frontiers | Frontiers in Chemical Engineering

TYPE Original Research
PUBLISHED 08 November 2023
DOI 10.3389/fceng.2023.1274349

OPEN ACCESS

EDITED BY
Hans-Jörg Bart,
University of Kaiserslautern, Germany

REVIEWED BY
Matthaeus Siebenhofer,
Graz University of Technology, Austria
Ville Alopaeus,
Aalto University, Finland

*CORRESPONDENCE
Andreas Jupke,
andreas.jupke@avt.rwth-aachen.de

RECEIVED 08 August 2023
ACCEPTED 05 October 2023
PUBLISHED 08 November 2023

CITATION
Palmtag A, Rousselli J, Gröschl H and Jupke A (2023), Hybrid modeling of drop breakage in pulsed sieve tray extraction columns.
Front. Chem. Eng. 5:1274349.
doi: 10.3389/fceng.2023.1274349

Hybrid modeling of drop breakage in pulsed sieve tray extraction columns

Andreas Palmtag, Johannes Rousselli, Henning Gröschl and Andreas Jupke*

Fluid Process Engineering (AVT.FVT), RWTH Aachen University, Aachen, Germany

Accurate models for pulsed sieve tray extraction columns (PSEs) depend on the correct prediction of the drop diameter to estimate extractive mass transfer across the phase boundary. Phenomenologically, the drop diameter is determined by a balance of drop breakage and coalescence. While for most industrial solvent systems, coalescence plays a minor role; breakage is mostly the dominant phenomenon determining the drop diameter. However, most modeling approaches for drop breakage in PSEs are characterized by a trade-off between a broad validity range and good prediction accuracy. To overcome this limitation, we developed a hybrid breakage model for drop breakage in PSEs in which a physical-empirical model basis is enhanced by data-driven parameter estimator models (PEMs). The hybrid model is based on a revised form of Garthe's breakage model, for which we developed a linear PEM for the model parameters and two data-driven PEMs for d_{stab} and d_{100}, respectively. The hybrid breakage model was validated on 743 experimental data sets and evaluated based on the pull metric. In a sensitivity analysis, the model correctly predicted the breakage probability over a wide range of solvent properties, operating conditions, and sieve tray geometries. In future studies, the hybrid breakage model can be incorporated into extraction column models without an initial parametrization.

KEYWORDS

pulsed sieve tray extraction column, liquid-liquid extraction, drop breakage, hybrid model, population balance model

1 Introduction

Pulsed sieve tray extraction columns (PSE) are one of the most common process equipment for industrial extraction, having found application in chemical and hydrometallurgical processes, in biotechnology, and particularly in nuclear fuel reprocessing (Lo et al., 1983; Schügerl, 1994; Gameiro et al., 2010). The numerous industrial applications have motivated intensive research on PSE, including experimental studies and the development of column models with various degrees of rigor. Particularly, the wish to limit the number of preliminary column experiments on a technical scale induced the development of models that correctly depict fluid dynamics and mass transfer over

Abbreviations: ANN, Artificial neural network; BW, Butyl acetate/water; BWA, Butyl acetate/water/acetone; EFCE, European Federation of Chemical Engineering; GP, Gaussian Process; ML, Machine learning; PBM, Population balance model; PEM, Parameter estimation model; PSE, Pulsed sieve-tray extraction column; Sec, Section; TW, Toluene/water; TWA, Toluene/water/acetone.

10.3389/fceng.2023.1274349

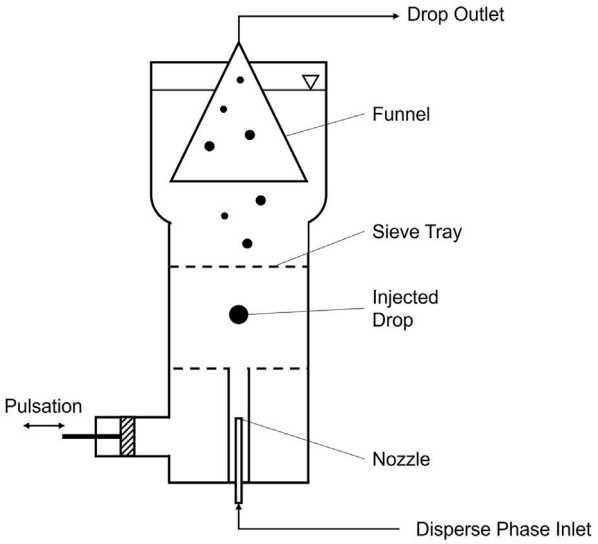

FIGURE 1
Single-drop cell for investigating drop breakage on sieve trays.

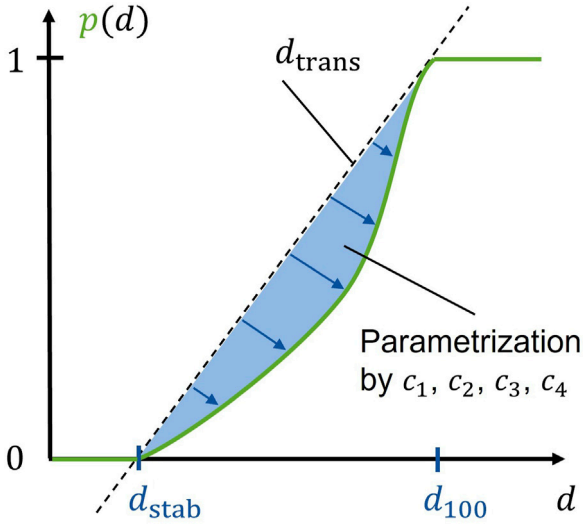

FIGURE 2
Qualitative representation of Garthe's breakage model.

several scales. A promising approach for this purpose is based on population balance models (PBMs), which track the evolution of the drop swarm along the column height by accounting for key phenomena such as drop sedimentation and mass transfer by physical-empirical sub-models (Goedecke, 2006; Weber et al., 2019). The accuracy of PBMs highly depends on the correct prediction of the drop diameter, which is crucial for accurate modeling of fluid dynamics and mass transfer in extraction columns (Hlawitschka et al., 2020; Weber and Jupke, 2020). Phenomenologically, the drop diameter is mainly determined by a balance of drop breakage and coalescence phenomena. Commonly, drop breakage is considered dominant since many industrial processes are characterized by significant coalescence inhibition, e.g., due to mass transfer and/or impurities (Henschke, 2003). Consequently, considerable research was conducted to investigate, model, and predict the drop breakage behavior in PSEs.

Most studies on drop breakage behavior are based on single-drop investigations in lab-scale devices to limit the experimental effort. A sketch of a single-drop cell is shown in Figure 1. The cell consists of a single compartment between two sieve trays, a pulsation unit, and a periphery to insert and remove single-drops into/from the cell. The setup in Figure 1 allows the investigation of organic drops (inserted at the bottom) submerged in an aqueous continuous phase. Within the single-drop experiments, the breakage behavior is commonly quantified based on the breakage probability $p(d)$—the fraction of drops breaking, the stable drop diameter d_{stab}—the maximal diameter of drops that never break ($p(d \le d_{\mathrm{stab}}) = 0$), and the number of daughter drops ϑ created during breakage. Usually, up to 50 drops are investigated to determine $p(d)$, d_{stab}, and ϑ for a set of phase properties, sieve tray geometry, and operating conditions (Haverland, 1988; Garthe, 2006).

Based on the experimental investigations, several modeling approaches for either of the breakage properties $p(d)$, d_{stab}, ϑ were developed in the literature that can be roughly distinguished between purely empirical and physically motivated models. An early empirical correlation was introduced by (Haverland, 1988), which quantifies the breakage behavior based on the breakage probability $p(d)$ as a function of d_{stab}:

$$p(d) = \left(\frac{d - d_{\mathrm{stab}}}{d_{100} - d_{\mathrm{stab}}} \right)^{c} \tag{1-1}$$

with c being an empirical parameter. Apart from d_{stab}, the correlation also contains d_{100}, which describes the minimal diameter of drops that always break ($p(d \ge d_{100}) = 1$) while passing through a sieve tray. The value of d_{100} is determined in single-drop experiments the same way as $p(d)$ and d_{stab}. Pursuing a similar empirical approach, Garthe extended Eq. 1–1 and performed an extensive experimental study with EFCE standard test systems (Mísek, 1985) at varying operating conditions and with different column internals (Garthe, 2006). Based on his investigations, Garthe introduced a further empirical model:

$$p = c_1 \pi_{\mathrm{af}}^{c_2} \frac{d_{\mathrm{trans}}^{c_3}}{c_4 + d_{\mathrm{trans}}^{c_3}} \tag{1-2}$$

$$d_{\mathrm{trans}} = \frac{d - d_{\mathrm{stab}}}{d_{100} - d_{\mathrm{stab}}} \tag{1-3}$$

$$\pi_{\mathrm{af}} = af \cdot \sqrt[3]{\frac{\rho_{\mathrm{c}}^2}{\eta_{\mathrm{c}} (\rho_{\mathrm{c}} - \rho_{\mathrm{d}}) g}} \tag{1-4}$$

Where d is the diameter of the mother drop, ρ_{c} and ρ_{d} are the densities of the continuous and dispersed phase, η_{c} is the viscosity of the continuous phase, and g the gravitational constant. The model's key terms are the droplet term d_{trans}, which is a linear interpolation of the drop diameter d between the diameters d_{stab} and d_{100}, and the solvent-dependent term π_{af}, which quantifies the influence of the pulsation intensity af. The parameters c_1, c_2, c_3, c_4 account for the non-linear form of $p(d)$ in the range $d_{\mathrm{stab}} < d < d_{100}$, also visualized in Figure 2.

