Experiments and Computations of Microfluidic Liquid–Liquid Flow Patterns

SCHOLARONE™
Manuscripts

Table of Contents Entry

Experiments and Computations of Microfluidic Liquid-Liquid Flow Patterns

Pierre Desir[δ,a], Tai-Ying Chen[δ,a], Mauro Bracconi[c], Basudeb Saha[b], Matteo Maestri[c], and Dionisios G. Vlachos[a,b,*]

[a]Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, Delaware 19716, United States

[b]Catalysis Center for Energy Innovation, 221 Academy Street, Newark, Delaware 19716, United States

[c]Laboratory of Catalysis and Catalytic Processes, Dipartimento di Energia, Politecnico di Milano, via La Masa 34, 20156 Milano, Italy

[δ]These authors contributed equally.

[*]Corresponding author; vlachos@udel.edu (D.G. Vlachos)

A high accuracy model is built using machine learning to predict flow patterns, providing a powerful tool for continuous flow microreactor design.

Experiments and Computations of Microfluidic Liquid-Liquid Flow Patterns

Pierre Desir[δ,a], Tai-Ying Chen[δ,a], Mauro Bracconi[c], Basudeb Saha[b], Matteo Maestri[c], and Dionisios G. Vlachos[a,b,*]

[a]Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, Delaware 19716, United States
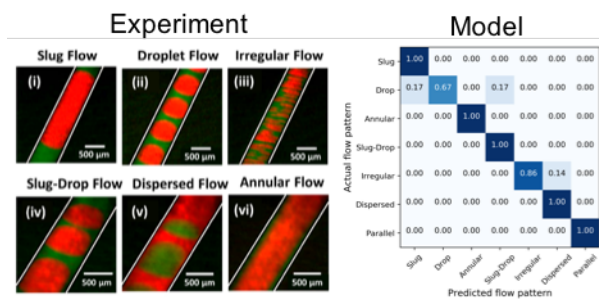
[b]Catalysis Center for Energy Innovation, 221 Academy Street, Newark, Delaware 19716, United States

[c]Laboratory of Catalysis and Catalytic Processes, Dipartimento di Energia, Politecnico di Milano, via La Masa 34, 20156 Milano, Italy

[δ]These authors contributed equally.

[*]Corresponding author; vlachos@udel.edu (D.G. Vlachos)

Abstract

We study two-phase liquid-liquid flow patterns in a 500 μm capillary microfluidic for four biphasic systems: ethyl acetate/water, 2-pentanol/water, methyl isobutyl ketone/water, and heptane/water. Flow visualization experiments using laser induced fluorescence (LIF) reveal a total of 7 different flow patterns for all solvent pairs, namely slug flow, droplet flow, slug-droplet flow, parallel, annular, dispersed, and irregular flow. A map of different flow patterns was built to delineate the origin of their formation. We find conventional dimensionless groups are insufficient to uniquely identify the flow patterns. Computational fluid dynamics (CFD) modeling in OpenFOAM shows agreement with the experimental flow patterns for most of the two-phase flows. Principal component analysis reduces the dimensionality of potential descriptors of flow patterns and, unlike prior work using two dimensionless numbers, determines six important features that describe >95% of the variance of the experimental flow patterns. These include the total flow rate, the flow rate ratio between the two phases, the capillary and Ohnesorge numbers of the aqueous phase, and the Weber number and velocity of the organic phase. We build a decision-tree model to further regress the data and identify the critical features and demonstrate an accuracy in predicting the flow patterns of up to 93%.

Keywords: Biomass, computational fluid dynamics, flow pattern, laser induced fluorescence, machine learning, two-phase flow

1

1 **Introduction**
2   The advances in microfluidic technology have enabled the effective miniaturization of
3 chemical processes. Continuous flow microreactors with characteristic length scales in the
4 micrometer range (hydraulic diameters <1 mm) possess large surface-to-volume ratios, which
5 result in heat and mass transport rates orders of magnitude greater than bench-scale and
6 conventional reactors. The laminar flow regime and short diffusion time-scale of microreactors
7 allow operation with more precise process windows including heating profiles, residence times,
8 mixing, and reaction times in the order of milliseconds for both single-phase gas and liquid
9 systems[1-3]. These characteristics are further enhanced in multiphase microsystems, in particular in
10 liquid-liquid biphasic microreactors, where two immiscible liquids come in contact and generate
11 various two-phase flow patterns by tuning the flow conditions[4-9]. The biphasic flow patterns in a
12 microchannel depend on multiple parameters such as the physical properties of the solvents
13 (density, viscosity, and surface tension), the micromixer where the solvent streams intersect, the
14 microchannel wall, diameter, length, and geometry, the flow rate, the fluid velocity, and the
15 fraction of each phase[2, 5, 8]. These flow patterns are usually generated at the inlet of the
16 microchannel using flow focusing geometries and mixing junctions (T-junctions and Y-junctions)
17 as the contacting point of the two liquids.

18   Given the flow rate-range in experimental studies using syringes or reciprocating-piston
19 pumps, the stable flow patterns commonly observed are segmented flow (slug flow and droplet
20 flow) and parallel flow[5]. In segmented flow, the two liquids form alternating segments in which
21 the wall-wetting (continuous) phase usually forms a thin film around the non-wetting (dispersed)
22 phase, whereas parallel flow is characterized by the two liquids flowing side by side through the
23 microchannel. The type of flow pattern formed by a solvent pair drives its ultimate application.
24 Slug and droplet flows have extensively been used for synthesis of nanoparticles and for
25 crystallization. In these flow patterns, the dispersed phase is not in contact with the wall and the
26 slug/droplets exhibit strong inner recirculation with no axial dispersion, producing particles with
27 a narrow particle size distribution[2, 3, 8, 10]. Segmented flow has also been heavily applied in biology
28 for flow injection analysis, blood analysis, DNA analysis[11], protein isolation, and cell
29 encapsulation[12]. Parallel flow has mainly been used in liquid-liquid extraction and characterization
30 of different products, such as metals, metal complexes, ions, and DNA[13-16]. Furthermore, slug and
31 droplet flows have been found to be suitable for short contact time liquid-liquid extraction with
32 high extraction efficiencies and millisecond time-span kinetic studies with in situ characterization
33 of the reaction network[4, 5, 7, 17-24]. For example, liquid-liquid biphasic microreactors have been
34 applied to the production of 5-hydroxylmethyl furfural (HMF), which is a key platform chemical
35 obtained from the acid-catalyzed dehydration of biomass-derived C_6 sugars (glucose and fructose)
36 in water, to prevent the side reactions and increase the HMF yield.

37   To leverage the advantages of different flow patterns for various applications, predicting the
38 liquid-liquid flow patterns for different solvent pairs is vital. Conventionally, dimensionless
39 analysis is applied, and the dimensionless groups are used to predict the flow patterns for different
40 systems. Darekar *et al.*[25] and Yagodnitsyna *et al.*[26] have investigated various maps of patterns
41 using different dimensionless groups and proposed combinations of them that are suitable for
42 mapping the flow patterns. However, these maps still fail to predict flow patterns at certain
43 operating conditions. Since the conventional method of mapping the flow patterns using
44 dimensionless groups is insufficient to predict flow patterns, CFD is a good alternative for flow

2

1　pattern predictions providing also insights into the mechanism of their formation[27]. Kashid and
2　Agar[28] have utilized CFD to investigate the influence of flow rate, channel diameter and Capillary
3　number to the flow regimes and slug size in a liquid-liquid capillary microreactor. On the other
4　hand, Ghaini *et al.*[29] combined experiments and CFD to provide insights into the effects of the
5　contact angle and the interfacial tension on the slug length and the slug formation time using
6　multiple organic solvents. Moreover, Nekouei and Vanapalli[30] have investigated the influence of
7　the viscosity ratio and the flow rate ratio between the two solvents on the droplet size using CFD
8　in a microchannel. In this context, CFD simulations can be employed as numerical experiments
9　which can reveal both the resulting flow conditions and the controlling parameters for the pattern
10　generation. However, CFD is computationally intensive. Rapid prediction of the flow patterns
11　corresponding to certain solvents and operating conditions might be achieved by employing
12　machine learning methodology. As a powerful data-driven computational tool, machine learning
13　could also be applied to learn the flow patterns from a dataset and serve as a robust tool for
14　prediction and classification problems. Das and Samanta[31] have applied artificial neural networks
15　to predict flow patterns for gas-liquid system with 95% accuracy. Furthermore, Nandagopal and
16　Selvaraju[32, 33] published several papers on using multiple neural network techniques to predict flow
17　patterns for liquid-liquid systems with an accuracy of ~95%. Nonetheless, these works mostly
18　focused on utilizing different descriptors, such as superficial velocity and confluence angle, to
19　build a neural network model for only one liquid-liquid system but not for various solvent pairs.

20　　The aim of this paper is to investigate the possibility of using CFD and machine learning
21　techniques to improve the predictive ability of liquid-liquid flow patterns for multiple solvents in
22　capillaries. First, we experimentally study the two phase liquid-liquid flow patterns for four
23　organic solvents with water relevant to HMF extraction[34], ethyl acetate (EtAc), 2-pentanol, MIBK,
24　and heptane. Second, we simulate the various flow patterns using CFD and compare them to our
25　experimental data. Finally, we combine experiments and machine learning techniques to predict
26　the two-phase flow patterns with 93% accuracy. We find that the system is inherently of higher
27　dimensionality of six key descriptors, of which the total flow rate and the capillary number of the
28　aqueous phase are the most important ones.

