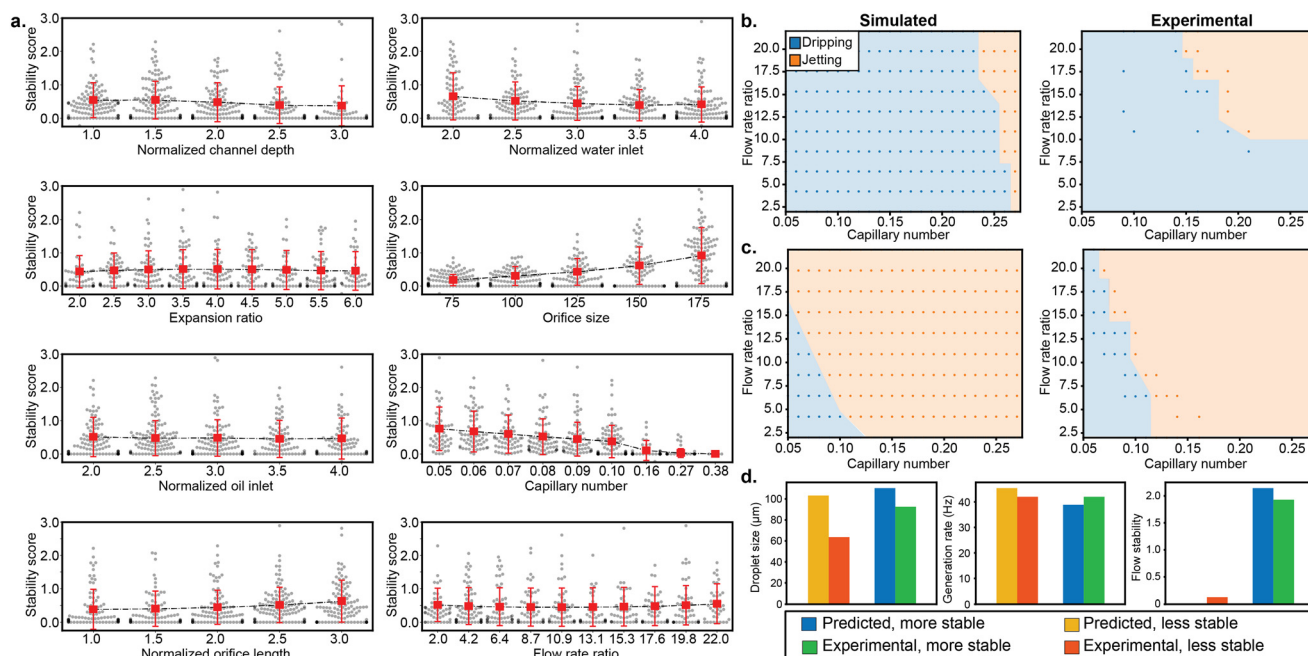
Fig. 5 The design parameters of a flow-focusing device can be adjusted to tune its stability. (a.) Main effect analysis is used on 4.2 million data points to quantify the effect of variations of geometric and flow parameters on stability. The effect of parameters on stability for just the dripping regime is shown here, and the effect of design parameters in the jetting regime is provided in Fig. S7.† The coefficient of determination (R^2) values for all parameters are provided in Table S2.