For sieve trays, Garthe provides eight parameter sets that account for four solvent systems toluene/acetone/water, butyl acetate/acetone/water, toluene/water, butyl acetate/water (abbreviated by TWA, BWA, TW, BW) and two sieve tray orifice diameters ($d_h = 2\,mm, 4\,mm$). Finally, the model predicts $p(d)$ in Garthe's experimental data with good accuracy and was also used in PBM-based simulations of PSEs (Jaradat et al., 2011a; Jaradat et al., 2011b).

In the two empirical models by Haverland and Garthe, the prediction of the breakage probability consists of a single mathematical expression, wherein d_{stab} is either determined by extrapolation or by experiments. In contrast, the physically motivated approaches focus initially on a correlation for d_{stab} which is subsequently integrated into an overall model for $p(d, d_{stab})$. For example, Wagner considered an energy balance around a drop passing a sieve tray orifice to deduce a correlation for d_{stab} (Wagner, 1994). Subsequently, the breakage probability p is estimated based on the ratio d/d_{stab}. Yet, Wagner's correlation only accounted for drop breakage during orifice passage, which is a valid approach for sieve trays with small free cross-sectional areas but disregards the drop breakage at the web between the orifices for larger free cross-sectional areas (Kalem, 2015). A further physically motivated modeling approach was introduced by (Gourdon et al., 1991). Gourdon's model is based on the Weber number We defined according to Kolmogoroff's turbulence theory by:

$$We = \frac{\rho_c \varepsilon^{2/3} d^{5/3}}{\sigma} \qquad (1-5)$$

with σ as the interfacial tension (Kolmogorov, 1991). The turbulent dissipation rate ε has to be estimated by further correlations (Kumar and Hartland, 1995). Subsequently, the breakage probability p is formulated as a function of We:

$$p(d) = \exp(-c/We). \qquad (1-6)$$

In the final validation of his model, Gourdon showed that Eq 1–6 correctly depicts the trend of the breakage probability over a wide range of operating conditions for the solvent systems TW and TWA despite some considerable deviations for single data sets.

In contrast to pursuing a single correlation for d_{stab} as Wagner and Gourdon, Henschke combined several models for d_{stab} to yield one functional correlation, which accounts for five different drop breakage mechanisms, such as during free sedimentation, during orifice passage, or in turbulent flow (Henschke, 2003). Henschke's overall model for p is subdivided into; first, a model predicting the number of daughter drops as a function of the stable drop diameter $\vartheta(d_{stab})$ and second, a model predicting the breakage probability as a function of the number of daughter drops $p(\vartheta(d_{stab}))$. Taking the three sub-models (d_{stab}, ϑ, and p) into account, the overall breakage model was fit to experimental data from (Haverland, 1988) based on seven parameters. Yet, Henschke's model proved little transferability when being tested on data it was not parametrized on (Kalem, 2015). Based on these findings, Kalem extended Henschke's d_{stab}-model to account for high-viscous solvents, added terms to account for pulsation and the web width b_s (see Figure 3) into the p-model, and introduced four further parameters (Kalem, 2015). Finally, Kalem parametrized and validated his breakage model based on 273 data sets from (Haverland, 1988; Wagner, 1994; Garthe, 2006). Comparably to (Gourdon et al., 1991), Kalem's model could predict the overall trend of the breakage probability with a root mean squared error of 32.5%.

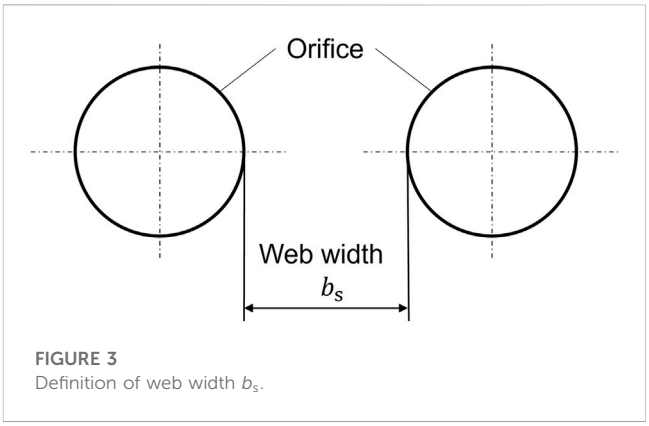

FIGURE 3
Definition of web width b_s.

Most of the modeling approaches for drop breakage in PSE are characterized by a trade-off between a broad validity range and good prediction accuracy. Thereby, good prediction accuracy highly depends on the model's parametrization, which can simultaneously limit the validity range to specific operating conditions, a specific solvent system, and a specific sieve tray geometry. Consequently, every new application of a breakage model would demand a re-parametrization to guarantee good accuracy. In this study, we want to overcome the need for a re-parametrization by introducing a hybrid modeling approach for drop breakage in PSEs. For this purpose, we have chosen a serial hybrid modeling approach (Thompson and Kramer, 1994) in which a physical-empirical model basis is enhanced by data-driven parameter estimation models (PEM). This way, we intend to combine the domain knowledge incorporated in developing physical-empirical models with the accuracy of data-driven models (McBride et al., 2020).

The study is organized as follows. The methods section gives an overview of the error metrics for model evaluation, the data-driven modeling approach, the breakage model, and the database used for model development. In the subsequent section, the results of the PEM development are presented and evaluated, and finally, the overall hybrid breakage model is validated on the breakage database. In a subsequent sensitivity analysis, we assess the hybrid breakage model's ability to predict the breakage behavior for several representative solvent systems, sieve tray geometries, and operating conditions. The final section closes with a brief conclusion and outlook on our future work.

2 Methods

2.1 Error metric

The accuracy of the developed models is evaluated based on different metrics for the error e (also residue), which quantifies the deviation between the experimental value $\hat{u}$ and the predicted value

TABLE 1 Error metrics used in this work.

Name	Equation	
Root mean squared error	$e_{\mathrm{rmse}} = \sqrt{\frac{1}{n_u}\sum_i (u_i - \hat{u}_i)^2}$	4–1
Coefficient of determination	$e_{R2} = 1 - \frac{\sum_i (\hat{u}_i - u_i)^2}{\sum_i (\hat{u}_i - \bar{u})^2}$	4–2
Pull error metric	$e_{\mathrm{pull}} = \frac{u - \hat{u}}{\sigma_e}$	4–3

u. An overview over the error metrics used in this work is given in Table 1, a brief review on the strengths and weaknesses of each error metric is provided in the Supplementary Materials.

