

An objective classification of condensation regimes in direct contact condensation

Citation for published version (APA):

Safavi Nic, S. S., van Veen, T. H. P., Buist, K. A., Verdurmen, R. E. M., & Kuipers, J. A. M. (2023). An objective classification of condensation regimes in direct contact condensation. *AIChE Journal*, 69(9), Article e18121. <https://doi.org/10.1002/aic.18121>

Document license:
CC BY-NC

DOI:
[10.1002/aic.18121](https://doi.org/10.1002/aic.18121)

Document status and date:
Published: 01/09/2023

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:


openaccess@tue.nl

providing details and we will investigate your claim.

RESEARCH ARTICLE

Transport Phenomena and Fluid Mechanics

An objective classification of condensation regimes in direct contact condensation

Sherwin S. Safavi Nic^{1,2} | Tijmen H. P. van Veen¹ | Kay A. Buist¹  |
Ruud E. M. Verdurmen² | J. A. M. Kuipers¹

¹Multiphase Reactors Group, Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, Eindhoven, the Netherlands

²Danone Nutricia Research, Utrecht, the Netherlands

Correspondence

Kay A. Buist, Multiphase Reactors Group, Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, P.O. Box 513, MB 5600 Eindhoven, the Netherlands.
Email: k.a.buist@tue.nl

Abstract

Intensified heat treatment, using direct contact condensation (DCC), is applied in the production of dairy products to ensure a high level of food safety. The key challenge with DCC is the fouling due to the protein reactions that limits operational efficiency and sustainability. Using a condensation regime map can improve operational decision-making. Pilot plant scale experiments were conducted for a wide range of steam mass fluxes and inlet temperatures at high and low channel pressures. High-speed images were recorded and analyzed to obtain penetration lengths and plume area. The experimental data and image analysis supplemented with temperature and pressure measurement, were processed using machine learning (ML) to develop a data driven model to predict the regime maps. The linear discriminant analysis (LDA) was found to be the most suitable model. From the ML models it was also found that the best parameters to make a condensation regime map are the steam pressure, channel pressure, subcooling temperature, water Prandtl number, and the relative velocity ratio between gas and liquid. The condensation outcomes were presented with various two-dimensional regime maps. New regime maps are proposed using the Prandtl number and velocity ratio as dimensionless parameters.

KEYWORDS

direct steam condensation, linear discriminant analysis, regime map

1 | INTRODUCTION

The applications of direct contact steam condensation (DCC) vary from a passive coolant injection system for nuclear power plants¹ to sterilization of dairy products.² Even though the objective for nuclear power plants is rapid cooling of steam whereas for dairy products the focus is rapid heating to achieve sterilization, the principle of DCC remains the same. The steam condensates at the equilibrium temperature and the hot condensate is mixed with the cold liquid.² The benefit of using DCC for sterilization is that high

temperatures can be reached in short times and as a result lowers the degree of browning and nutrient losses that occur when applying an ultra-high temperature-treatment (UHT) compared to indirect methods.³ The challenge with DCC is that it can generate a hot surface of steam bubbles upon which milk proteins may aggregate.^{3,4} It is therefore important that the flow patterns in the injector are such that the residence times are as short as possible and that hot condensate mixes rapidly with the cold milk. Mapping the behavior of steam condensation into a condensation regime map can improve operational decision-making.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *AIChE Journal* published by Wiley Periodicals LLC on behalf of American Institute of Chemical Engineers.

A regime map provides an overview of the type of condensation behavior that can be observed under certain operational conditions. The most relevant operational regimes can be found in Table 1. In addition to the mentioned regimes some researchers also identified several. However, for many of these subcategories it is not clear at what exact conditions these experiments were performed.^{8,9} This was also reported by Heinze et al.¹⁰ when using the experimental data from Xu et al.¹¹ to validate the model in their work. An overview of most relevant research is provided in Table 2. Parameters such as steam pressure and channel pressure are not reported precisely, usually only as an indicative range or not reported.^{5,6,12} These parameters are important because they influence the physical properties of steam and the saturation temperature of water. Clerx et al.¹³ reported the exact experimental conditions per run, however, the results were not used to generate reliable regime maps.

The goal of the present work is to investigate the critical parameters that determine the boundaries of condensation regimes of direct steam injection in industrial applications. A data driven model was used to predict the regime boundaries to elucidate the role of critical parameters. This study is divided into three parts: (1) producing a condensation regime map using a pilot-plant scale setup with industrial conditions; (2) determining which parameters are most influential in controlling the condensation regime and finally; and (3) developing a data driven model to predict the regime map.