† To validate the findings on stability based on the main effect analysis, two droplet generators were designed using the geometric parameters that resulted in (b.) the highest or (c.) lowest stability. (b.–d.) The performance of these devices was both tested experimentally and predicted for the dripping regime and was shown to behave as expected both experimentally and *in silico*.
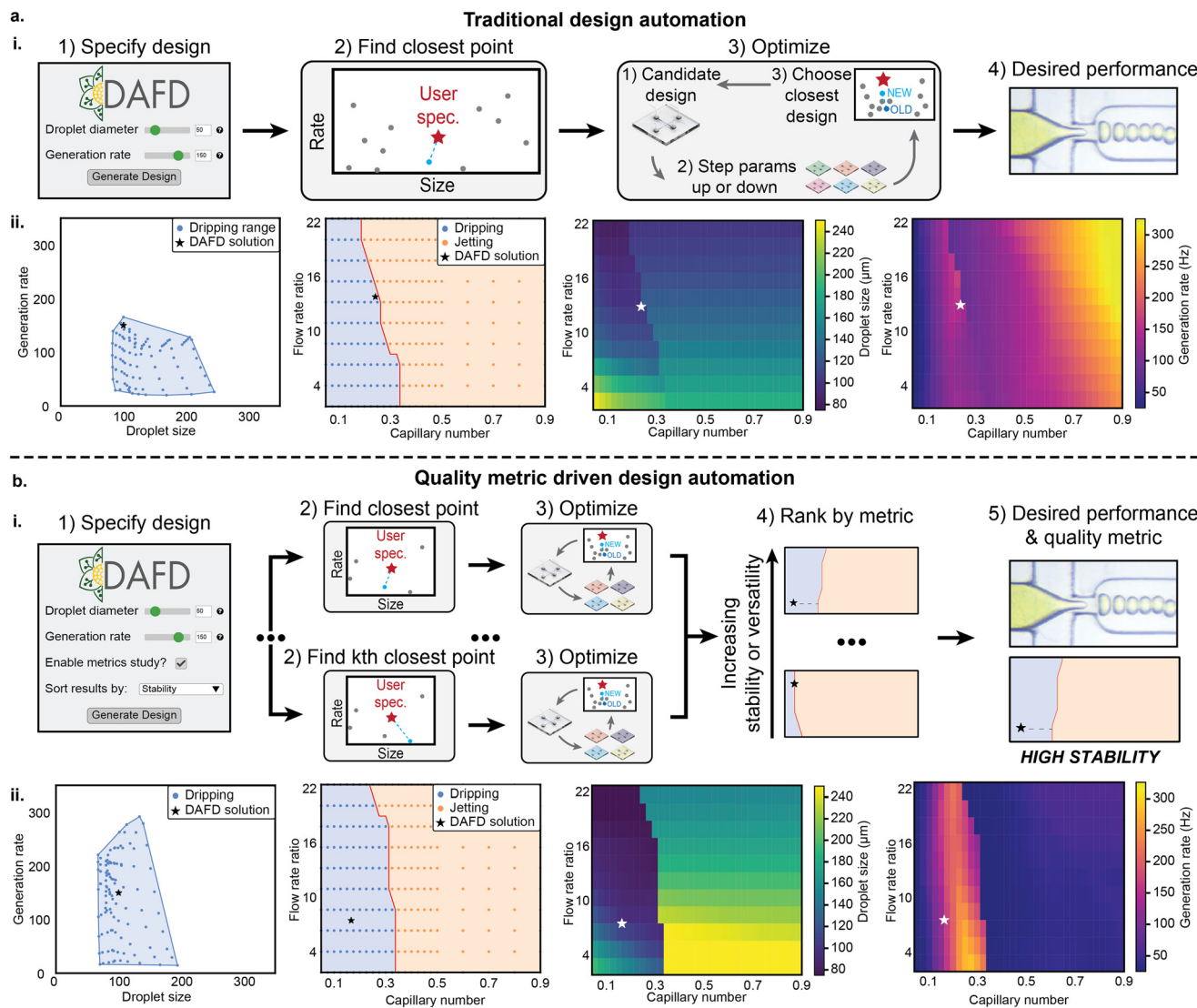
Design automation integration