29　**Methods**
30　*Experimental Methods*
31　　Two syringe pumps (Harvard Apparatus PHD2000 and New Era PumpSystemsInc. NE-300)
32　were used to pump the aqueous and organic solvent feeds into a 500 μm square cross-section T-
33　junction (Valco Instruments) made of polyether ether ketone (PEEK). The details of internal
34　structure of T-junction are provided in the Supplementary material. The feeds intersect in a cross-
35　flow configuration at the T-junction to generate the biphasic flow patterns. The biphasic mixture
36　then enters a capillary microchannel made of perfluoroalkoxy alkane (PFA) tubing (Idex Health)
37　with alternating coiled and straight segments of ID = 500 μm and OD = 1600 μm. Deionized water
38　(Milli-Q) was used as the aqueous solvent. EtAc 99% (Sigma Aldrich), 2-pentanol 99% (Sigma
39　Aldrich), MIBK 99% (Sigma Aldrich), and n-heptane 99% (Sigma Aldrich) were used as the
40　organic solvents. Sodium fluorescein 99% (Sigma Aldrich) and 9,10-diphenyl anthracene 99%
41　(Sigma Aldrich) were used as the aqueous fluorescent dye and the organic fluorescent dye,
42　respectively, in order to contrast the two liquid phases during flow visualization.

3

1   The biphasic flow patterns were characterized using laser induced fluorescence (LIF) of a 250
2   μM solution of sodium fluorescein in water and a 10 mM of 9,10-diphenyl anthracene solution in
3   one of the selected organic solvents, using a high-speed confocal microscope (Highspeed LSM 5
4   Live Duo) mounted with an inverter. Two laser sources with wavelength of 488 nm and 405 nm
5   were used for the fluorescence excitation of the aqueous and organic solvents, respectively. Images
6   were captured using a Zeiss 1.25X and a 2.5X Plan-Neofluar objective lens at frame rates ranging
7   from 30 to 108.1 fps. Further image analysis and processing of the flow patterns were conducted
8   in ImageJ.

9
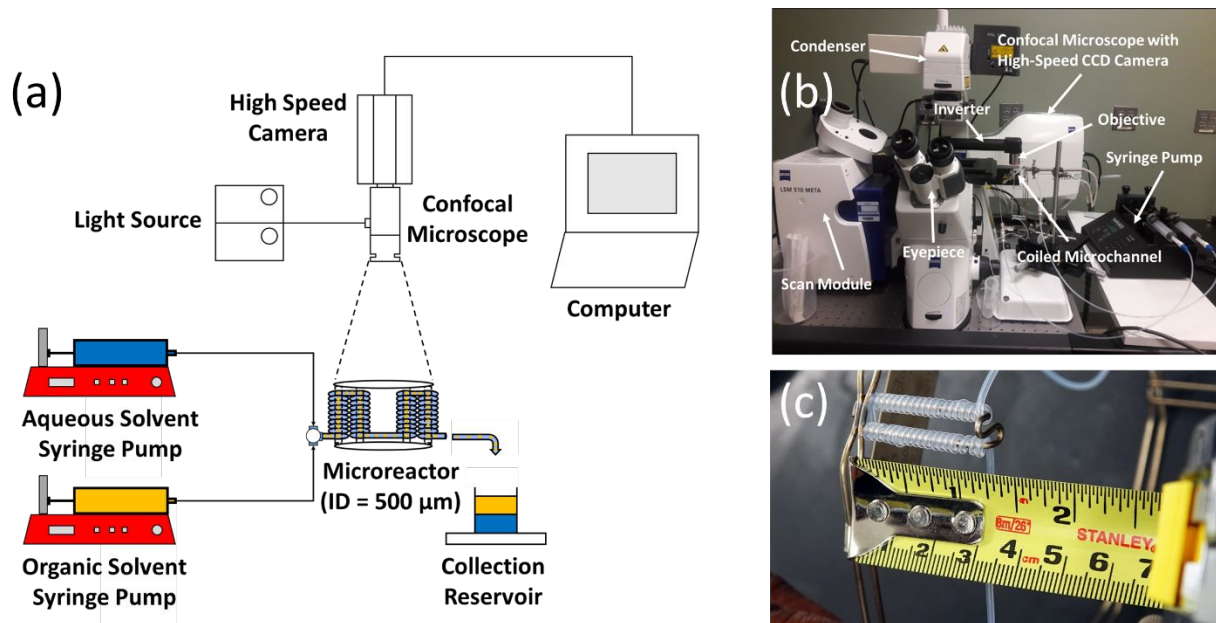
Scheme 1: Schematic diagram of the experimental setup used for flow visualization of the two-phase flow patterns (a), photograph of the high-speed confocal microscope setup (b), and zoom-in view of the coiled microchannel (c).

10

11 *Computational Methods*
12   CFD simulations of two-phase flow are conducted using the single-field incompressible
13 Navier-Stokes equations modeled with the volume-of-fluid (VOF) method[35]. The VOF model is a
14 surface-tracking technique using a fixed mesh system to resolve sharp interfaces. A set of single-
15 field equations is employed to describe the fluid dynamics of two-phase flows, which are obtained
16 by means of conditional volume-averaging of the local instantaneous conservation equations of
17 mass and momentum. The location of the interface is obtained from an indicator function (α),
18 which is equal to the volume fraction of a phase in each cell. The solution of the continuity equation
19 Eq. (2) for the volume fraction (α) of one of the phases enables to track the evolution of the
20 interphase surface[35]

21
$$\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{U}) = 0 \#(1)$$

4

where α is the volume fraction and $\mathbf{U}$ is the velocity field. Since α is discontinuous at the phase boundaries, and it is important to have this discontinuity sharp enough to precisely capture the interface between two phases, an additional term has been introduced, which is a function of the relative velocity between two phases $\mathbf{U_r}$, as in Eq. (2)

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{U}) + \nabla \cdot (\alpha(1-\alpha)\mathbf{U_r}) = 0. \#(2)$$

It is worth noticing that the equation is exact because no assumptions are employed in its derivation. The relative velocity is usually modelled as a compressive velocity normal to the interface with a magnitude proportional to the maximum velocity of the field[36]. The compressive velocity acts to compensate the numerical diffusion across the interface and is calculated according to Eq. (3)

$$\mathbf{U_r} = \mathbf{n_f} \min\left[\frac{C_\gamma |\varphi|}{|S_f|}, \max\left(\frac{|\varphi|}{|S_f|}\right)\right] \#(3)$$

where n_f denotes a vector normal to the cell surface, S_f represents the surface area of each face, φ is the volumetric flux, and C_γ is the compression factor equal to 1 in this study. The conservation of mass and the momentum equation, which is dependent on the volume fractions of the phases through the mixture properties, are expressed by Eqs. (4) and (5)

$$\nabla(\mathbf{U}) = 0 \#(4)$$

$$\frac{\partial(\rho \mathbf{U})}{\partial t} + \nabla \cdot (\rho \mathbf{U} \times \mathbf{U}) = \mu \nabla^2 \mathbf{U} - \nabla P + \rho \mathbf{g} + \mathbf{F_{st}} \#(5)$$

where P is the pressure, ρ is the density, μ is the viscosity, and F_{st} is the surface tension force. The surface tension force acts at the mathematical surface which represents the interface between the phases. In the Finite Volume method, it has to be transformed into a three-dimensional body force to be accounted for. The continuum surface force (CSF) approach[37] is generally employed in VOF simulations

$$\mathbf{F_{st}} = \sigma \kappa \nabla \alpha \#(6)$$

where σ is the surface tension and κ represents the curvature. Brackbill *et al.*[37] pointed out that the curvature of an interface modelled by a volume fraction can be computed by taking the divergence of a normal field, $\kappa = \nabla \cdot \mathbf{n}$. In this work, we employ as normal field the normal to the volume fraction. Hence, the normal field is computed as in Eq. (7)

$$\mathbf{n} = \frac{\nabla \alpha}{|\nabla \alpha|} \#(7)$$

A special treatment is required for the cells near the walls where the adhesion force is considered using a fixed contact angle at the liquid/liquid/solid triple point[37]. According to the CSF model, the unit normal vector is computed at the walls as a function of the contact angle as in Eq. (8)

$$\mathbf{n} = \mathbf{n_w}\cos\theta_0 + \mathbf{t_w}\sin\theta_0 \#(8)$$

5

1   where θ_0 is the contact angle, and $\mathbf{n_w}$ and $\mathbf{t_w}$ are the unit normal and tangential vectors to the wall
2   surface. The value of contact angle defines the wettability. When θ_0 is smaller than 90°, the fluid
3   spreads over most of the area of the wall, and the solid wall is hydrophilic. On the other hand,
4   when θ_0 is larger than 90°, the wetting of the surface is unfavorable and the fluid minimizes the
5   contact with the wall, indicating the wall is hydrophobic.