2 | MATERIALS AND METHODS

2.1 | Experimental set-up

The experimental setup is schematically represented in Figure 1. As can be seen from the diagram, the setup has a fully closed water cycle

TABLE 1 Description of the distinguished condensation regimes.^{5–7}

Chugging (C)	Steam–water interface fluctuates around the nozzle exit and moves in and out of the nozzle
Bubbling (B)	Bubble formation around the nozzle exit, cyclic growth, detachment, and collapse
Jetting (J)	Steam velocity was (super)sonic and the steam–water interface was stable and smooth

TABLE 2 Overview of various research on steam injection into stagnant pool and flowing water bodies.

Injector type	Orientation	Water temperature (°C)	Steam mass flux (kg m ⁻² s ⁻¹)	Channel pressure (bar)	Steam pressure (bar)	References
Stagnant pool	Vertical	60–90	0–50	atm	NR	6,12
Co-Current	Vertical	10–90	0–1500	NR	NR	5
Cross flow	Vertical	25–74	40–135	2.44–3.16	2.2–3.5	13
Co-current	Vertical	20–70	110–500	1.5	2–8	8,11
Co-current	Horizontal	20–60	200–650	1–5	1–5	9

Note: Experimental data are listed however not all values are reported (NR).

and a steam generation input. Water is stored in a 339 L cylindrical stainless steel tank (1) and is fed to the rectangular injection channel using a Optidrive E3 IP66/NEMA 4X pump with a capacity of 275–890 L/h. Steam is generated (3) using a CERTUSS E6-72M electrical steam generator with a capacity of 0–48 kg/h at 9 bar. A pressure reduce valve (4) is used to control the pressure and mass flow rate toward the injector. The latter is measured using a Proline Promass 80F Coriolis mass flow sensor. An ADCA FLT16 float and thermostatic steam trap (5) is introduced before injection into the channel to remove remaining condensate and inert gases to ensure a steady flow of saturated steam into the channel. The channel pressure is regulated using a back pressure control valve (6). The injector channel is illuminated by two LED lights (10) and are placed on both sides of the camera. The condensation phenomenon is recorded using a FASTCAM SA-Z 2100K-M-128GB high-speed camera at 20 kHz (11) with a Sigma 105 mm F/2.8 EX DG OS HSM Macro lens. The heated water is stored in a second 339 L stainless steel water buffer tank (7) before it is taken on by the heat-exchanger (9) to cool and returned to the feed tank. An overview of the experimental conditions is reported in Table 3.

The steam injector is a rectangular channel with two injection points, as shown schematically in Figure 2. The first segment of the injector is 415 mm long and is required for flow stabilization. The length was taken as 40 times the hydraulic diameter. The injection points are placed after the flow development segment and both have a diameter of 4 mm. The steam can be directed either through the top, bottom, or both injectors using valves. The mixing segment is optically accessible using Sapphire glass, placed on the front and back-side of the injector. The injector has in total 8 sensor slots placed on the top side (A) and 8 on the bottom (B). The first sensor after the injection point is placed 6.5 mm after the injection point. The distance to the other sensor slots is 19 mm with exception of a 27 mm distance between sensor slot 6, 7, and 8.

The sensor slots are suitable to place temperature and pressure sensors. Temperature is measured with type 12 mineral insulated T-type thermocouples. Pressure is measured with MAKS-6 (X) ultraminiature high temperature pressure transducers from Kulite. All sensor data are collected using a NI cDAQ-9185 data acquisition system (12) (with modules NI-9220, NI 9401 and NI 9213), which is transferred to a computer (13) and logged via LabVIEW-software. The recordings of the high-speed camera are processed via Photron FASTCAM Viewer 4.

FIGURE 1 Schematic representation of the experimental rig. (1) Water tank, (2) pump, (3) steam generator, (4) pressure reduce valve, (5) condensate drain, (6) back pressure valve, (7) hot water tank, (8) pump, (9) plate heat exchanger, (10) LED front lighting, (11) high speed camera, (12) data logger, and (13) computer.