To facilitate the utility of quality metrics in the microfluidics community, both versatility and stability were implemented in a new design automation algorithm to develop DAFD 2.0.[22] In the previous version, DAFD automated droplet generator design by picking the closest experimental point in the dataset and then making adjustments to the design parameters, iterating until the difference between the specified and predicted performance is within a set threshold (Fig. 6a). In DAFD 2.0, a user-specified number of closest points are selected and optimized in parallel to produce multiple candidate results (Fig. 6b). The closest points are ignored before optimization if they have the same geometric features as those already chosen. Optimized points are scored by versatility or stability and then ranked according to the user-specified quality metric. Thus, the resulting point would have both the desired behavior and a contextual understanding of its versatility or stability. If a user would like to optimize by both versatility and stability, we recommend that the user first optimize by versatility to get a specific device with a broad operating range, and then optimize by stability, fully constraining the design automation algorithm to the previous solution's geometric design parameters. This would then give the user an output with both high versatility and stability. A companion report is generated to report different metric scores and visualize the user's device information on its deliverable diameter and

generation range in each regime and the operating regime of the device based on its capillary number and flow rate ratio (Fig. S8†). An additional sensitivity analysis (which would evaluate the changes in droplet size or generation rate as any of the input parameters are changed) can be generated by the user *via* a previously developed "Tolerance Test."[22]

The efficacy of design automation with quality metrics was validated by comparing design automation results for 100 μm diameter droplets generated at 150 Hz in the dripping regime. Two solutions were generated by using either traditional design automation or design automation with quality metrics. Both solutions had a predicted performance within 1 μm or 1 Hz of the specified droplet size and rate, respectively (Fig. 6ai and bi). In DAFD 2.0, the same device ranked highest in both overall versatility and stability scores. The quality-metric driven solution had a predicted versatility of 24 796, 40% higher than the default solution (see Fig. 6aii and bii), as well as approximately a 6-fold higher stability score (0.71 for the ranked device, 0.13 for the original solution).

A regime change can cause a significant change in the observed performance, as shown in Fig. 6aii and bii, demonstrating the importance of the stability metric to improve the robustness of droplet generation against small perturbations in fabrication and testing. Quality metric integration can play an important role in improving the accuracy and reproducibility of microfluidic design automation while increasing the understanding and

Fig. 6 The newly established quality metrics were integrated into the DAFD design automation tool to deliver the desired droplet size and rate while maximizing the user-specified quality metric. (ai.) Previously, a user would specify a desired performance, and design automation would start by selecting the closest experimental point (*i.e.*, diameter and rate) in the dataset and iteratively optimizing the design parameters until the desired performance was achieved. (aii.) This approach can generate designs that have a narrow range of diameters and rates that are less ideal for integration with other microfluidic components or operate close to a regime boundary, making it susceptible to large changes in performance from small errors. (bi.) In the newly developed metric-driven design automation tool (*i.e.*, DAFD 2.0), the "top-*k*" closest data points are selected from the dataset, simultaneously optimized to achieve the desired performance, and then ranked according to the user-specified quality metric (*i.e.*, performance versatility or stability). (bii.) The candidate with the highest quality metric is then selected, producing a desired behavior with maximized quality metric. While inputting the same performance of 150 μm droplets at 100 Hz into DAFD and specifying versatility and stability as quality metrics, the suggested design was observed to deliver a broader range of possible performance. The suggested design and operating point were also farther away from the boundary of regime change, thus making it more robust against small errors in fabrication and testing.

accessibility of microfluidics to the broader research community. DAFD 2.0 is freely available to users at **http://dafdcad.org** and the source code can be found at **https://github.com/CIDARLAB/DAFD_Metrics**.