6   In the VOF method, we solve only the single continuity and momentum equations for the entire
7   domain, an approach known as the "one-fluid" approach[38]. Therefore, the variables and material
8   properties are defined as volume-averaged of the two phases in each cell, as shown in Eqs. (9) and
9   (10)

10   $$\rho = \alpha\rho_1 + (1 - \alpha)\rho_2 \#(9)$$

11   $$\mu = \alpha\mu_1 + (1 - \alpha)\mu_2. \#(10)$$

12   The governing equations are solved with the finite-volume based, open-source CFD toolbox,
13   OpenFOAM[39]. The simulation mesh is established as a 3D T-shape micromixer as shown in
14   Scheme 2, which shares the same geometric parameters and flow conditions with our experimental
15   study. Herein, the channel size is set to 0.5 mm. A total of 1,856,000 cells within the simulation
16   domain were used, and results regarding flow pattern and slug length are found to be independent
17   of discretization; additional information can be found in the Supplementary material. The
18   dispersed phase, water, flows into the T-mixer from one inlet, while the continuous organic phase,
19   e.g., EtAc, flows into the T-mixer from the other inlet and mixes with the water. The maximum
20   Courant number is set to 0.25 to control the time step without reducing the accuracy and quality
21   of the simulation. Zero gradient for pressure and no-slip boundary condition for velocity are
22   implemented at the wall of the microchannel, and the contact angle is set to be 180°. At the inlet,
23   a zero pressure gradient and constant velocity are used as boundary conditions. At the outlet, the
24   velocity and volume fraction are set to be zero-gradient, and the pressure to be atmospheric.

25
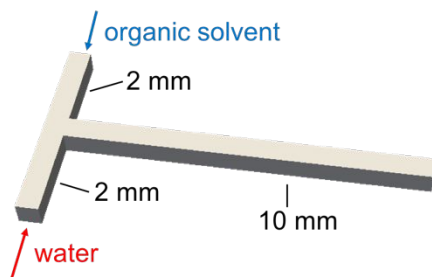
Scheme 2: Flow configuration of the organic and aqueous feeds into the T-junction used for both
the flow visualization experiments and the CFD simulations.

26

27   **Results and Discussion**
28   *Liquid-Liquid Flow Patterns*
29   Figure 1 shows the two-phase flow patterns observed experimentally among a total of 260 data
30   flow visualization points collected. Overall, 7 different biphasic flow patterns are obtained: slug

6

1   flow, droplet, slug-droplet, dispersed, parallel, annular, and irregular flow. The flow patterns were
2   visualized in both coiled and straight segments of the microchannel and there were no observable
3   effects of centrifugal forces arising from the channel curvature on the flow patterns. The organic
4   solvent is the continuous phase in most flow patterns (slug, droplet, slug-drop, parallel, annular,
5   and irregular flows) as it wets the hydrophobic microchannel wall while the aqueous solvent (red
6   phase in Figure 1) is the dispersed phase. However, in the case of the dispersed flow, water is the
7   continuous phase into which the organic solvent is dispersed[40]. The slug and droplet flow regimes
8   are marked by elongated slugs and droplets of uniform and constant size and shape. When the
9   length of the dispersed phase segments is longer than the inner diameter, we define the flow pattern
10   as slug flow; when it is equal or shorter than the inner diameter, as droplet flow. The slug-drop
11   and dispersed flows consist of alternating slugs and droplets with non-uniform size. The parallel
12   flow is characterized by a layer of water flowing next to a layer of organic solvent whereas in the
13   annular flow regime, the aqueous layer flows in between two parallel layers of organic solvent. In
14   the irregular flow regime, the dispersed aqueous phase flows as varying deformed slugs and
15   droplets in the organic solvent.

16   Figure 2 maps the flow conditions that generate the biphasic flow patterns. Interestingly, for
17   the same total volumetric flow rate and org/aq (v/v) ratio, different organic solvents generate
18   different flow patterns. At low flow rates (<1 mL/min), the slug flow is the prevalent for all solvent
19   pairs. At high flow rates (>1 mL/min), the irregular flow is more frequently observed with EtAc
20   and MIBK, the annular flow is dominant with 2-pentanol, and the droplet flow is preferred with
21   heptane. Moreover, parallel flow is only observed when MIBK is used as an organic solvent.
22   Furthermore, the dispersed flow regime is usually encountered at high flow rates and low org/aq
23   (v/v) ratios where the inertial forces in the aqueous phase overcome the surfaces forces in the
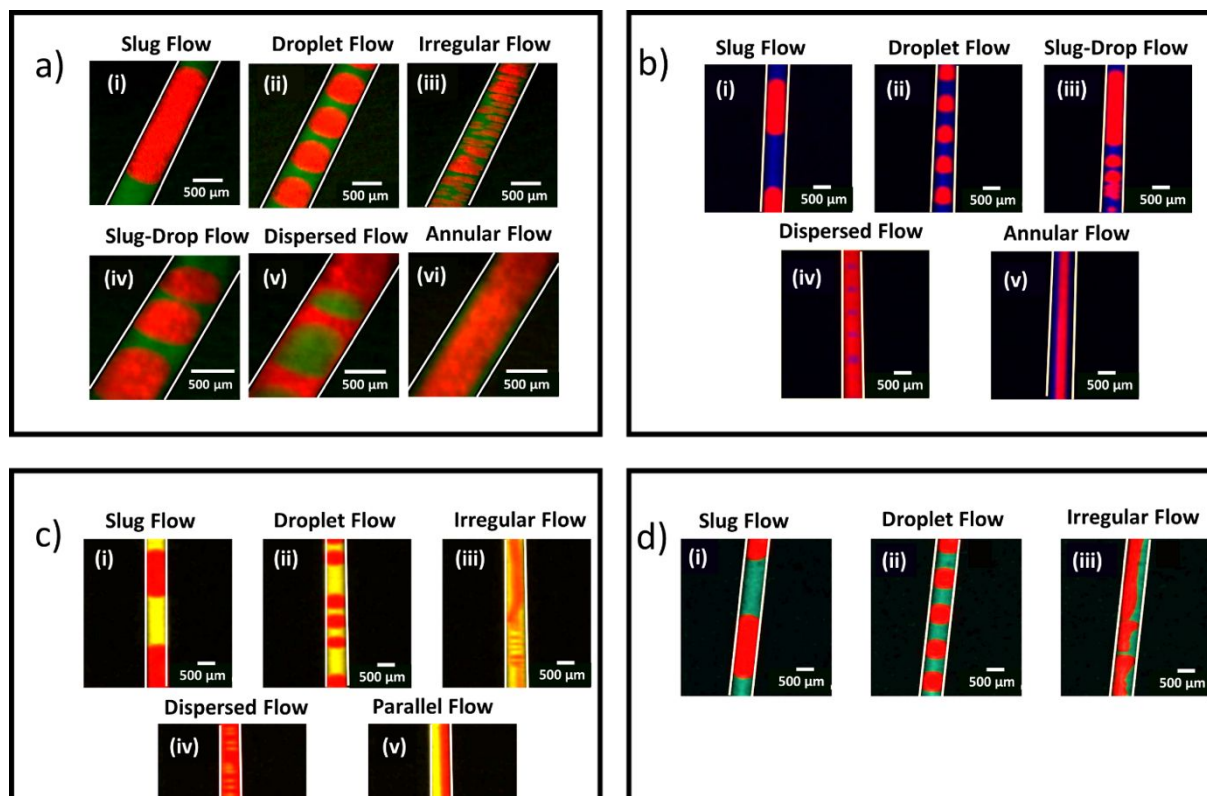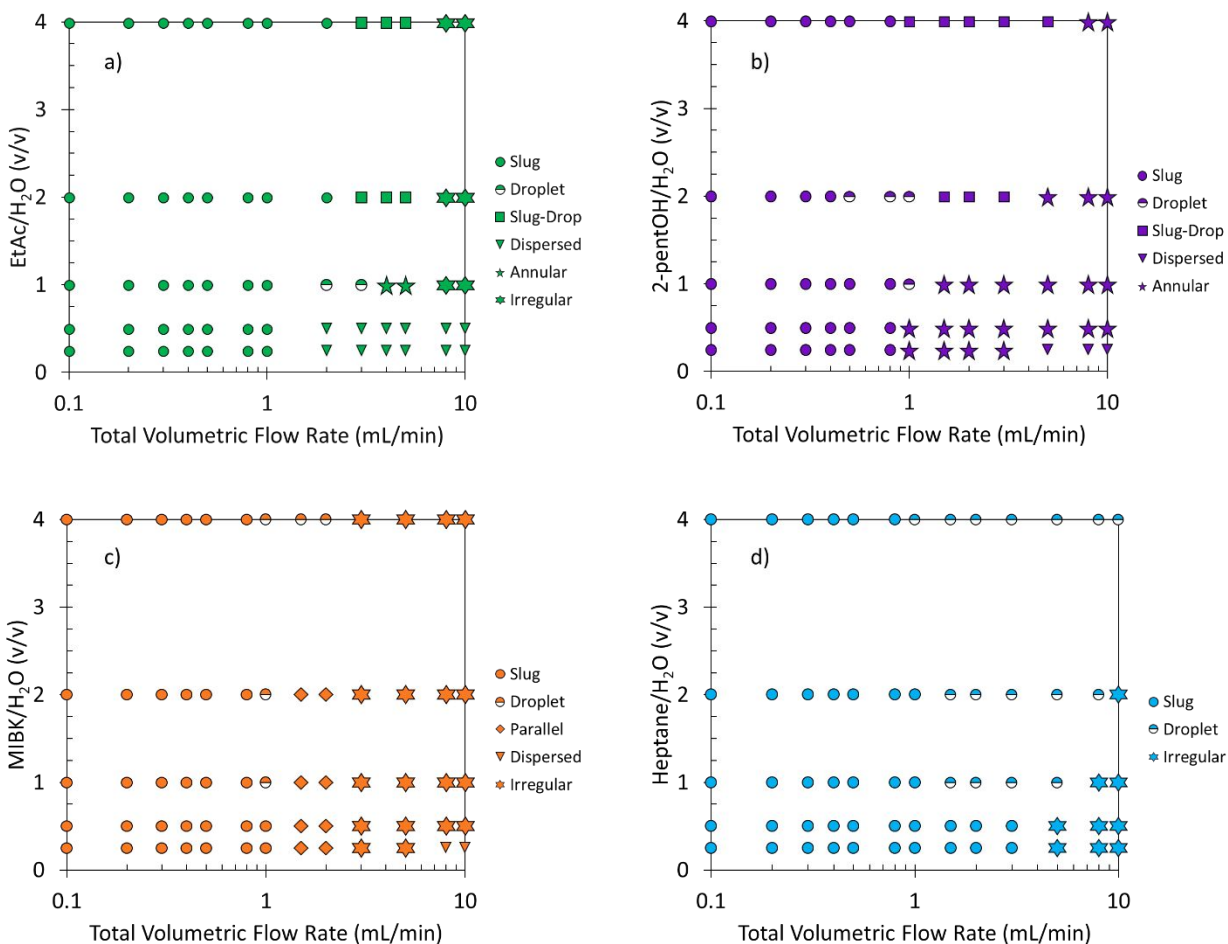24   organic phase, resulting in the aqueous solvent wetting the hydrophobic wall[40].