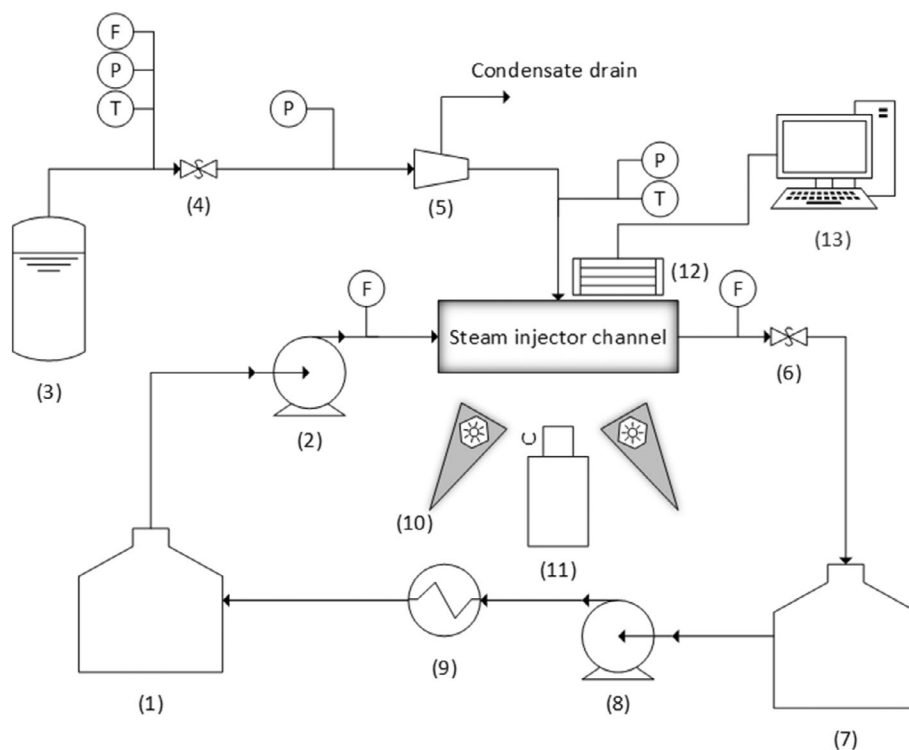


TABLE 3 Overview of the experimental conditions.

Variable	Value	Accuracy (%)	Unit
Liquid inlet temperature	15–85	0.4	°C
Channel pressure	1.2–4.2	1.1	bar
Steam pressure	1.2–8.2	0.5	bar
Steam mass flux	10–650	0.8	kg/(m ² s)
Subcooling	130–280	0.4	°C

2.2 | Image analysis

Objective classification methods for condensation regime determination are necessary to study the subtle transitions between regimes that cannot be determined visually and are subjected to operator biases. To this end, automated processing of the recording is preferred to analyze and compute relevant condensation characteristics such as the steam plume length. In this work, the Image Processing Toolbox in MATLAB R2020b is used to analyze the images and determine properties such as the plume length, area, and oscillation characteristics.

The MATLAB scripts loads the original image (Figure 3A) and proceeds, in the case of jetting and bubbling with binarization (Figure 3B) of this image, depending on a manually fixed threshold value that is applied to the complete image. The use of a manually fixed threshold can be sensitive to a bias affecting the objective regime classification. For this reason in section 2.3 the sensitivity of this parameter is assessed. As a result the image is processed such that only the steam bodies of interest remain. This processing involves two steps; filling

and clearing. First, holes in the image are filled. After the filling step, small objects that are not of interest are removed resulting in Figure 3C. When these steps have been performed, the resulting binary image shows only the steam bodies of interest. These steam bodies are then labeled to allow for individual analysis. Individually connected areas where a steam body is present are now labeled with ascending integers such that the first body is labeled with 1's in the matrix. Finally, the properties of these regions, such as area, centroid, and extrema of the steam body are determined. An example can be found in Figure 3D.

For the chugging regime, some additional processing steps are required before the binarization of the image. Due to weak internal reflection of light (see Figure 4A), some steam pockets appear darker than the background. Therefore, these steam pockets are rejected in the binarization due to the global thresholding. The visibility of steam bodies can be improved by using filters and contrasting to enhance edges of the pockets, and subsequently filling these pockets. Accordingly, a new step is added to the existing procedure: texture analysis, which is the characterization of various regions by texture in an image. This is done by quantifying texture qualities (e.g., rough/smooth) as a function of spatial variation in pixel intensities.¹⁴ The images are processed using a local standard deviation filtering which returns an image that has equal dimensions as its original, but each pixel entry is now the result of the standard deviation of a defined neighborhood, 3-by-3 in this work. On edges in the original image, the local standard deviation is high due to high local variation in pixel intensity. An example is provided by Figure 4B. After rescaling and adjusting the image, the pixels on the edges have now a high intensity, allowing the edge to pass the binarization process as shown by Figure 4C. The hole