Discussion

In this study, we introduced versatility and stability, two quality metrics for microfluidic flow-focusing droplet generators. Operating points with a high stability score are

far away from a fluidic regime change (*e.g.*, from dripping to jetting) and therefore are more robust to performance errors from small fabrication or flow-based errors as well as inaccuracies in predictive models. Stable operating points can be particularly helpful for multi-component devices, where droplet size and generation rate are less affected by pressure fluctuations or the required operating conditions of other components.[8] Versatile droplet generators are capable of delivering a wider range of diameters and generation rates for a given range of flow conditions. Versatile designs could

be used for rapid data generation, minimizing the number of devices needed to explore the output space and reducing time and cost requirements. For example, the same performance space that was previously mapped with 25 orthogonal devices[21] can be mapped with only 5 versatile devices with more than 99% coverage, a substantial reduction in the number of devices that have to be fabricated and tested (Fig. S9†). While helpful, manual analysis of quality metrics can be time-consuming and require significant expertise to understand what parameters can be changed to improve each metric while adhering to the desired droplet size and rate. To this end, versatility and stability were integrated into the DAFD design automation software to generate device designs that can both deliver user-specified performance as well as maximizing versatility or stability.

These metrics can facilitate the integration of droplet generators with other microfluidic components, selecting for large overlaps in their operating range or high stability, reducing the need for multiple devices to achieve different droplet properties. Double emulsion generation is one application where versatility and stability can be used in tandem to optimize performance.[25] In double emulsion generation, matching the generation rates at the two flow-focuser junctions is essential for producing single-core double emulsions.[4,26] Minimizing size differences between the inner and outer emulsion (*i.e.*, oil shell thickness) can also be essential if double emulsions are going to be processed in a size-restricted system such as a commercial FACS machine.[25] High size versatility and low rate versatility in both of the linked droplet generators would enable the generation of different inner and outer diameters while limiting changes in generation rates that make unwanted products (*e.g.* multiple or no cores). High stability in performance would avoid failure modes stemming from a regime change in either of the two flow-focusing droplet generators.

Although these metrics are created and validated with a droplet generation dataset with DI water and mineral oil in a polycarbonate device, transfer learning could be used to expand DAFD and the developed metrics to other fluids and fabrication methods such as cell media, fluorinated oils, and PDMS. In transfer learning, a small-scale dataset is used to tune an existing model previously generated on a larger dataset. The overall fluid dynamics of droplet generation are expected to be similar even when different fluids are used, therefore, leveraging geometry to improve versatility and stability is also expected to hold true for different fluids.

To this end, we evaluated the versatility and stability of the devices used in Fig. 4 and 5 with a different fluid combination, switching out DI water for lysogeny broth (LB) bacterial cell media, which is 80% more viscous than DI water (1.8 mPa s).[27] In the versatile and less versatile devices, a marked difference in dripping versatility was observed, despite a reduction in versatility in both cases compared to DI water and mineral oil (Fig. 7a). A cause of this versatility reduction was observed in the stability devices: in both cases, a significant shift to lower capillary numbers was observed in the regime change boundary. Despite this shift, the more stable device exhibited a larger number of fluid conditions in the dripping regime than the less stable device (Fig. 7b).