25

Figure 1: LIF images of the experimental biphasic flow patterns observed for EtAc (green) (a), 2-pentanol (blue) (b), MIBK (yellow) (c), and heptane (cyan) (d). In all the images, water is shown as the red phase. Conditions: total volumetric flow rate from 0.1 to 10 mL/min, org/aq (v/v) from 0.25 to 4. The detailed conditions are indicated in Figure 2.

Figure 2: Maps of flow patterns vs. the org/aq (v/v) ratio and the total volumetric flow rate using various organic solvents: EtAc (a), 2-pentanol (b), MIBK (c), heptane (d).

CFD Modeling of Flow Patterns

We adopt the same configuration employed in the experimental study (Scheme 2) to assess the capabilities of CFD in predicting the flow patterns. By alternating the flow rate and flow rate ratio between organic and aqueous phase, we analyzed the experimental behavior via numerical simulations. The properties of the solvents employed in CFD are reported in Table 1. Our results revealed the prediction of five different flow patterns, namely slug, droplet, slug-drop, annular, and irregular, as shown in Figure 3a. The flow patterns are generally in agreement with experimental results under the same conditions for all the solvents considered. At low flow rates, the interfacial tension forces dominate, enabling the sharp formation of slugs at the T-junction section[41, 42]. We quantitatively assessed the model predictions for this flow pattern by comparing the slug length at different flow rates at unitary flow rate ratio. The slug length in the simulations is evaluated by drawing the isocontour at 0.5 of volume fraction (aqueous phase) to isolate the slug, and then, the actual length is measured using Paraview. Figure 3b shows that the experimental and numerical results are in good agreement with relative deviations at most of 25%. The slug

8

1 length decreases as the flow rate increases due to the gradual reduction of the relative importance
2 of the surface tension which turns to be less effective in building long slugs as the inertial and
3 viscous contributions become more relevant.

4

5 Table 1. Physical properties of the organic solvents used in CFD simulations.

Organic solvent	Dynamic viscosity (Pa-s)	Density (kg/m^3)	Interfacial tension (N/m)
EtAc[43]	4.41×10^{-4}	900	0.0074
2-pentanol	3.47×10^{-3}	812	0.0034
MIBK	5.85×10^{-4}	802	0.0157
Heptane[44]	3.86×10^{-4}	684	0.0507

6

7   As the flow rate increases, the viscous force increases and starts playing a vital role in the
8 pattern formation. At this condition, the interfacial tension force is not sufficient for breakup, and
9 shear-off takes place after some distance from the T-junction to form a droplet[41, 42] as shown in
10 Figure 3a. When the flow rate is high and the inertial force is large enough, the breakup no longer
11 happens and annular flow forms[2], which is accurately predicted by the numerical simulation, as
12 well as the slug-droplet flow regime is observed, which appears only at a high organic-to-aqueous
13 flow rate ratio, as shown in Figure 3a. In this case, the organic flow rate is sufficiently large giving
14 rise to segments of different size because the slugs are no longer stable. At the very high flow rate,
15 rather irregular flow appears, as shown in Figure 3a.

16   Comparison between the numerical simulations and the experimental data reveals an accuracy
17 of 68% of the CFD in predicting the correct pattern. Figure 4 shows a confusion matrix for each
18 of the solvent. Each element of the confusion matrix represents the fraction of experimental flow
19 patterns (x-axis labels) predicted by CFD (y-axis labels). Diagonal (off diagonal) elements
20 represent correct (incorrect) predictions. For example, the 1st row/3rd column element 0.19 in
21 Figure 4a means that 19% of the slug flows (1st row) seen experimentally are predicted to be
22 annular by CFD (3rd column). Most of the wrong predictions are related to the dispersed flow
23 which is never forecasted by the simulations. The remaining errors are related to the misprediction
24 of the slug flow at both low flow rate ratio (0.25) and low total flow rate. The prediction errors are
25 related to distinct reasons in the two cases stemming from the limitations of the current model.

26   The dispersed flow is experimentally observed at low flow rate ratios and high total flow rate.
27 In the experiment, it has been noticed that the water is actually wetting the hydrophobic wall due
28 to the large flow rate compared to the organic phase (low flow rate ratio). In this scenario, the
29 continuous and dispersed phase are reversed with respect to the other experiments because of the
30 particular operating conditions, even if the channel walls are hydrophobic. This finally results in a

9

1   boundary condition mismatch between simulations and experiments. Indeed, the simulation setup
2   imposes a hydrophobic wall due to the nature of the material used in the experiments, using a fixed
3   and pre-defined contact angle whose value is constant in time according to the current surface
4   tension model, i.e., the wettability of the wall does not change during a simulation. This boundary
5   keeps the organic phase touching the wall and avoids the water becoming the continuous phase,
6   thereby preventing the proper prediction of this flow pattern from the simulation. To confirm this
7   hypothesis, we carried out a dedicated simulation using a hydrophilic wall. In doing this, we
8   observed the formation of the dispersed flow as expected from the experiments. CFD simulations
9   are able to reproduce the flow pattern even for this condition. However, they could not be used as
10   a predictive tool since they require knowledge on which phase wets the wall.

11   The misprediction of the slug flow is confined in the region of low flow rate ratio and low total
12   flow rate. In these cases, the annular flow is predicted from the simulation. These conditions are
13   dominated and governed by the surface tension force since the low fluid velocity keeps the
14   contribution of the viscous and inertial force low. The misprediction could be attributed to an
15   under-prediction of the surface tension force between the two phases. As a result, annular flow is
16   obtained since the surface tension is not sufficient for breaking up the water phase. This indicates
17   that an improvement of the current surface tension model might to be needed to precisely predict
18   these specific flow patterns.

19   In summary, CFD simulation is a good tool to provide useful information in designing the
20   actual microreactor and understanding the mechanism for flow pattern generation[30, 45, 46]. It is
21   widely used for Taylor flow regime[27, 30, 45-48], but exposes some limitations in predicting the
22   conditions of different flow patterns observed in experiments. Moreover, CFD simulations are
23   computationally expensive preventing fast screening of different conditions. It takes around 12 to
24   20 hours computing time using 36 CPUs (Intel E5-2695V4) on a high-performance computing
25   cluster to obtain one flow pattern. Thus, it is important to develop an alternative tool to accurately
26   and quickly predict flow patterns for different conditions and water/organic solvent pairs.
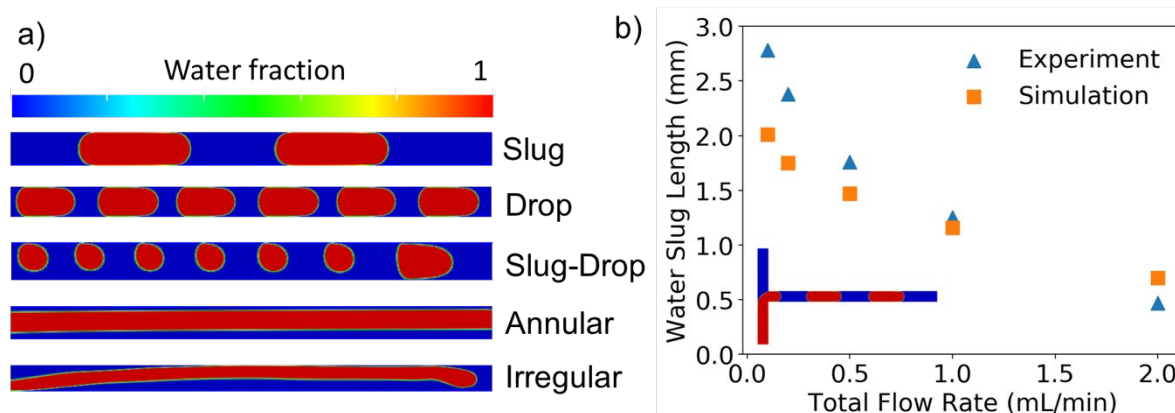
27