In contrast to the first two, the pull metric e_{pull} considers the measurement uncertainty σ_e explicitly, standardizing the residue between experimental and predicted value by σ_e. Thus, the prediction u with a pull value of $-1 \leq e_{\mathrm{pull}} \leq 1$ can be explained within the boundaries of the measurement uncertainty. Considering a whole database where a number of n_u predicted and measured values are compared, the resulting pull distribution e_{pull} can be assessed by its mean $\bar{e}_{\mathrm{pull}}$ and its standard deviation $\tilde{e}_{\mathrm{pull}}$. Consequently, a good model is characterized by a pull distribution with a mean close to zero ($\bar{e}_{\mathrm{pull}} = 0$) and a standard deviation smaller than one ($\tilde{e}_{\mathrm{pull}} < 1$). (Demortier and Lyons, 2008)

2.2 Data flow for machine learning

The development of the data-driven models follows the data flow introduced by (Brockkötter et al., 2020; Brockkötter et al., 2021). The data flow consists of five steps, including (i) *data transformation*, (ii) *data split*, (iii) *machine learning* (ML), (iv) *selection of the data-driven algorithm*, and (v) *wrapper feature selection*. In the first two steps, the database is *transformed* by applying a Min Max scaler and *split* randomly into a train and a test part with a ratio of 85/15. The split is performed once before training and not altered thereafter to ensure comparability of the following development steps. The transformed and split data set is used to train six potential *machine learning* algorithms, including (i) linear regression, (ii) k-Nearest Neighbor, (iii) support-vector-regression, (iv) Gaussian processes, (v) decision tree, and (vi) random forest. In contrast to (Brockkötter et al., 2020), we did not consider Artificial Neural Networks (ANN) as the mismatch between the size of our data sets and the model complexity of ANNs would not justify the use of such a complex algorithm. The training of all ML algorithms is divided into two consecutive steps. First, the training of all ML algorithms is performed based on an exhaustive grid search with k-fold cross-validation with $k = 5$. After each k training step, the e_{rmse} (Eq. 4-1) is determined for each combination of hyperparameters. The combination of hyperparameters resulting in the lowest e_{rmse} is selected for further evaluation. Second, each algorithm undergoes training using the complete training data set, followed by testing on the test set. The algorithm with the lowest e_{rmse} on the test set is then *selected* for further optimization. In the final optimization step, a sequential *feature selection wrapper method* is used to reduce the feature space of the best algorithm. Since the feature selection is not exhaustive, a trade-off is made between the size of the feature set and the quality of the prediction accuracy. This step is necessary to ensure that the validity range, e.g., convex hull, of the PEM is sufficiently large to use the PEM in future column simulations (Kahrs and Marquardt, 2007). Finally, in this work, we use the same wording for differentiating between black-box models (purely data-driven) and white-box models (based on theoretical considerations) as (Brockkötter et al., 2020).

2.3 Revised breakage model

The white-box part of the hybrid breakage model is based on a revised form of Garthe's breakage model. Garthe's model was chosen for this purpose as, in principle, it is not bound to a specific breakage mechanism, but rather replicates the form of the breakage probability between the boundary values d_{stab} and d_{100} (see Figure 2). Consequently, Garthe's model should be applicable under various solvent systems, sieve tray geometries, and operating conditions. Nevertheless, a revision of Garthe's model was conducted to guarantee a physically consistent prediction of p. Particularly, Eq. 1–2 can predict breakage probabilities exceeding $p > 1$ for large values of π_{af} ($c_2 > 0$) or small values of π_{af} ($c_2 < 0$). Additionally, when the drop diameter approaches $d \to d_{100}$ the physically consistent consequence $p = 1$ is not automatically predicted by Eq. 1–2. Considering the detected limitations and by preserving Garthe's modeling idea, we re-arranged the model yielding:

$$p(c_1, c_2, c_3, c_4) = \left(\frac{\pi_{\mathrm{af}}}{c_1 + \pi_{\mathrm{af}}}\right)^{c_2(1-d_{\mathrm{trans}})} \cdot \frac{d_{\mathrm{trans}}^{c_3}}{c_4 + d_{\mathrm{trans}}^{c_3}} \cdot (c_4 + 1) \quad (2-1)$$

For $d_{\mathrm{stab}} < d < d_{100}$ with $d_{\mathrm{trans}} = \frac{d - d_{\mathrm{stab}}}{d_{100} - d_{\mathrm{stab}}}$.

The first term on the right side of Eq. 2–1 limits the $\pi_{\mathrm{af}}-$ term to 1, while the first term's exponent $c_2(1 - d_{\mathrm{trans}})$ guarantees that the breakage probability approaches 1 for $d \to d_{100}$. Additionally, Eq. 2–1 is multiplied by $(c_4 + 1)$, so that the second term on the right side also approaches 1 for $d \to d_{100}$. Therefore, the new form of the breakage model ensures the model's continuity within the physical constraints $p(d = d_{\mathrm{stab}}) = 0$ and $p(d = d_{100}) = 1$ for all parameters $c_1, c_2, c_3, c_4 \in \mathbb{R}_{>0}$.

2.4 Breakage database

The breakage database for PSEs consists of 743 data sets retrieved from literature (Haverland, 1988; Eid et al., 1991; Wagner, 1994; Garthe, 2006). One data set corresponds to one entry in the breakage database which specifies the experimental value of the breakage probability p (target feature) characterized by 15 features including the solvent system, the sieve tray geometry, and the operating conditions. Table 2 provides an overview over the features and the statistical information of the database. The complete database is provided in the Supplementary Materials.

TABLE 2 Statistical information on the breakage database.

Name	Unit	Mean	Std	Min	25%	50%	75%	Max
Breakage probability, p	-	0.48	0.42	0.00	0.00	0.49	0.96	1.00
Density continuous phase, ρ_c	kg/m^3	1019.8	73.1	990.9	997.0	998.2	998.8	1204.2
Density dispersed phase, ρ_d	kg/m^3	870.9	98.8	726.4	858.3	866.7	877.5	1246.0
Viscosity continuous phase, η_c	Pa·s	0.006	0.015	0.001	0.001	0.001	0.001	0.050
Viscosity dispersed phase, η_d	Pa·s	0.008	0.018	0.001	0.001	0.001	0.001	0.053
Interfacial tension, σ	N/m	0.031	0.010	0.011	0.024	0.035	0.035	0.052
Diameter of mother drop, d	m	0.004	0.002	0.0003	0.002	0.004	0.005	0.010
Pulsation amplitude, a	m	0.009	0.005	0.004	0.008	0.008	0.010	0.018
Pulsation frequency, f	1/s	2.4	1.4	1.3	1.9	2.5	2.5	10.0
Pulsation intensity, af	m/s	0.021	0.008	0.003	0.015	0.020	0.025	0.048
Velocity continuous phase, u_c	m/s	0.001	0.003	0.000	0.000	0.000	0.000	0.025
Column diameter, D_{col}	m	0.072	0.008	0.0472	0.07245	0.07245	0.079	0.079
Orifice diameter of sieve tray, d_h	m	0.004	0.002	0.002	0.002	0.004	0.004	0.012
Relative free cross-sectional area, φ	-	0.31	0.12	0.07	0.23	0.26	0.40	0.60
Web width, b_s	m	0.002	0.001	0.002	0.002	0.002	0.003	0.005
Tray spacing, h_{st}	m	0.082	0.032	0.05	0.05	0.1	0.1	0.2
Sieve tray material: 100% stainless steel								

The experimental values for p are determined in single-drop cells (see Figure 1) which partly differ in diameter and height. Considering that the drop breakage appears primarily in the direct vicinity of the sieve tray orifice (Haverland, 1988), we assume that the difference in diameter and height does not affect the comparability of the experimental studies with each other. Most of the authors performed at least 50 single drop measurements to determine p for one specific solvent system, sieve tray geometry, and operating condition (Haverland, 1988; Wagner, 1994; Garthe, 2006). The definition of the experimental values for d_{stab} and d_{100} vary throughout the database. After investigating the breakage behavior for a range of drop diameters, Haverland and Wagner performed 50 additional single-drop experiments to define d_{stab} as $p(d) = 0$, and d_{100} as $p(d) = 1$. Instead, Garthe defines d_{stab} for $p(d) \leq 0.03$ and d_{100} for $p(d) \geq 0.97$. In principle, Garthe's definition does not contradict the previous one, it rather accounts for the stochastic effects in the experimental determination of d_{stab}/d_{100}. In contrast, Eid et al. did not explicitly specify their approach to determine either p or d_{stab}/d_{100} (Eid et al., 1991). Considering that the research groups of (Haverland, 1988; Eid et al., 1991) worked partly together, e.g., (Haverland et al., 1987), we assume that the experimental approach did not substantially deviate from (Haverland, 1988; Wagner, 1994; Garthe, 2006).

3 Results and discussion

3.1 Hybrid breakage model

The revised breakage model introduced in section 2.3 poses the prerequisite for a robust breakage model, due to a physically consistent prediction of drop breakage between d_{stab} and d_{100}. Two limitations remaining in the revised model are (i) the limited availability of experimental data for d_{stab} and d_{100} and (ii) the discrete validity of the model parameters c_1, c_2, c_3, c_4 which demand additional experiments and/or re-parametrization for each new set of solvent properties and operating conditions.

In the following, we address these limitations by introducing PEM for d_{stab}, d_{100}, and c_1, c_2, c_3, c_4,. For this purpose, the breakage database is split into three parts. First, the database is filtered for entries with $p \leq 0.03$ and $p \geq 0.97$, yielding two data sets for d_{stab} (235 data sets) and d_{100} (188 data sets), respectively. The remaining entries for $0.03 < p < 0.97$ (320 data sets) constitute the database for the consecutive parametrization of c_1, c_2, c_3, c_4. Figure 4 visualizes the data split and the approach for PEM development.