These results suggest both versatility and stability can be generalized to other fluid combinations; however, as the fluid combinations begin to differ more significantly from DI water and mineral oil, we anticipate the conservation of such properties within the same geometric designs to be limited. Additional machine learning models capable of predicting performance and regime change across fluid combinations are needed. Machine learning has been used to extend these predictive models to fluids commonly used in life science applications in a fluid-agnostic manner, broadening the resource of droplet generator design automation across the microfluidic community.[26] Currently, such models are unable
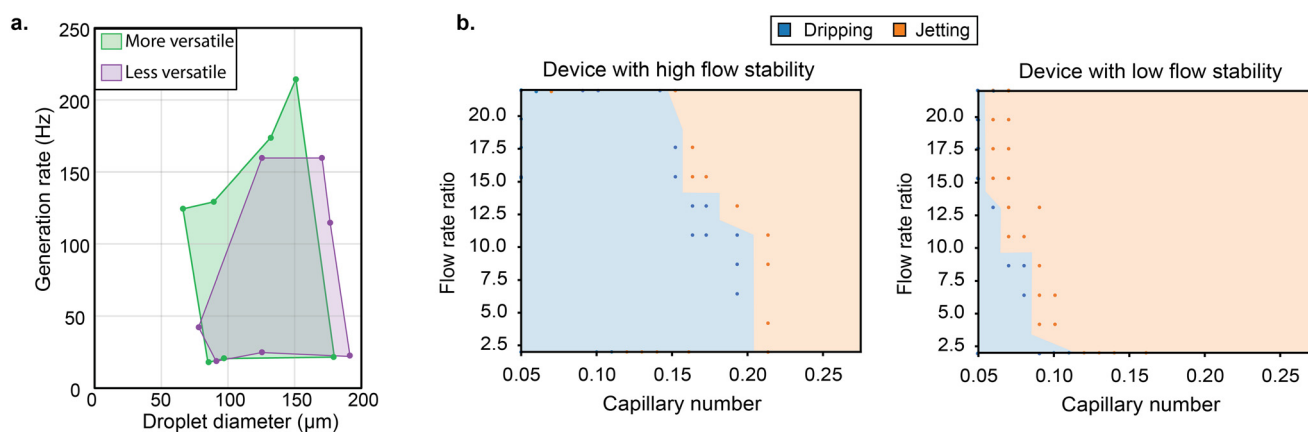
Fig. 7 Generalization of versatility and stability to novel fluid combinations (LB media and mineral oil). (a.) Using the droplet generators from Fig. 4, differences in versatility were still apparent between the more and less versatile devices. (b.) With the droplet generators from Fig. 5, a marked shift in regime change boundary was observed in both devices. However, relative stability in the dripping regime between the more and less stable devices.

to predict regime changes in different fluid types, limiting the adoption of the presented quality metrics.

By combining machine learning,[12] device standardization,[10] and both rapid and high-resolution fabrication techniques,[28–34] metric-driven microfluidic design automation can be applied across fluid combinations and droplet microfluidic component libraries and play an important role in reducing the barrier to adoption in microfluidics.

Conclusions

The development of versatility and stability quality metrics was made possible by leveraging large-scale predictions using machine learning-based models for flow-focusing droplet generators. These metrics were also experimentally validated by fabricating and testing devices with high and low versatility and stability scores. The quality metric-driven devices were demonstrated to significantly improve versatility and stability when compared to traditionally designed devices. The use of both quality metrics by the broader community was made available through integration with an open-source and online design automation tool, DAFD, which now generates user-specified performance while also optimizing for stability or versatility. To our knowledge, this is the first integration of quality metrics in droplet microfluidics that are specifically made to improve the reproducibility and robustness of droplet generators while reducing design iterations and facilitating integration with downstream microfluidic components.

Methods

Simulation of droplet generation with machine learning-based predictive models using the DAFD tool

Droplet generation with DI water and NF 350 mineral oil was simulated using the "performance prediction module of the DAFD (design automation of fluid dynamics) tool previously developed by our group.[22] In brief, neural networks predicting droplet diameter and generation rate across two generation regimes of dripping and jetting (four models in total) are trained on a large experimental dataset consisting of 888 data points from 43 unique flow-focusing devices. These models can then be used to predict output droplet size (25–250 μm) and generation rate (5–500 Hz) from input design parameters and flow conditions. DAFD was used to simulate a full-factorial design space of the input parameters, totaling approximately 4.2 million data points from 28 125 unique flow-focusing geometries (Table 1). The 150 flow conditions include 10 flow rate ratios between 2 and 22 and 15 capillary numbers, comprised of 6 evenly spaced capillary numbers between 0.05 and 0.1 and 9 evenly spaced capillary numbers between 0.161 and 1.05. These capillary numbers were chosen to give a roughly equal distribution of data points in the dripping and jetting regimes. The models can be accessed online at **http://dafdcad.org/** and the source code

https://github.com/CIDARLAB/DAFD/ used in this study can be generated at **https://github.com/CIDARLAB/DAFD_Metrics**.