Figure 3: CFD generated two-phase flow patterns for the EtAc/water solvent pair (a). Conditions
for the CFD simulations: total volumetric flow rate varies from 0.2, 2, 3, 5 to 8 mL/min from top
to bottom; and org/aq (v/v) ratio is 1, 1, 4, 1 and 2 for each flow pattern; (b) Comparison between
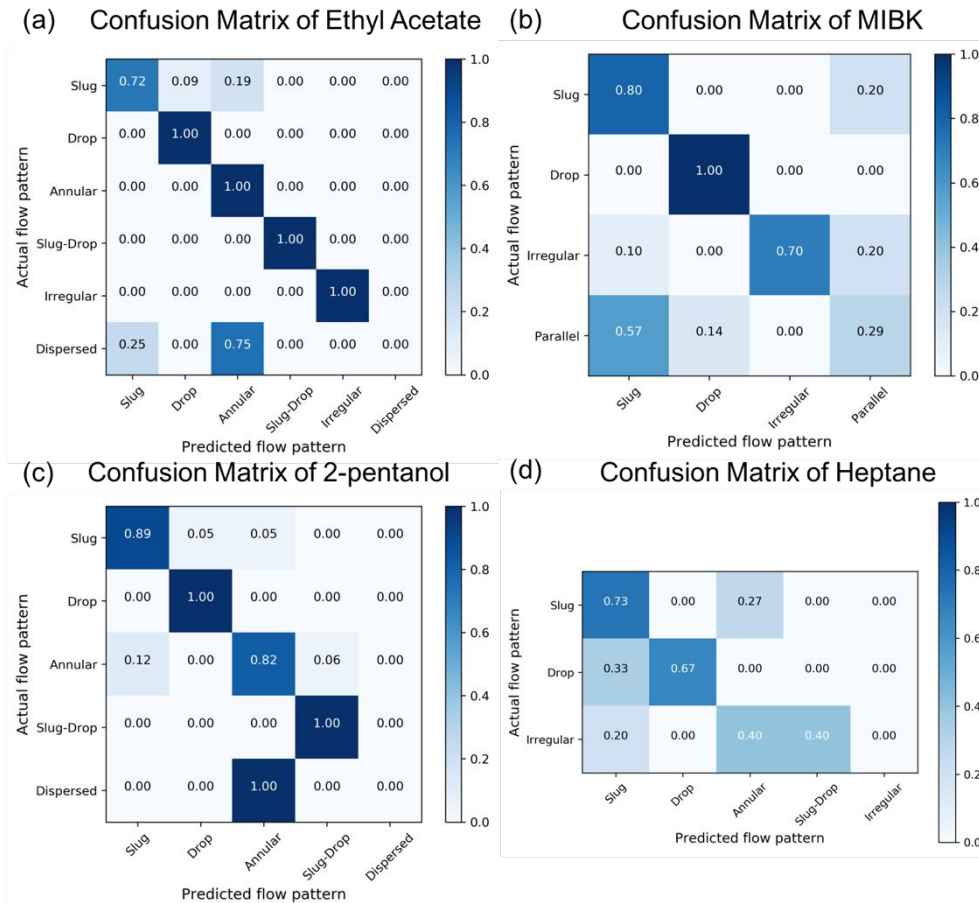the experimental and the CFD predicted slug length of the EtAc/water solvent pair.

10

Figure 4: Confusion matrix showing the accuracy of CFD in predicting the flow patterns in the microchannel for EtAc (a), MIBK (b), 2-pentanol (c), and heptane (d).

1

2   ***Map of Flow Patterns***
3   The micro-flow system is mainly influenced by surface forces, viscous forces, and inertial
4   forces. To better understand the dynamics of the forces, four key dimensionless groups are
5   introduced, namely the Reynolds number (Re), the Capillary number (Ca), the Weber number (We
6   ), and the Ohnesorge number (Oh). The solvent properties in Table 1 are used in the following
7   equations (their physical interpretation is also shown) to estimate their values:

8
$$\mathrm{Re} = \frac{\rho U d_{\mathrm{H}}}{\mu} = \frac{\text{Inertial forces}}{\text{Viscous forces}} \#(11)$$

9
$$\mathrm{Ca} = \frac{\mu U}{\sigma} = \frac{\text{Viscous forces}}{\text{Surface forces}} \#(12)$$

10
$$\mathrm{We} = \frac{\rho U^2 d_{\mathrm{H}}}{\sigma} = \frac{\text{Inertial forces}}{\text{Surface forces}} \#(13)$$

11

$$Oh = \frac{\mu}{\sqrt{\rho \sigma d_H}} = \frac{Viscous\ forces}{\sqrt{Inertial\ forces \times Surface\ forces}} \#(14)$$

where ρ is the density of the solvent, d_H is the hydraulic diameter of the microchannel, μ is the dynamic viscosity of the solvent, and σ is the surface tension between the two liquids.

Flow maps, plotting the various patterns vs. one or two dimensionless numbers for both the continuous and dispersed phase, are typically employed for understanding the dynamics of the surface, viscous, and inertial forces and predicting flow patterns. Figure 5 shows the experimentally observed flow patterns for the various biphasic systems under similar flow conditions vs. the Capillary number of the organic phase (the continuous phase for most of the flow patterns), Ca_{Org}, and the Weber number of the aqueous phase (the dispersed phase in most cases), We_{Aq}. For all organic solvents used, slug flow is obtained at $Ca_{Org}< 0.01$ and $We_{Aq}< 1$. These data underscore the fact that slug flow is characterized by strong surface forces. Nonetheless, the maximum Ca number below which slug flow occurs depends on the organic solvent: slug flow is obtained up to $Ca_{Org}\sim 0.1$ for 2-pentanol and $We_{Aq}\sim 2$ for heptane, reflective of relatively stronger viscous forces in the 2-pentanol/water system and stronger inertial forces in the heptane/water system. Furthermore, at the same values of Ca_{Org} and We_{Aq}, different organic solvents generate different flow patterns. For example, in Figure 3, at $We_{Aq} = 1$ and $Ca_{Org} = 0.02$, EtAc produces droplet flow, 2-pentanol generates slug flow, MIBK forms parallel flow, and heptane droplet flow. These observations suggest that the common dimensionless numbers are inadequate to encompass the force dynamics in both solvents of a biphasic system. Flow maps, like Figure 5, are limited in accurately predicting the flow patterns since Ca_{Org} does not account for the contribution of the inertial forces in the organic phase and We_{Aq} does not consider the relative strength of the viscous forces in the aqueous phase.