In the following, the results of the PEM development for d_{stab} and d_{100} (Sec. 3.1.1) and for c_1, c_2, c_3, c_4 (Sec. 3.1.2) are presented and discussed. Finally, the PEMs are incorporated in the model basis (Eq. 2–1), and the hybrid model is validated on the complete breakage database (Sec. 3.1.3).

3.1.1 Modeling of d_{stab} and d_{100}

The data sets for d_{stab} and d_{100} are inserted into the data flow described in section 2.2 to develop two separate data-driven PEMs for each diameter. The process was performed several times, reducing the feature set during each training. The results are summarized in Table 3. The final models enable the prediction of the characteristic properties d_{stab} and d_{100} based on two reduced feature sets, including one property describing the operating condition (af), one property describing the solvent system (σ), and two properties characterizing the

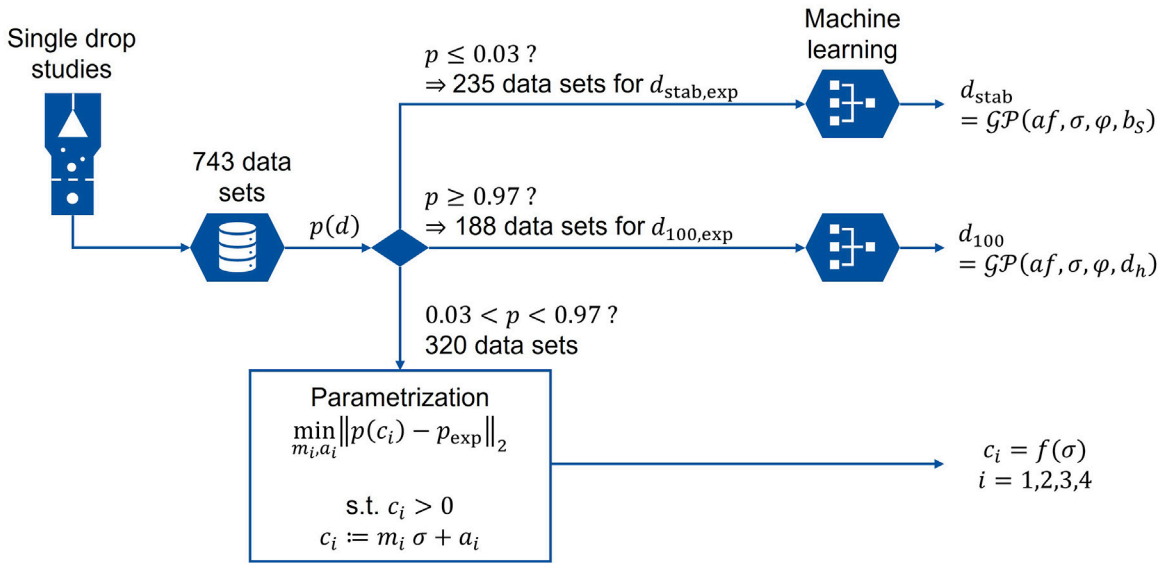

FIGURE 4
Dataflow for the development of the hybrid breakage model.

TABLE 3 Final data-driven models for d_{stab} and d_{100}. The error metrics refer exclusively to the test set.

	Features	e_{rmse} [mm]	e_{R2} [-]	Best algorithm
d_{stab} [mm]	af, σ, φ, b_s	0.461	0.916	Gaussian Process Regressor
d_{100} [mm]	af, σ, φ, d_h	0.463	0.847	Gaussian Process Regressor

sieve tray geometry (φ and b_s/d_h, respectively). For both diameters, Gaussian Process (GP) Regressors were identified as the best models with e_{rmse} of 0.461 *mm* and 0.463 *mm* and e_{R2} of 0.916 and 0.847, respectively.

The feature set included in the final models agrees well with most physical-empirical correlations for d_{stab} reported in the literature. Particularly, the features depicting the operating conditions (af) and the sieve tray geometry (φ and b_s/d_h) are also considered by (Wagner, 1994; Henschke, 2003; Kalem, 2015). In contrast, the characterization of the solvent systems deviates from most models reported in the literature. While the wrapper feature selection selects the interfacial tension, the densities of both phases are disregarded in the final models. Obviously, the feature selection mirrors the fact that most of the experiments in the data sets were conducted with water (68.6%) as the continuous phase and toluene (47,3%)/butyl acetate (11.9%) as the dispersed phase. A similar trend is investigated for d_{100}, the according data set is also dominated by water (90.0%) and toluene (51.1%)/butyl acetate (18.4%). This problem is compounded by toluene and butyl acetate having comparably similar densities (compare Table 6). Consequently, the wrapper feature selection does not detect a benefit in keeping the densities for the final model. Brockkötter et al. (2020) yielded similar results and drew the same conclusion for their data-driven flooding model. Nevertheless, the densities of both phases are included in the white-box part (Eq. 2–1) of the final breakage model; hence, the absence of the densities in the feature space of d_{stab} and d_{100} is mitigated for the prediction of the breakage probability after all.

3.1.2 Continuous parameter estimation for c_1, c_2, c_3, c_4

For the development of PEMs for the breakage model parameters c_1, c_2, c_3, c_4, a different approach was pursued than for d_{stab} and d_{100}. Particularly, the introduction of four black-boxes or one multi-output black-box for c_1, c_2, c_3, c_4 could increase the feature space furthermore. In order to limit the feature space, we chose a linear correlation, which predicts the parameters c_1, c_2, c_3, c_4 depending on one characteristic feature. Considering Garthe's eight parameter sets for the original model, the characteristic feature could be a property describing the solvent system or characterizing the sieve tray geometry. For this purpose, the interfacial tension σ seems as a consistent choice as it may be considered characteristic of a solvent system, particularly regarding the breakage behavior of the dispersed phase. Implicitly, the influence of the interfacial tension was accounted by Garthe's parameter sets which were adapted individually to the solvent system TW, BW, TWA, and BWA. Specifically, these solvent systems can be discriminated between each other mainly due to different interfacial tensions (Mísek, 1985). Therefore, the fitting parameters are formulated as a linear function of the interfacial tension σ: $c_i = m_i \cdot \sigma + a_i$ for $i = 1, 2, 3, 4$, creating two degrees of freedom per parameter c_i, which are determined by the optimization problem:

$$\min_{m_i,a_i} \left\| p(c_1, c_2, c_3, c_4) - \hat{p} \right\|_2 \qquad (3-1)$$

TABLE 4 Results of constant and linear parametrization of the hybrid breakage model and residues of the hybrid breakage model and Garthe's model. The original parameterization of Garthe's model depends on multiple influencing factors. Thus, the entries are marked with the non-zero marker ∗. The numerical values for ∗ are listed in the Supplementary Material.

Equation	c_1		c_2		c_3		c_4		Residue	
	m_1	a_1	m_2	a_2	m_3	a_3	m_4	a_4	e_{rmse} [-]	e_{R2}
1-2 Garthe	0	*	0	*	0	*	0	*	0.198	0.515
2-1 constant	0	2.64	0	0.86	0	1.44	0	0.06	0.148	0.630
2-1 linear	−36.07	3.12	30.22	0.06	−5.00	1.55	−4.61	0.24	0.139	0.761

TABLE 5 Error metrics of Garthe's and the hybrid breakage model for the prediction of the breakage probability on the complete breakage database.

	e_{rmse}	e_{R2}	$\bar{e}_{pull}$	$\tilde{e}_{pull}$
Garthe (2006)	0.221	0.723	0.634	2.116
Hybrid Model	0.132	0.902	−0.068	1.315

w.r. $c_i > 0$ with $c_i = m_i \cdot \sigma + a_i$ for $i = 1, 2, 3, 4$, with $\hat{p}$ being the experimental breakage probability and $p(c_1, c_2, c_3, c_4)$ the predicted breakage probability by Eq. 2–1. For the optimization, values for d_{stab} and d_{100} have to be assigned at each experimental data set. Data sets where no experimental values for d_{stab} and d_{100} have been determined yet are estimated by the PEMs introduced in the previous section. The optimization problem in Eq. 3–1 searches for the best combination of four linear models within the framework of the breakage model Eq. 2–1, which describes the measurement $\hat{p}$. The best set of linear functions c_1, c_2, c_3, c_4 is realized by a grid search in Matlab™. For this purpose, an eight-dimensional search grid of starting values m_i and a_i is selected, and the objective Eq. 3–1 is optimized by the solver *fminsearch*, which is based on the simplex method (Lagarias et al., 1998). The smallest local minimum is selected from the solutions of the grid.