Versatility scoring

To find the versatility of a device, the 2D convex hull of the performance space (droplet size and generation rate) was calculated using the SciPy spatial library (**https://www.scipy.org/**).[35] The total versatility score for the device was calculated as the area of the convex hull. Droplet size and generation rate versatility scores were calculated by their respective ranges (maximum predicted value minus minimum predicted value). Any devices with less than 3 points in a droplet generation regime are given a versatility score of −1 and are excluded from downstream analysis as a convex hull cannot be formed. This was repeated using points in the dripping or jetting regime to generate regime-specific versatility scores.

Stability scoring

The stability of a single point is found by first labeling droplet generation regime boundary points within the device's performance space that has an adjacent point of a different regime (from a step in capillary number or flow rate ratio). Next, the Euclidean distance (d) in oil and water flow rates in μL min^{-1} from the point in question to each boundary point was calculated, where:

$$d = \sqrt{\left(Q_{\text{oil\_boundary}} - Q_{\text{oil\_point}}\right)^2 + \left(Q_{\text{water\_boundary}} - Q_{\text{water\_point}}\right)^2}$$

(1)

The stability score is then set as the minimum distance to a boundary point.

Metric integration with design automation

The design of flow-focusing droplet generators is automated by finding the closest experimental point to the user specification, as described previously.[22] Next, this starting point is optimized by increasing and decreasing each of the eight design parameters to produce 16 candidate designs. The design with a predicted performance that is closest to the desired performance is chosen and this process is repeated until the predicted point is below a set threshold value. User constraints can be added to limit the search space. The experimental starting point is returned if it is already within the threshold.

Metric-driven design automation is achieved by using a similar algorithm. Rather than picking one starting point, user-specified "top-k" closest points are chosen. To guarantee a diversity of candidates, new points are not considered if they are within 10 μm or 0.25 from the orifice size or normalized geometric parameters of the previous point, respectively. The multiple starting positions are then simultaneously optimized in the same way as the previous

version. Once optimized, the versatility and stability scores are calculated. A total of 230 flow conditions are used to define device boundaries, consisting of 10 flow rate ratios evenly spaced between 2 and 22 as well as 23 capillary numbers comprised of 18 evenly spaced points between 0.05 and 0.5 and 5 evenly spaced points between 0.5 and 1.

Candidates are then ranked by the metric specified by the user, with the most versatile or stable point recommended. All candidate points are available in a separate .csv file that users can download for reference. The source code for metric-driven design automation is available at https://github.com/CIDARLAB/DAFD_Metrics/.

Main effect analysis

As described previously[23] and used before on a similar dataset,[21] main effect analysis was used to approximate the relative influence that each design parameter had on stability and versatility. The 4.2 million data points were binned into each unique input parameter, and the average metric value for each bin was calculated. The effect of each value is quantified using the correlation coefficient (R^2).

Microfluidic fabrication and operation

Microfluidic geometries were directly etched into polycarbonate slabs using a desktop CNC micromill (Bantam Tools), as described previously.[29] Once etched, devices are cleaned with first sonication in IPA and DI water and then a soft brush. Next, devices are sealed with an 81 μm thick double-sided adhesive (Adhesives Research ArCare 90445). Microfluidic devices are then placed in a vacuum to remove any air bubbles and ensure proper bonding between the adhesive and the device layers.

Droplet generation with colored DI water and NF 350 mineral oil with 5% V/V Span 80 surfactant (Sigma Aldrich) was actuated with syringe pumps (Harvard Apparatus). Images were captured with a high-speed camera (IDT X-Stream) mounted to an inverted microscope (Zeiss). For illumination at high frame rates, an 18 000 lumen LED light source (Expert Digital Imaging) is used. The droplet size and generation rate of each experiment were measured by manually analyzing droplet generation videos to measure the generation rate, and subsequently calculate the diameter using the water flow rate and conservation of mass equation.