The above observations clearly show that biphasic flow patterns are multi-dimensional systems that depend on the physical properties of the solvents, the flow rate, the flow rate ratio, the common dimensionless numbers, the microchannel diameter, the channel wall, and the contacting geometry. To mitigate the limitations in predicting biphasic flow patterns in microchannels, a universal map of patterns vs. the WeOh product of each phase was developed by Yagodnitsyna and coworkers[26]. Figure 6 shows the flow map of experimental patterns for the four biphasic systems used in this study. Slug flow is well described for all organic solvents for $(WeOh)_{Org}< 0.003$ for the organic phase and $(WeOh)_{Aq}< 0.01$ for the aqueous phase, which is in the range of WeOh values reported by Yagodnitsyna *et al.*[26]. However, this universal map fails at accurately predicting the flow patterns observed at higher WeOh values, as shown in Figure 6. The red circle highlights the considerable overlap of different flow patterns at similar WeOh values. This again indicates the need for an alternative tool to accurately and quickly predict flow patterns at different conditions and water/organic solvent pairs.
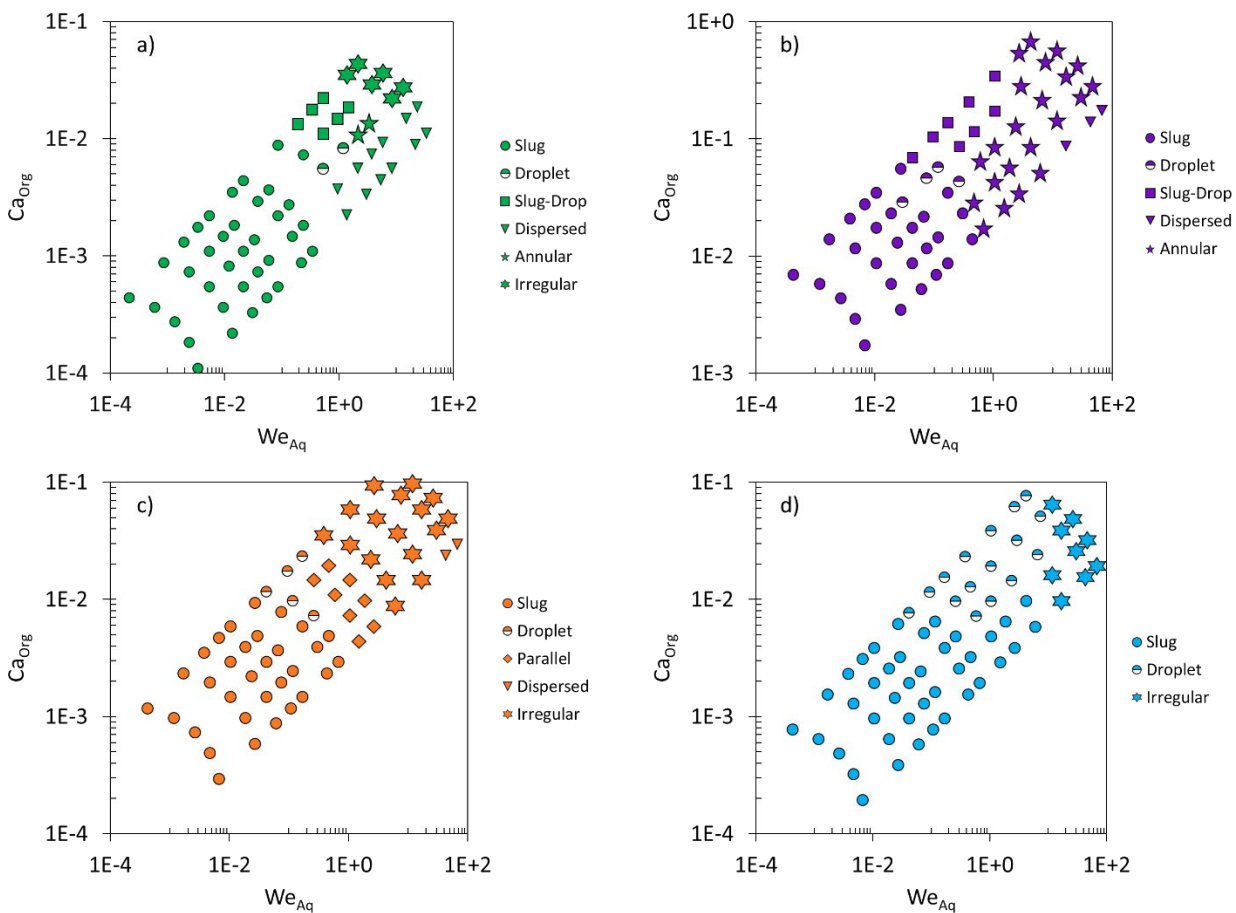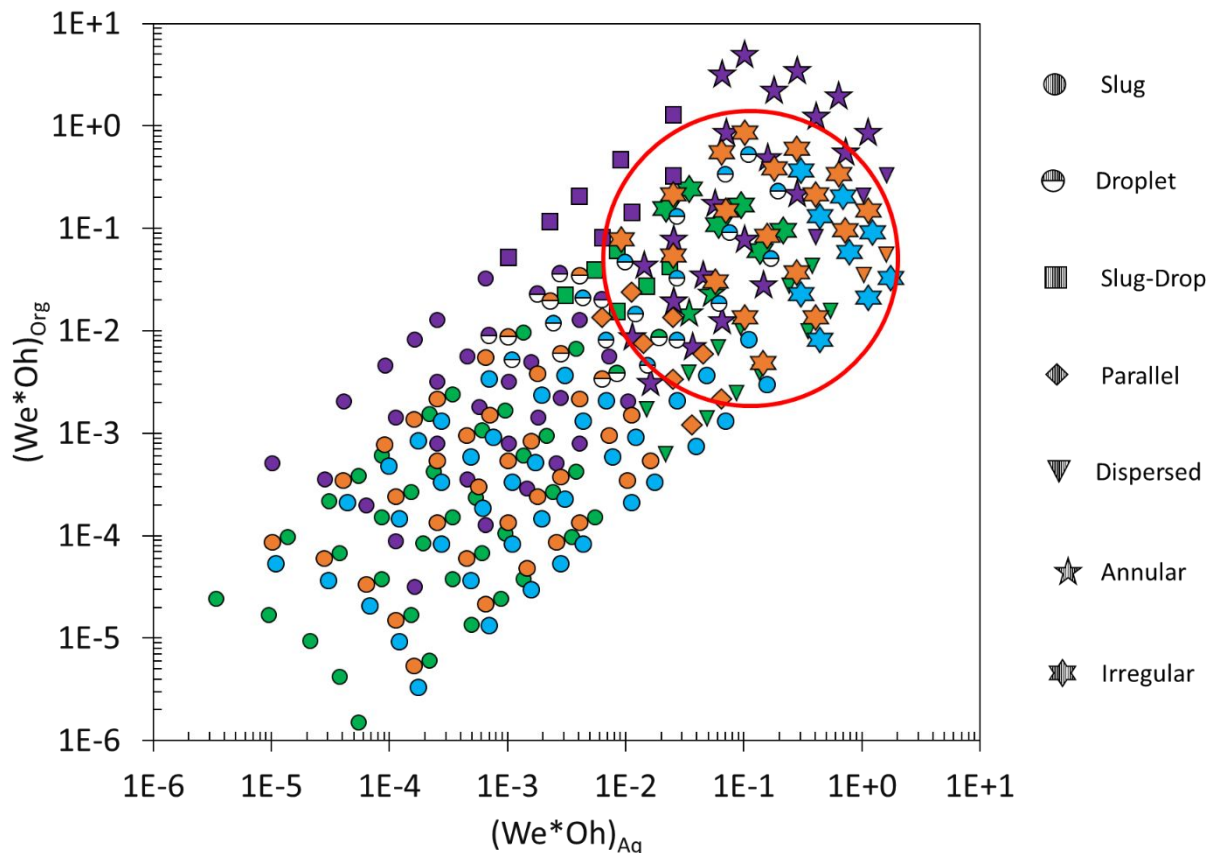
12

Figure 5: Dimensionless flow map plotting the Capillary number of the organic phase against the Weber number of the aqueous phase for the two-phase flow patterns observed experimentally with various organic solvents: EtAc (a), 2-pentanol (b), MIBK (c), and heptane (d). Flow conditions: total volumetric flow rate from 0.1 to 10 mL/min, org/aq (v/v) ratio from 0.25 to 4.

1

13

1

2 Figure 6: Map of flow patterns against the product of the Weber number times the Ohnesorge
3 number for both the organic and aqueous phase for all the organic solvents: EtAc (green), 2-
4 pentanol (purple), MIBK (orange), and heptane (cyan). Flow conditions: total volumetric flow rate
5 from 0.1 to 10 mL/min, org/aq (v/v) ratio from 0.25 to 4.

6

7 *Principal Component Analysis (PCA)*
8   PCA is a useful statistical tool for dimensionality reduction and analysis[49]. It can identify the
9 dimensionality of the problem and show potential correlations, and is fairly simple to use.
10 Considering material properties and dimensionless groups of both aqueous and organic phases,
11 there are 19 features in total, shown in Table 2, selected for our analysis. The explained variance
12 of each principal component is shown in Figure 7a, and the cumulative explained variance in
13 Figure 7b. The top 2 principal components only explain 38% and 23% of the variance, respectively,
14 and the variance explained by the top 6 principal components is significantly higher than the
15 remaining. Hence, at least 6 features are needed to describe >95% of the variance of the data.
16 Unlike prior work using two dimensionless groups to describe flow patterns, it is clear that PCA
17 suggests that the problem is of higher dimensionality. Since principal components are linear
18 combinations of the original features, it is hard to directly identify the key descriptors and
19 characterize the flow pattern. Another method is therefore needed to efficiently predict the patterns
20 as well as to identify the relevant descriptors.

14

1   Table 2. Features determined by principal component analysis (PCA) to predict the two-phase flow
2   patterns.

|	Features
Aqueous phase (Aq)	Kinematic viscosity (Nu), density (Rho), velocity (V), Capillary number (Ca), Reynolds number (Re), Weber number (We), Ohnesorge number (Oh) and We*Oh
Organic phase (Org)	Kinematic viscosity, density, velocity, Capillary number, Reynolds number, Weber number, Ohnesorge number and We*Oh
System-wide	Interfacial tension (Sigma), total flow rate (Q), flow rate ratio (Org/Aq)

3
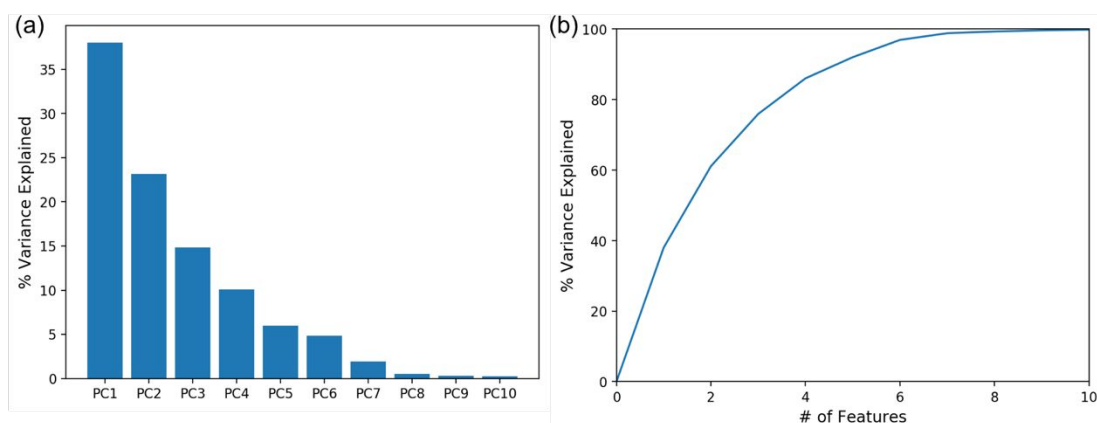
Figure 7: Explained variance by principal components (a) and cumulative explained variance (b).

4

5   ***Decision Tree Model***
6   Decision tree learning is a tree-based predictive modeling for regression and classification that
7   provides interpretive predictions[50-52]. As a commonly used supervised machine learning method
8   in data mining, decision tree could assist in predicting either the value (regression) or the class
9   (classification) of target variable(s) from input observations[53]. Since our variables are discrete, we
10   employ classification trees. To construct a decision tree, the dataset is split into different subsets
11   by choosing certain variables that could best split the set of observations at each node of the tree.
12   The building process of a simple decision tree is illustrated in the Supplementary information.