The results of the parameterization are summarized in Table 4. The table shows the eight parameters m_i and a_i of the *linear parameterization* for c_i and the deviation between the predicted and the measured breakage probability in terms of e_{rmse} and e_{R2}. In addition to the *linear*, a *constant parametrization* was performed where m_i was set to 0; thus, the interfacial tension is not included in this parameterization (middle row). For consistency, the data set for parametrization was also predicted by *Garthe's* original model to compare its prediction accuracy with the new model. The residues given for *Garthe* in Table 4 refer exclusively to the solvent systems TW, BW, TWA, and BWA for which Garthe provided parameter sets. Since the original parameterization of *Garthe's* model depends on multiple influencing factors, the entries are marked with the non-zero marker ∗. The numerical values for ∗ are listed in the Supplementary Material S1. Despite covering a considerably wider range of solvent systems and operating conditions, the parametrization results of the revised models surpass *Garthe's* model on both error metrics e_{rmse} and e_{R2}. Even the *constant parameterization* shows a significant improvement compared to *Garthe's* model, which is enhanced by including the interfacial tension in the *linear* model. The *linear parametrization* based on the interfacial tension improves both error metrics by approximately 30% compared to Garthe's model. Overall, the results indicate that the revised model (Eq. 3–1) in combination with the linear PEM can depict the breakage probability for $0 < p < 1$ ($d_{stab} < d < d_{100}$).

3.1.3 Model validation

In the final step of the model development, the revised breakage model is extended by the data-driven PEM for d_{stab} and d_{100} and the linear PEM for c_1, c_2, c_3, c_4. The new hybrid breakage model is validated on the complete breakage database consisting of 743 data sets and compared to Garthe's model. The error metrics assess both models' prediction accuracy e_{rmse} and e_{R2} as well as the mean $\bar{e}_{pull}$ and the standard deviation $\tilde{e}_{pull}$ of the pull distribution. To calculate the pull, we assume a constant measurement uncertainty of $\sigma_e = 0.1$ for the experimental breakage probability. This assumption is necessary as (to the best of our knowledge) no author has specified a value for the uncertainty and/or standard deviation of the experimentally determined breakage probability. The error metrics of Garthe's and the hybrid breakage model are compared in Table 5. Additionally, the pull distributions for both models are visualized in Figure 5.

Regarding the pull distribution, both models are centered around 0 and have a qualitatively reasonable spread around the center point. The hybrid model predicts a larger fraction of the database within the assumed measurement accuracy, while the variance is obviously smaller. The qualitative observation is numerically confirmed: the hybrid model achieves a prediction accuracy close to the optimal value of 0, underestimating it slightly ($\bar{e}_{pull} = -0.068$), while Garthe's model tends to overestimate the breakage probability ($\bar{e}_{pull} = 0.634$). A similar trend is observed for the standard deviation $\tilde{e}_{pull}$: the hybrid model achieves $\tilde{e}_{pull} = 1.315$, which is lower than Garthe's model with a value of $\tilde{e}_{pull} = 2.116$ and closer to optimum. Apart from the pull metric, the hybrid model surpasses Garthe's model also regarding e_{rmse} (40.2% lower than Garthe) and e_{R2} (24.8% higher than Garthe).

Finally, a qualitative evaluation based on Figure 5 indicates that both models can correctly depict the experimental breakage probability in the database. Nevertheless, the hybrid model achieves better scores in all error metrics and has an extended validity range due to the PEMs for d_{stab}, d_{100} and c_1, c_2, c_3, c_4. Based on these results, we deduce that the hybrid breakage model was successfully validated on the breakage database. Therefore, in the

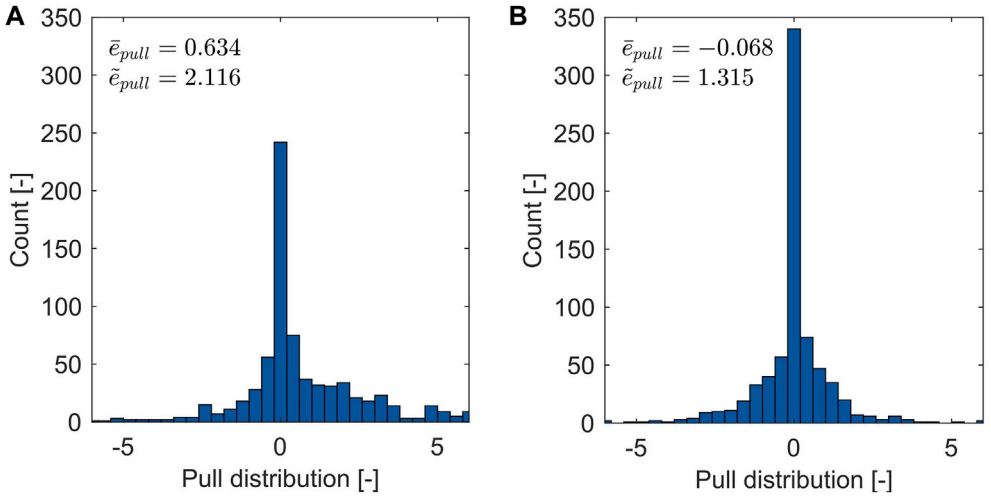

FIGURE 5
Pull distribution of Garthe's **(A)** and the hybrid breakage model **(B)** for the prediction of the breakage probability on the complete breakage database.

TABLE 6 Properties of the solvents considered in the sensitivity analysis.

Reference	Solvent system Organic/Aqueous phase	Density [kg/m^3]		Viscosity [$mPa \cdot s$]		Interfacial tension [mN/m]	Mass transfer direction	Figure
		Aqueous phase	Organic phase	Aqueous phase	Organic phase			
Haverland (1988)	Toluene/water	998.2	866.7	1.003	0.586	35.40	-	Figure 6 a Figure 7 (a)
	Butyl acetate/water	998.2	881.5	1.003	0.730	13.50	-	Figure 6 b
Garthe (2006)	Toluene/water	998.8	867.5	1.029	0.596	34.31	-	Figure 7 (a) & (b)
	Toluene/acetone/water	992.0	863.3	1.134	0.566	24.41	c → d	Figure 7 (b)
	Butyl acetate/water	996.4	881.3	1.022	0.738	13.97		Figure 7 (b)
	Butyl acetate/acetone/water	990.2	877.5	1.163	0.709	10.96	c → d	Figure 7 (b)
Wagner (1994)	Paraffin oil (53%) + n-hexane (47%)/water	999.2	745.6	1.001	1.230	40.08	-	Figure 8
	Paraffin oil 63.75%, n-hexane 36.25%/water	999.2	858.3	1.001	53.200	43.70	-	Figure 8
	Paraffin oil 53%, n-hexane 47%/water 21.3%, glycerine 78.7%	1204.2	745.6	50.080	1.230	31.00	-	Figure 8
	Paraffin oil 63.75%, n-hexane 36.25%/water 21.3%, glycerine 78.7%	1204.2	858.3	50.080	53.200	32.20	-	Figure 8

following result section, the discussion focuses exclusively on the hybrid breakage model.

3.2 Sensitivity analysis

The purpose of the sensitivity analysis is to test the model's ability to predict physically consistent trends of d_{stab}, d_{100} and $p(d)$ based on several representative and coherent data sets from different authors. The data sets cover experiments with the EFCE systems TW, BW, TWA, BWA by (Haverland, 1988; Garthe, 2006), as well as experiments with deliberately adjusted viscosities of the solvents by (Wagner, 1994). The according solvent properties are summarized in Table 6. For a complete overview of the data sets, the reader is referred to the breakage database in the Supplementary Material S1. The sensitivity analysis is primarily based on a qualitative comparison between experimental data (symbols) and model prediction

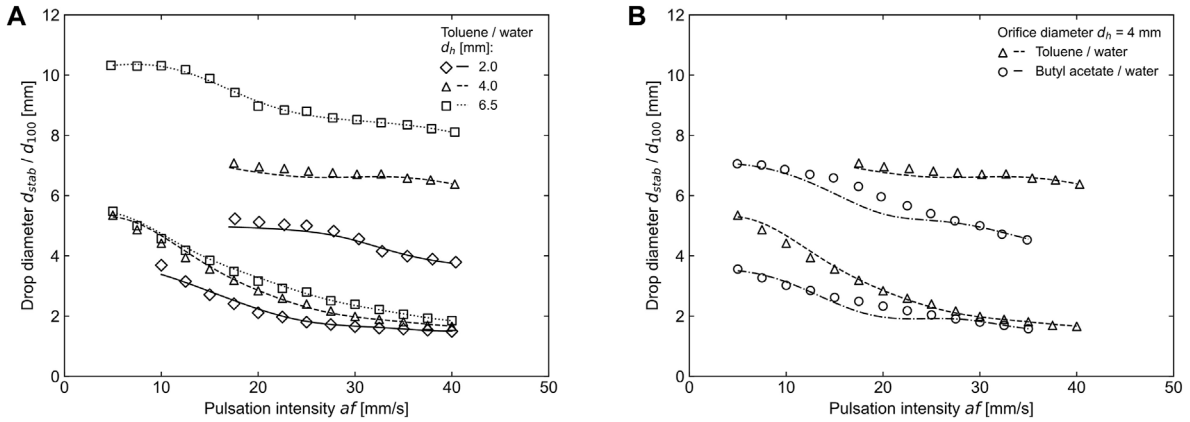

FIGURE 6
(A) Comparison of d_{stab}, and d_{100} for toluene/water for different orifice diameters d_h and pulsation intensities af; and **(B)** comparison of d_{stab}, and d_{100} for toluene/water and butyl acetate/water. Experimental data: symbols, prediction: lines. The upper lines indicate d_{100}. Figures according to (Haverland, 1988).