Flow rate calculation

Capillary number and flow rate ratio are commonly used as dimensionless flow parameters to describe and characterize the fluid flow in flow-focusing droplet generators.[24] Here, the flow rate of water and oil are calculated based on the capillary number, flow rate ratio, fluid properties, and device geometry, as given in eqn (2):

$$Q_{oil} = \frac{Ca. \cdot \sigma \cdot H \cdot W_{oil}}{\mu_{oil} W_{water} \left[\dfrac{1}{Or.} - \dfrac{1}{2 W_{oil}} \right]}$$

$$Q_{water} = \frac{Q_{oil}}{\phi}, \tag{2}$$

where Q_{oil} is the oil flow rate, Ca. denotes capillary number, σ represents the surface tension between the continuous and dispersed phases, H is channel depth, μ_{oil} denotes dynamic viscosity of oil, ϕ represents flow rate ratio, and W_{water}, W_{oil}, and Or. are water inlet, oil inlet, and orifice widths, respectively. The viscosity of the NF 350 mineral oil is 57.2 mPa s and the surface tension between the oil and DI water is 0.005 N m^{-1}.[36] The flow rate ratio and the flow rates for both water and oil can be readily calculated using eqn (2).

Data and source-code availability

All the data generated in this study and the original 888 experimental data points are freely available at https://dafdcad.org. Additionally, all code used in this study is available at https://github.com/CIDARLAB/DAFD_Metrics.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 Y. Wang, R. Jin, B. Shen, N. Li, H. Zhou, W. Wang, Y. Zhao, M. Huang, P. Fang and S. Wang, *et al.*, *Sci. Adv.*, 2021, **7**, eabe3839.

2 A. H.-H. Wong, H. Li, Y. Jia, P.-I. Mak, R. P. da Silva Martins, Y. Liu, C. M. Vong, H. C. Wong, P. K. Wong and H. Wang, *et al.*, *Sci. Rep.*, 2017, **7**, 1–15.

3 P. C. Gach, K. Iwai, P. W. Kim, N. J. Hillson and A. K. Singh, *Lab Chip*, 2017, **17**, 3388–3400.

4 K. K. Brower, M. Khariton, P. H. Suzuki, C. Still, G. Kim, S. G. Calhoun, L. S. Qi, B. Wang and P. M. Fordyce, *Anal. Chem.*, 2020, **92**, 13262–13270.

5 A. J. de Mello and N. Beard, *Lab Chip*, 2003, **3**, 11N–20N.

6 M. Khan, S. Mao, W. Li and J.-M. Lin, *Chem. – Eur. J.*, 2018, **24**, 15398–15420.

7 S. G. Calhoun, K. K. Brower, V. C. Suja, G. Kim, N. Wang, A. L. McCully, H. Kusumaatmaja, G. G. Fuller and P. M. Fordyce, *Lab Chip*, 2022, **22**, 2315–2330.

8 A. H. Ng, S. Peng, A. M. Xu, W. J. Noh, K. Guo, M. T. Bethune, W. Chour, J. Choi, S. Yang and D. Baltimore, *et al.*, *Lab Chip*, 2019, **19**, 3011–3021.

9 E. E. Tsur, *Annu. Rev. Biomed. Eng.*, 2020, **22**, 285–307.

10 D. R. Reyes, H. van Heeren, S. Guha, L. Herbertson, A. P. Tzannis, J. Ducrée, H. Bissig and H. Becker, *Lab Chip*, 2021, **21**, 9–21.

11 S. Battat, D. A. Weitz and G. M. Whitesides, *Lab Chip*, 2022, **22**, 530–536.

12 D. McIntyre, A. Lashkaripour, P. Fordyce and D. Densmore, *Lab Chip*, 2022, **22**, 2925–2937.

13 L. Rosenfeld, T. Lin, R. Derda and S. K. Tang, *Microfluid. Nanofluid.*, 2014, **16**, 921–939.

14 A. Grimmer, W. Haselmayr, A. Springer and R. Wille, *Proceedings of the 54th Annual Design Automation Conference 2017*, 2017, pp. 1–6.