13   The Gini impurity is employed to each candidate subset to evaluate the quality of the split and
14   decide the variable for splitting data. The Gini impurity is the measure that a randomly chosen data
15   point would be incorrectly classified by randomly using one of the distributions of classes in the
16   subset[51-53]. Here k represents the number of classes and p_i represents the fraction of elements that
17   are classified as class *i*. The split that minimizes the Gini impurity for a system is chosen to divide
18   the dataset and construct the tree. This process is repeated on each subset recursively to build a

15

1   decision tree until reaching stopping criteria, i.e., every subset contains only one unique class.
2   Then, the averaged Gini impurity decreases for each variable in a node over the entire tree and can
3   be used to estimate the variable importance[51]. The decision tree algorithm is carried out using
4   scikit-learn, which is a well-known open-source machine learning package in Python[54].

5   The response space in our problem contains different flow patterns, and the predictors are
6   features selected from Table 2. There is a total of 260 data points collected, and the data is
7   randomly split into a training set and a testing set; 80% of data is used for training and the
8   remaining 20% is kept for testing. A 3-fold cross validation is applied during model training to
9   avoid overfitting. The testing data is used to evaluate the performance of the model. The decision
10   tree model is built using the same features, listed in Table 2. Figure 8a shows the confusion matrix
11   of the flow pattern predictions and Figure 8b shows the feature importance. Figure 8a shows the
12   decision tree model reaches 87% overall accuracy among all flow patterns. The model correctly
13   predicts the slug-drop flow, dispersed flow, and parallel flow. On the other hand, the drop flow is
14   the most wrongly predicted. This may be attributed to the very narrow operating condition window
15   of the drop flow. This pattern has been observed in a very limited number of operating conditions,
16   which in some cases overlaps with the other flow regimes. The machine learning algorithm thereby
17   predicts the drop flow as the common slug or the slug-drop flow in some cases. Besides, the top 6
18   important features are identified from the feature importance using the decision tree model. These
19   include the total flow rate, the Capillary number of aqueous phase, the Weber number of the
20   organic phase, the flow rate ratio, the velocity of the organic phase, and the Ohnesorge
21   dimensionless group of the aqueous phase. The total flow rate and flow rate ratio are identified as
22   the most important factors that influence patterns, consistent with these being the key variables we
23   tune to obtain different flow patterns. Moreover, the Capillary number of the aqueous phase and
24   the Weber number of the organic phase are identified as vital factors as well, indicating that
25   whether the system is dominated by surface tension force is important for dictating the flow
26   patterns. Using the top 6 important features to build the decision tree model again, the confusion
27   matrix of new model (Figure 9a) and the evaluation metrics are shown in Figure 9b, where the F1
28   score is the harmonic average of precision and recall. The precision is the number of true positives
29   divided by the true positives plus false positives[55]. The recall is the correct identifications of the
30   class over the total population of the class[55]. The new decision tree model reaches 93% accuracy
31   and correctly predicts slug flow, annular flow, slug-droplet flow, dispersed flow, and parallel flow.
32   The accuracy and F1 score are improved due to reduced overfitting. This indicates that these are
33   the key features of controlling flow patterns. The decision tree model reaches much higher
34   accuracy (>90%) by identifying the key features for this classification problem.

35   The methodology of using a decision tree to predict flow pattern could also be extended to
36   liquid-liquid microchannels with hydrophilic walls and gas-liquid systems. When the channel wall
37   is changed from hydrophobic to hydrophilic, the organic phase will become the dispersed phase
38   and the aqueous phase will then be the continuous phase, and different flow patterns might be
39   generated under the same flow conditions. On the other hand, for gas-liquid systems, there would
40   be significant density and surface tension differences, leading to the greater range of dimensionless
41   numbers. In general, the key features identified in this study would also be important in these
42   systems, since these same parameters are tuned by varying operating conditions and determine
43   when the surface tension force dominates. However, the relative importance of key features may
44   differ because of the much different range of dimensionless numbers and operating conditions.
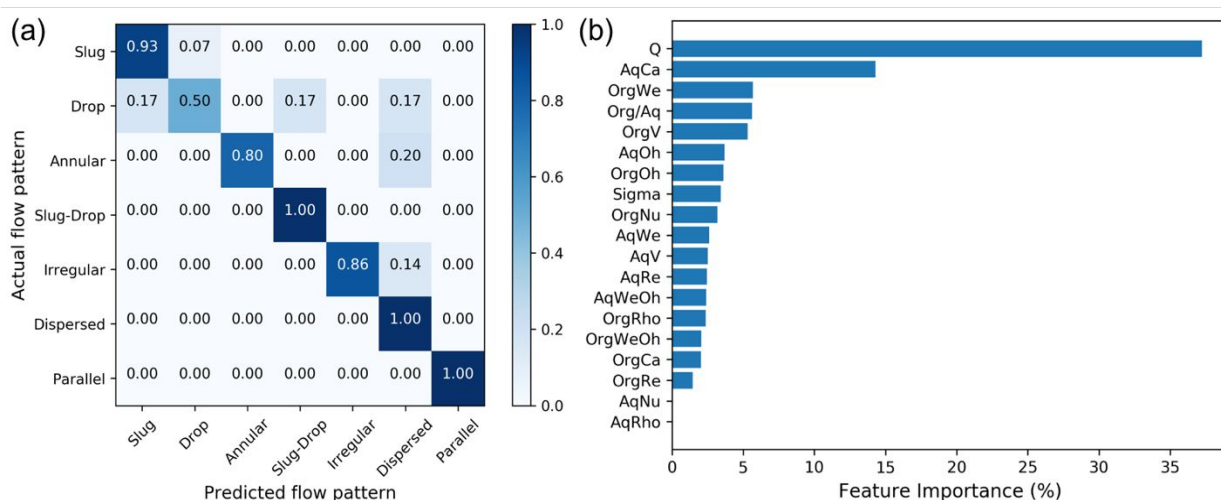
1

Figure 8: Confusion matrix of decision tree model (a) and feature importance (b). Symbols are defined in Table 2.
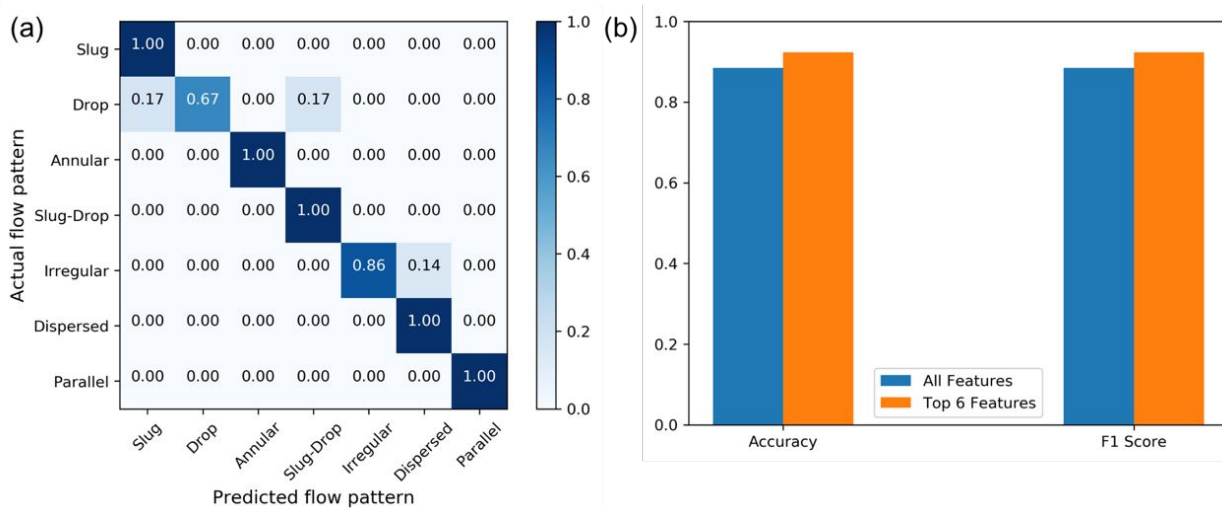
2

Figure 9: Confusion matrix of the decision tree model (a) and performance of the decision tree model (b).

3

4 **Conclusions**
5   Two-phase flow patterns were studied in a microchannel for systems of relevance to the
6 reactive extraction of HMF, using 4 different organic/aqueous biphasic systems, with water as the
7 aqueous phase and organic solvents including EtAc, 2-pentanol, MIBK, and heptane. The various
8 biphasic flow regimes were generated using a T-junction and were studied experimentally using
9 LIF flow visualization. In total, 7 different two-phase flow patterns were observed: slug flow,
10 droplet flow, slug-drop flow, parallel flow, annular flow, dispersed flow, and irregular flow. Maps

17

for the different flow patterns were created for different flow conditions and all solvent pairs. Interestingly, different flow patterns were observed for different organic solvents under similar flow conditions. Nonetheless, the slug flow regime was found to be common to all solvent pairs at low flow rates <1 mL/min. The flow patterns were, then, modeled using CFD simulations with fairly good agreement with the experimental results for most of the flow patterns. However, dispersed flow and slug flow at low flow rate ratio are usually wrongly predicted by numerical simulations. In the former, the specific operating conditions, i.e. large amount of water, determines that water is in contact with the wall, an aspect not captured by simulations where as wall is assumed to be hydrophobic. In the latter, the flow is governed by the surface tension due to the very small capillary number. In this case, the closure model for the surface tension plays a vital role for the accurate description of patterns. The model employed in this work is found to under-predict the surface tension resulting in misprediction of this flow pattern.