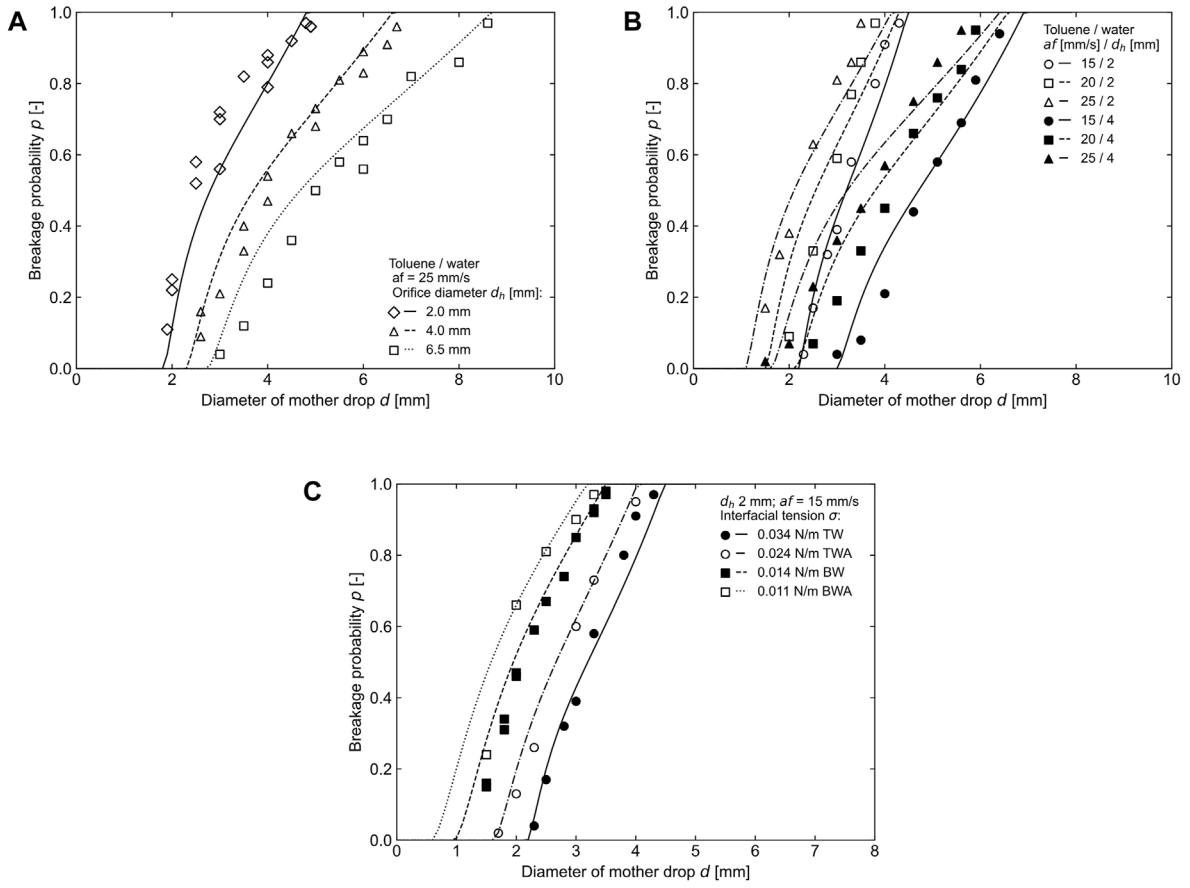

FIGURE 7
Breakage probability for toluene/water for **(A)** different orifice diameters, **(B)** different orifice diameters and pulsation intensities, and **(C)** for different interfacial tensions. Experimental data: symbols, prediction: lines. Figures according to (Haverland, 1988; Garthe, 2006).

(lines) in Figures 6–8. Each figure highlights one or several characteristic features, which are indicated in the figure key. Additionally, Table 7 summarizes the e_{rmse} for the predicted sub-data sets of each Figure. At first, the discussion of the results focuses on the EFCE system for which d_{stab}/d_{100} and $p(d)$ are visualized in Figures 6, 7, respectively.

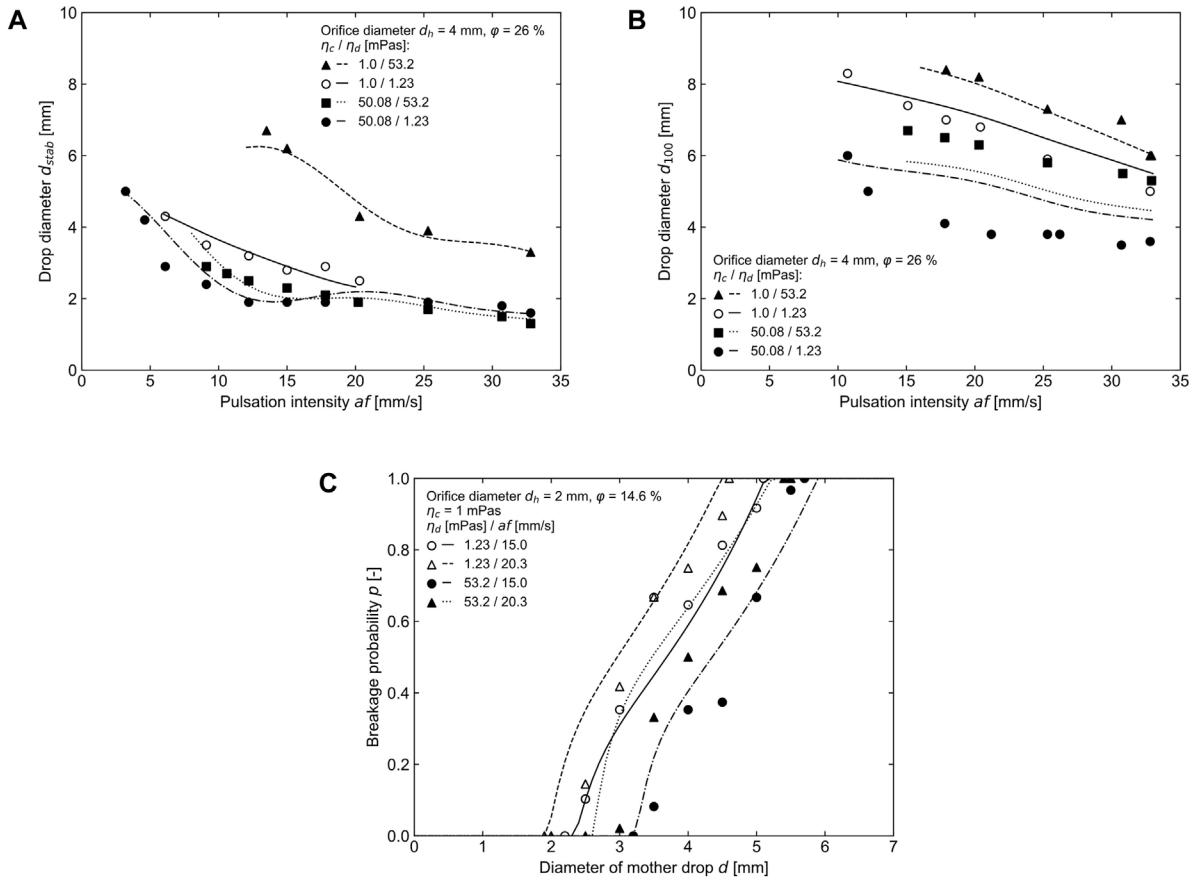

FIGURE 8
Comparison of **(A)** d_{stab}, **(B)** d_{100}, and **(C)** the breakage probability for varying viscosities. Experimental data: symbols, prediction: lines. Figures according to (Wagner, 1994).

Three overall t rends can be deduced from the experimental data:

- Trend 1: Higher af values lead to a decrease in d_{stab}, and d_{100}, thus the $p(d)$ curve is shifted to the left (Figure 6A, B; Figure 7B)
- Trend 2: Lower σ values lead to a decrease in d_{stab}, and d_{100}, thus the $p(d)$ curve is shifted to the left (Figure 6B; Figure 7C)
- Trend 3: Higher d_h/φ values lead to an increase in d_{stab}, and d_{100} thus the $p(d)$ curve is shifted to the right, the slope of the $p(d)$ curve decreases (Figure 6A; Figure 7A, B).