15 M. Hamidović, W. Haselmayr, A. Grimmer, R. Wille and A. Springer, *Nano Commun. Netw.*, 2019, **19**, 33–46.

16 S. Wiedemeier, M. Eichler, R. Römer, A. Grodrian, K. Lemke, K. Nagel, C.-P. Klages and G. Gastrock, *Eng. Life Sci.*, 2017, **17**, 1271–1280.

17 S. Lignel, A.-V. Salsac, A. Drelich, E. Leclerc and I. Pezron, *Colloids Surf., A*, 2017, **531**, 164–172.

18 C. Ren and A. Lee, *Droplet Microfluidics*, The Royal Society of Chemistry, 2021, pp. 1–14.

19 S. Xu, Z. Nie, M. Seo, P. Lewis, E. Kumacheva, H. A. Stone, P. Garstecki, D. B. Weibel, I. Gitlin and G. M. Whitesides, *Angew. Chem.*, 2005, **117**, 734–738.

20 W. Lee, L. M. Walker and S. L. Anna, *Phys. Fluids*, 2009, **21**, 032103.

21 A. Lashkaripour, C. Rodriguez, L. Ortiz and D. Densmore, *Lab Chip*, 2019, **19**, 1041–1053.

22 A. Lashkaripour, C. Rodriguez, N. Mehdipour, R. Mardian, D. McIntyre, L. Ortiz, J. Campbell and D. Densmore, *Nat. Commun.*, 2021, **12**, 1–14.

23 D. C. Montgomery, *Design and analysis of experiments*, John wiley & sons, 2017.

24 T. Fu, Y. Wu, Y. Ma and H. Z. Li, *Chem. Eng. Sci.*, 2012, **84**, 207–217.

25 K. K. Brower, C. Carswell-Crumpton, S. Klemm, B. Cruz, G. Kim, S. G. Calhoun, L. Nichols and P. M. Fordyce, *Lab Chip*, 2020, **20**, 2062–2074.

26 A. Lashkaripour, D. P. McIntyre, S. G. Calhoun, K. Krauth, D. M. Densmore and P. M. Fordyce, *bioRxiv*, 2023, preprint, DOI: 10.1101/2023.05.31.543128.

27 K. Sklodowska, P. R. Debski, J. A. Michalski, P. M. Korczyk, M. Dolata, M. Zajac and S. Jakiela, *Micromachines*, 2018, **9**, 251.

28 A. K. Au, W. Huynh, L. F. Horowitz and A. Folch, *Angew. Chem., Int. Ed.*, 2016, **55**, 3862–3881.

29 A. Lashkaripour, R. Silva and D. Densmore, *Microfluid. Nanofluid.*, 2018, **22**, 1–13.

30 D. J. Guckenberger, T. E. de Groot, A. M. Wan, D. J. Beebe and E. W. Young, *Lab Chip*, 2015, **15**, 2364–2378.

31 H. Klank, J. P. Kutter and O. Geschke, *Lab Chip*, 2002, **2**, 242–246.

32 J. H. So and M. D. Dickey, *Lab Chip*, 2011, **11**, 905–911.

33 D. McIntyre, A. Lashkaripour and D. Densmore, *Lab Chip*, 2020, **20**, 3690–3695.

34 A. Sciambi and A. R. Abate, *Lab Chip*, 2014, **14**, 2605–2609.

35 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt and SciPy 1.0 Contributors, *Nat. Methods*, 2020, **17**, 261–272.

36 P. S. Dittrich, M. Jahnz and P. Schwille, *ChemBioChem*, 2005, **6**, 811–814.