Finally, the flow patterns were analyzed using various descriptors to delineate the physics of flow pattern formation. Using the two common dimensionless groups of Weber number and Ohnesorge number results fails in uniquely identifying the patterns. Principal component analysis determines the most important features and shows that at least 6 features are necessary to explain >95% of the variance of the experimental data. A decision-tree model was then built to further analyze the data and provide insights into the most important features. The model improved the accuracy in predicting flow patterns up to 93% and exposed six critical features that include the total flow rate, the flow rate ratio between the two phases, the Capillary and Ohnesorge numbers of the aqueous phase, and the Weber number and velocity of the organic phase. This work introduces machine learning in accurately describing complex flow patterns to increase the ability to predict liquid-liquid flow patterns and reduce the computational cost compared to using CFD simulations. Moreover, this work underscores that these systems are inherently of higher dimensionality than thought before.

Conflicts of interest

There are no conflicts to declare.

References

1. M. Hoffmann, M. Schluter and N. Rabiger, *Chemical Engineering Science*, 2006, **61**, 2968-2976.
2. C. X. Zhao and A. P. J. Middelberg, *Chemical Engineering Science*, 2011, **66**, 1394-1411.
3. Y. C. Zhao, G. W. Chen and Q. Yuan, *Aiche Journal*, 2006, **52**, 4052-4060.
4. S. L. Anna, N. Bontoux and H. A. Stone, *Applied Physics Letters*, 2003, **82**, 364-366.
5. N. Assmann, A. Ladosz and P. R. von Rohr, *Chemical Engineering & Technology*, 2013, **36**, 921-936.

6. G. F. Christopher and S. L. Anna, *Journal of Physics D-Applied Physics*, 2007, **40**, R319-R336.

7. A. L. Dessimoz, L. Cavin, A. Renken and L. Kiwi-Minsker, *Chemical Engineering Science*, 2008, **63**, 4035-4044.

8. A. Gunther and K. F. Jensen, *Lab on a Chip*, 2007, **7**, 399-399.

9. K. Wang, L. Li, P. Xie and G. Luo, *Reaction Chemistry & Engineering*, 2017, **2**, 611-627.

10. B. K. H. Yen, N. E. Stott, K. F. Jensen and M. G. Bawendi, *Advanced Materials*, 2003, **15**, 1858-1862.

11. M. A. Burns, B. N. Johnson, S. N. Brahmasandra, K. Handique, J. R. Webster, M. Krishnan, T. S. Sammarco, P. M. Man, D. Jones, D. Heldsinger, C. H. Mastrangelo and D. T. Burke, *Science*, 1998, **282**, 484-487.

12. S. Sugiura, T. Oda, Y. Izumida, Y. Aoyagi, M. Satake, A. Ochiai, N. Ohkohchi and M. Nakajima, *Biomaterials*, 2005, **26**, 3327-3331.

13. M. Surmeian, A. Hibara, M. Slyadnev, K. Uchiyama, H. Hisamoto and T. Kitamori, *Analytical Letters*, 2001, **34**, 1421-1429.

14. H. Hisamoto, T. Horiuchi, M. Tokeshi, A. Hibara and T. Kitamori, *Analytical Chemistry*, 2001, **73**, 1382-1386.

15. Y. Kikutani, H. Hisamoto, M. Tokeshi and T. Kitamori, *Lab on a Chip*, 2004, **4**, 328-332.

16. V. Reddy and J. D. Zahn, *Journal of Colloid and Interface Science*, 2005, **286**, 158-165.

17. A. Adamo, P. L. Heider, N. Weeranoppanant and K. F. Jensen, *Industrial & Engineering Chemistry Research*, 2013, **52**, 10802-10808.

18. A. Ghaini, M. N. Kashid and D. W. Agar, *Chemical Engineering and Processing*, 2010, **49**, 358-366.

19. J. Jovanovic, E. V. Rebrov, T. A. Nijhuis, M. T. Kreutzer, V. Hessel and J. C. Schouten, *Industrial & Engineering Chemistry Research*, 2012, **51**, 1015-1026.

20. M. N. Kashid, Y. M. Harshe and D. W. Agar, *Industrial & Engineering Chemistry Research*, 2007, **46**, 8420-8430.

21. P. Mary, V. Studer and P. Tabeling, *Analytical Chemistry*, 2008, **80**, 2680-2687.

22. D. Tsaoulidis and P. Angeli, *Chemical Engineering Journal*, 2015, **262**, 785-793.

23. A. Woitalka, S. Kuhn and K. F. Jensen, *Chemical Engineering Science*, 2014, **116**, 1-8.

24. M. Shang, T. Noël, Y. Su and V. Hessel, *AIChE Journal*, 2017, **63**, 689-697.

25. M. Darekar, K. K. Singh, S. Mukhopadhyay and K. T. Shenoy, *Industrial & Engineering Chemistry Research*, 2017, **56**, 12215-12226.

26. A. A. Yagodnitsyna, A. V. Kovalev and A. V. Bilsky, *Chemical Engineering Journal*, 2016, **303**, 547-554.

27. D. A. Hoang, V. van Steijn, L. M. Portela, M. T. Kreutzer and C. R. Kleijn, *Computers & Fluids*, 2013, **86**, 28-36.

28. M. N. Kashid and D. W. Agar, *Chemical Engineering Journal*, 2007, **131**, 1-13.

29. A. Ghaini, A. Mescher and D. W. Agar, *Chemical Engineering Science*, 2011, **66**, 1168-1178.

30. M. Nekouei and S. A. Vanapalli, *Physics of Fluids*, 2017, **29**.

31. H. Sharma, G. Das and A. N. Samanta, *AIChE Journal*, 2006, **52**, 3018-3028.

32. M. S. Giri Nandagopal and N. Selvaraju, *Industrial & Engineering Chemistry Research*, 2016, **55**, 11346-11362.

19

33. M. S. G. Nandagopal, E. Abraham and N. Selvaraju, *Chemical Engineering Journal*, 2017, **309**, 850-865.

34. Y. Muranaka, H. Nakagawa, R. Masaki, T. Maki and K. Mae, *Industrial & Engineering Chemistry Research*, 2017, **56**, 10998-11005.

35. C. W. Hirt and B. D. Nichols, *Journal of Computational Physics*, 1981, **39**, 201-225.

36. E. Berberovic, N. P. van Hinsberg, S. Jakirlic, I. V. Roisman and C. Tropea, *Physical Review E*, 2009, **79**.

37. J. U. Brackbill, D. B. Kothe and C. Zemach, *Journal of Computational Physics*, 1992, **100**, 335-354.

38. D. Gueyffier, J. Li, A. Nadim, R. Scardovelli and S. Zaleski, *Journal of Computational Physics*, 1999, **152**, 423-456.

39. H. G. Weller, G. Tabor, H. Jasak and C. Fureby, *Computers in Physics*, 1998, **12**, 620-631.

40. D. Tsaoulidis, V. Dore, P. Angeli, N. V. Plechkova and K. R. Seddon, *International Journal of Multiphase Flow*, 2013, **54**, 1-10.

41. M. De Menech, P. Garstecki, F. Jousse and H. A. Stone, *Journal of Fluid Mechanics*, 2008, **595**, 141-161.

42. P. Garstecki, M. J. Fuerstman, H. A. Stone and G. M. Whitesides, *Lab on a Chip*, 2006, **6**, 437-446.

43. M. Z. Shahid, M. R. Usman, M. S. Akram, S. Y. Khawaja and W. Afzal, *Journal of Chemical & Engineering Data*, 2017, **62**, 1198-1203.

44. S. Zeppieri, J. Rodríguez and A. L. López de Ramos, *Journal of Chemical & Engineering Data*, 2001, **46**, 1086-1088.

45. L. Dai, W. F. Cai and F. Xin, *Chemical Engineering & Technology*, 2009, **32**, 1984-1991.

46. D. Y. Qian and A. Lawal, *Chemical Engineering Science*, 2006, **61**, 7609-7625.

47. M. J. Nieves-Remacha, L. Yang and K. F. Jensen, *Industrial & Engineering Chemistry Research*, 2015, **54**, 6649-6659.

48. L. Yang, M. J. Nieves-Remacha and K. F. Jensen, *Chemical Engineering Science*, 2017, **169**, 106-116.

49. E. Lopez-Guajardo, E. Ortiz-Nadal, A. Montesinos-Castellanos and K. D. P. Nigam, *Chemical Engineering Communications*, 2017, **204**, 467-475.

50. D. A. C. Beck, J. M. Carothers, V. R. Subramanian and J. Pfaendtner, *AIChE Journal*, 2016, **62**, 1402-1416.

51. T. Hastie, R. Tibshirani and J. H. Friedman, *The elements of statistical learning : data mining, inference, and prediction*, 2009.

52. G. James, D. Witten, T. Hastie, R. Tibshirani, *An introduction to statistical learning : with applications in R*, 2013.

53. B. Partopour, R. C. Paffenroth and A. G. Dixon, *Computers & Chemical Engineering*, 2018, **115**, 286-294.

54. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825-2830.

55. M. Sokolova and G. Lapalme, *Information Processing & Management*, 2009, **45**, 427-437.