Trends 1 and 2 can be accounted for by a force balance of stabilizing and disruptive forces acting on the drop surface and deforming the shape of the drop. In principle, the force balance is incorporated in the We number, which represents the ratio of the fluid's inertia compared to its surface tension (compare 1-5). Hence, an increasing energy input (increasing af) shifts the equilibrium towards the disruptive forces, e.g., increasing We (trend 1), while a decreasing interfacial tension σ reduces the stabilizing forces, e.g., increasing We (trend 2). Trend 3 can be attributed to an increasing steric hindrance for small d_h/φ values, causing more drop breakage if the drop diameter surpasses the orifice diameter.

Considering the prediction of d_{stab}, d_{100} and $p(d)$, the experimental trends are correctly predicted by the according PEM and the overall hybrid breakage model, which is also reflected in the e_{rmse} values, respectively. The experimental trends are also reflected in the features for the PEM, e.g., af, σ, and φ,

TABLE 7 Root mean squared error e_{rmse} of the predictions in Figure 6, Figure 7, and Figure 8.

Diagrams	Property	Unit	e_{rmse}
Figure 6A	d_{stab}	[mm]	0.100
	d_{100}	[mm]	0.117
Figure 6B	d_{stab}	[mm]	0.145
	d_{100}	[mm]	0.258
Figure 7B	$p(d)$	[-]	0.082
Figure 7B	$p(d)$	[-]	0.083
Figure 7C	$p(d)$	[-]	0.072
Figure 8A	d_{stab}	[mm]	0.277
Figure 8B	d_{100}—total	[mm]	0.692
	d_{100}—high η_c	[mm]	0.880
	d_{100}—low η_c	[mm]	0.367
Figure 8C	$p(d)$	[-]	0.114

which were detected as pivotal to model drop breakage of the EFCE systems in previous works (see also Section 3.1.1).

In contrast, Wagner's data sets include rather unconventional solvent systems, which can be distinguished primarily based on the viscosities η_c, η_d and not as commonly in extraction research based on the interfacial tension σ. It is worth highlighting that neither the PEM for d_{stab} nor the PEM for d_{100} include η_c, η_d as features. Regarding solvent properties, the distinction is based only on the interfacial tension σ, which varies between Wagner's solvent systems too, but not as distinctively as for the EFCE systems (see Table 6). The according experimental values for d_{stab}, d_{100} and $p(d)$ and their prediction are visualized Figure 8.

Apart from the trends already discussed for the EFCE systems, one additional trend can be deduced from Wagner's experiments.

- Trend 4: Higher η_c values lead to a decrease in d_{stab}, and d_{100}.

Wagner argues that a highly viscous continuous phase increases the shear stress on the drops during orifice passage, increasing the tendency to drop breakage (see d_{stab} in Figure 8 (a)). In contrast, a high viscous dispersed phase rather stabilizes the drop. Thus, the combination of a high η_d and a low η_c leads to the largest d_{stab}, and d_{100} values, which is comparable to the conclusions drawn from fluid dynamic studies in agitated systems (Stamatoudis and Tavlarides, 1985; Kraume et al., 2004; Maaß et al., 2012; Hasan, 2017).

Although both viscosities η_c, η_d are not included in the feature set for the PEM the qualitative experimental trends are predicted correctly. In comparison to EFCE systems, the d_{stab}/d_{100} prediction is obviously less accurate. However, while the prediction of d_{stab} is in a comparable range as for EFCE systems ($e_{rmse} = 0.277mm$), the prediction of d_{100} deviates significantly ($e_{rmse} = 0.692mm$), particularly for high η_c ($e_{rmse} = 0.880mm$, see Table 7). Potentially, the ratio of Wagner's solvent system in the training data for d_{100} does not suffice to depict the physical effects determining d_{100}. Afterall, in the d_{100} database, only 30% (56/188) of the data sets are not EFCE systems, while in the d_{stab} database at least 43% (103/235) of the data sets are not EFCE systems.

The validity of the $p(d)$ trend is more difficult to assess since the experimental data shows a certain degree of ambiguity, e.g., in some cases, the data sets overlap for different solvent systems. Considering these limitations, the prediction of $p(d)$ covers at least the same range and reflects the same trends as the experimental data.

Overall, the prediction of the hybrid breakage model is consistent with the trends in the experimental data for d_{stab}, d_{100} and $p(d)$. Therefore, we conclude that the hybrid breakage model correctly depicts the underlying physical effects that cause drop breakage. Regarding the EFCE systems, the prediction of the key properties describing drop breakage d_{stab}, d_{100} and $p(d)$ can be assessed as quantitatively and qualitatively good for the considered operating conditions and sieve tray geometries. For highly viscous solvent systems, the prediction accuracy is worse than for the EFCE systems. Considering the smaller share in the training databases for d_{stab}/d_{100}, the prediction can still be considered as reasonable for d_{stab} and for d_{100} at low η_c, and moderate for d_{100} at high η_c and for $p(d)$. To improve the prediction of d_{100} and $p(d)$, future experimental studies of drop breakage should focus on solvents beyond the EFCE systems to increase their share in the PEM training data.

4 Conclusion

Within this study, we have developed a hybrid breakage model for PSEs based on a serial hybrid modeling approach. The hybrid breakage model consists of an empirical model basis which is enhanced by data-driven PEMs. For the model basis, Garthe's breakage model was revised to guarantee a physically consistent prediction of the breakage probability p. To establish the data-driven PEM, an experimental database for breakage in PSEs was retrieved from the literature, including data sets for d_{stab}, d_{100} and $p(d)$. The database consists of 743 data sets, including various solvent systems and sieve tray geometries and covering PSE-typical operating conditions. Based on according subsets of the database, two data-driven PEMs based on GPs were developed to predict the experimental parameters d_{stab} and d_{100}. The feature space of both PEMs could be limited to only four features while achieving a good accuracy of $e_{R2} = 0.916$ for d_{stab} and $e_{R2} = 0.847$ for d_{100}. In addition to that, the four model parameters c_1, c_2, c_3, c_4 were formulated as a linear function of the interfacial tension σ. Finally, the linear and the data-driven PEMs were incorporated into the model basis to form the hybrid breakage model.

For validation purposes, the complete breakage database was predicted by Garthe's and the hybrid breakage model, respectively. Subsequently, the accuracy of both models was compared based on the pull distribution of their predictions. Thereby, the hybrid model not only surpassed Garthe's model regarding the prediction error (hybrid: $e_{R2} = 0.9$, Garthe: $e_{R2} = 0.72$), but also achieved pull scores close to optimum, e.g., mean $\bar{e}_{pull} = -0.068$ and standard deviation $\tilde{e}_{pull} = 1.315$. A sensitivity analysis demonstrated that the hybrid breakage model correctly predicts the experimental trends of d_{stab}, d_{100} and $p(d)$ indicating that the underlying physical effects are accurately considered. In future studies, we plan to integrate the hybrid breakage model into a PBM-based PSE model.

Regarding the research on drop breakage, we would suggest two aspects to be considered for future studies. As demonstrated in Section 3.2, a limited variety in the training data of the data-driven PEM can substantially affect the quality of the model. Consequently, future studies should focus on solvent systems beyond the EFCE solvents. Special effort should be put in the investigation of highly viscous solvents as was done for agitated systems in recent years. The second aspect we would suggest concerns the model development itself. In this study, we focused primarily on the reduction of the number of features considered for the PEM development. The reduction of the feature space addresses the problem that every additional feature in a data-driven model improves the error metrics, yet simultaneously reduces the validity range, e.g., convex hull. In future studies, a reduction of the number of parameters should be considered as a possible objective too. For example, the breakage probability might be expressed by a simpler breakage model eventually with fewer parameters than c_1, c_2, c_3, c_4. Additionally, instead of linearizing all parameters c_1, c_2, c_3, c_4, the linearization could also be limited to only one parameter. Consequently, the final objective should be a trade-off between a broad validity range, a good model accuracy and a simple model for drop breakage at sieve trays.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

AP: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. JR: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing–review and editing, Writing–original draft, Validation. HG: Data curation, Methodology, Writing–review and editing, Software. AJ: Supervision, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fceng.2023.1274349/full#supplementary-material

References

Brockkötter, J., Ahndorf, J., and Jupke, A. (2021). Prediction of flooding in packed liquid-liquid and high-pressure extraction columns using a Gaussian process. *Chem. Ing. Tech.* 93, 1907–1916. doi:10.1002/cite.202100073

Brockkötter, J., Cielanga, M., Weber, B., and Jupke, A. (2020). Prediction and characterization of flooding in pulsed Sieve Plate Extraction columns using data-driven models. *Ind. Eng. Chem. Res.* 59, 19726–19735. doi:10.1021/acs.iecr.0c03282

Cramer, E., and Kamps, U. (2017). *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik: Eine Einführung für Studierende der Informatik, der Ingenieur-und Wirtschaftswissenschaften*. Berlin: Springer Spektrum.

Dahmen, W., and Reusken, A. (2022). *Numerik für Ingenieure und Naturwissenschaftler*. Berlin [u.a.]: Springer.

Demortier, L., and Lyons, L. (2008). *Everything you always wanted to know about pulls*.

Eid, K., Gourdon, C., Casamatta, G., and Muratet, G. (1991). Drop breakage in a pulsed sieve-plate column. *Chem. Eng. Sci.* 46, 1595–1608. doi:10.1016/0009-2509(91)87008-Z

Gameiro, M. L. F., Machado, R. M., Ismael, M. R. C., Reis, M. T. A., and Carvalho, J. M. R. (2010). Copper extraction from ammoniacal medium in a pulsed sieve-plate column with LIX 84-I. *J. Hazard. Mater.* 183, 165–175. doi:10.1016/j.jhazmat.2010.07.006

Garthe, D. (2006). [München]: Technical University of Munich. Fluiddynamics and Mass Transfer of Single Particles and Swarms of Particles in Extraction Columns. [Dissertation].

Gourdon, C., Casamatta, G., and Angelino, H. (1991). Single drop experiments with liquid test systems: a way of comparing two types of mechanically agitated extraction columns. *Chem. Eng. J.* 46, 137–148. doi:10.1016/0300-9467(91)87005-U

Hasan, B. O. (2017). Breakage of drops and bubbles in a stirred tank: a review of experimental studies. *Chin. J. Chem. Eng.* 25, 698–711. doi:10.1016/j.cjche.2017.03.008

Haverland, H. (1988). *Untersuchungen zur Tropfendispergierung in flüssiigkeitspulsierten Siebboden-Extraktionskolonnen. [Dissertation]*. Clausthal: Technical University of Clausthal.

Haverland, H., Vogelpohl, A., Gourdon, C., and Casamatta, G. (1987). Simulation of fluid dynamics in a pulsed sieve plate column. *Chem. Eng. Technol.* 10, 151–157. doi:10.1002/ceat.270100119

Henschke, M. (2003). *Auslegung pulsierter siebboden-extraktionskolonnen. [Habilitation]*. [Aachen]: RWTH Aachen University.

Hlawitschka, M. W., Schulz, J., Wirz, D., Schäfer, J., Keller, A., and Bart, H.-J. (2020). Digital Extraction Column: measurement and modeling techniques. *Chem. Ing. Tech.* 92, 914–925. doi:10.1002/cite.202000043

James, G., and Witten, D. (2013). *#x88;An introduction to statistical learning: with applications in R*. New York: Springer. [etc.].

Jaradat, M., Attarakih, M., and Bart, H.-J. (2011a). Advanced prediction of pulsed (packed and sieve plate) extraction columns performance using population balance modelling. *Chem. Eng. Res. Des.* 89, 2752–2760. doi:10.1016/j.cherd.2011.05.009

Jaradat, M., Attarakih, M., and Bart, H.-J. (2011b). Population balance modeling of pulsed (packed and sieve-plate) extraction columns: coupled hydrodynamic and mass transfer. *Ind. Eng. Chem. Res.* 50, 14121–14135. doi:10.1021/ie201041q

Jareš, J., and Procházka, J. (1987). Break-up of droplets in Karr reciprocating plate extraction column. *Chem. Eng. Sci.* 42, 283–292. doi:10.1016/0009-2509(87)85058-3

Kahrs, O., and Marquardt, W. (2007). The validity domain of hybrid models and its application in process optimization. *Chem. Eng. Process. Process Intensif.* 46, 1054–1066. doi:10.1016/j.cep.2007.02.031

Kalem, M. (2015). *Einzeltropfenbasierte Simulation von pulsierten Siebbodenextraktionskolonnen für die Reaktivextraktion. [Dissertation]*. Aachen: RWTH Aachen University.

Kolmogorov, A. N. (1991). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proc. R. Soc. Lond. A* 434, 9–13. doi:10.1098/rspa.1991.0075

Kraume, M., Gäbler, A., and Schulze, K. (2004). Influence of physical properties on drop size distribution of stirred liquid-liquid dispersions. *Chem. Eng. Technol.* 27, 330–334. doi:10.1002/ceat.200402006

Kumar, A., and Hartland, S. (1995). A unified correlation for the prediction of dispersed-phase hold-up in liquid-liquid extraction columns. *Ind. Eng. Chem. Res.* 34, 3925–3940. doi:10.1021/ie00038a032

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM J. Optim.* 9, 112–147. doi:10.1137/S1052623496303470

Maaß, S., Rojahn, J., Hänsch, R., and Kraume, M. (2012). Automated drop detection using image analysis for online particle size monitoring in multiphase systems. *Comput. Chem. Eng.* 45, 27–37. doi:10.1016/j.compchemeng.2012.05.014

McBride, K., Sanchez Medina, E. I., and Sundmacher, K. (2020). Hybrid semi-parametric modeling in separation processes: a review. *Chem. Ing. Tech.* 92, 842–855. doi:10.1002/cite.202000025

R. Goedecke (Editor) (2006). *Fluidverfahrenstechnik: geschichte, methodik, praxis, technik* (Weinheim, Chichester: Wiley-VCH Verlag GmbH; John Wiley & Sons).

Schügerl, K. (1994). *Solvent extraction in biotechnology: recovery of primary and secondary metabolites*. Springer-Verlag.

Stamatoudis, M., and Tavlarides, L. L. (1985). Effect of continuous-phase viscosity on the drop sizes of liquid-liquid dispersions in agitated vessels. *Ind. Eng. Chem. Proc. Des. Dev.* 24, 1175–1181. doi:10.1021/i200031a047

Thompson, M. L., and Kramer, M. A. (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE J.* 40, 1328–1340. doi:10.1002/aic.690400806

T. C. Lo, M. H. I. Baird, and C. Hanson (Editors) (1983). *Handbook of solvent extraction* (New York: Wiley).

T. Mísek (Editor) (1985). *Standard test systems for liquid extraction* (Rugby: Institution of Chemical Engineers).

Wagner, G. (1994). *Der Einfluss der Viskosität auf die Strömung in Apparaten für die Flüssig-flüssig-Extraktion. [Dissertation]*. [Munich]: Technical University of Munich.

Weber, B., and Jupke, A. (2020). Compartment-model for the simulation of the separation performance of stirred liquid–liquid-extraction columns. *AIChE J.* 66. doi:10.1002/aic.16286

Weber, B., Meyer, C., and Jupke, A. (2019). Performance map for the design of liquid-liquid extraction columns. *Chem. Ing. Tech.* 91, 1674–1680. doi:10.1002/cite.201900057

Nomenclature

Symbols

a	[m]	Amplitude of pulsation
af	[m/s]	Pulsation intensity
b_s	[m]	Orifice spacing
c	[-]	Constant model parameter
d	[m]	Drop diameter
d_{stab}	[m]	Maximal stable diameter
d_h	[m]	Orifice diameter in sieve tray
d_{trans}	[m]	Linear interpolation of the d between the d_{stab} and d_{100}
d_{100}	[m]	Characteristic drop diameter due to a breakage probability of 100%
D	[m]	Diameter of column
e_{pull}	[*]	Pull error metric
$\bar{e}_{pull}$	[*]	Mean of pull distribution
$\tilde{e}_{pull}$	[*]	Standard deviation of pull distribution
e_{rmse}	[*]	Root mean squared error
e_{R2}	[-]	Coefficient of determination
f	[1/s]	Frequency of pulsation
g	[m/s^2]	Gravitational constant
$\mathcal{GP}$	[-]	Gaussian process
p	[-]	Breakage probability
u	[*]	Predicted values
$\hat{u}$	[*]	Experimental value
We	[-]	Weber number
ε	[m^2/s^2]	Turbulent dissipation rate
η	[Pa·s]	Viscosity
ϑ	[-]	Number of daughter drops
π_{af}	[-]	Pulsation term in breakage model
ρ	[kg/m^3]	Density
σ	[N/m]	Interfacial tension
σ_e	[-]	Measurement uncertainty
φ	[-]	Free cross-sectional area of sieve tray

***Unit depends on property.**

Sub- and Superscripts

c	Continuous phase
col	Column
d	Dispersed phase
exp	Experimental
i, n	Count variables