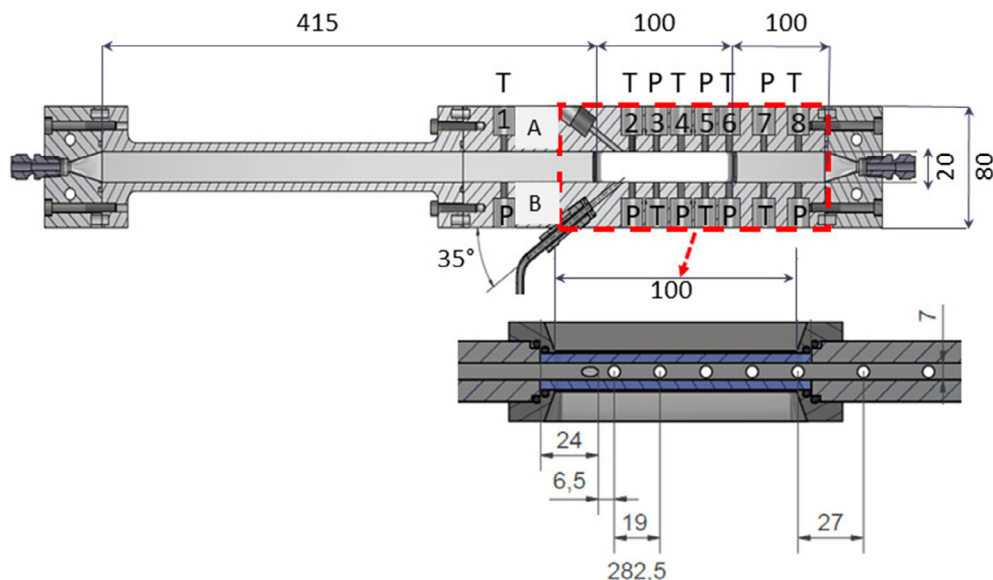


FIGURE 2 Schematic representation of the steam injector. $L = 615$ mm, $H = 20$ mm, and $W = 7$ mm. The channel has a 100 mm long sapphire glass window placed front and back for optical access of the injector.

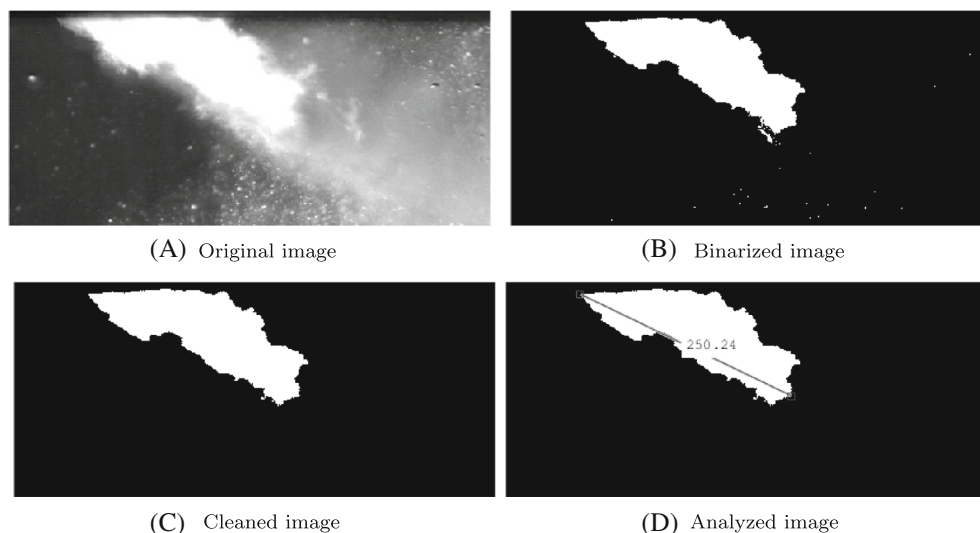


FIGURE 3 Steps of image preprocessing for bubbling and jetting regimes.

filling function is especially helpful in the chugging regime where the absolute pixel intensity of the regions cells are relatively low, but the variation between the bubble and the background is relatively high.

2.3 | Thresholding verification

In order to apply image binarization it is necessary to determine a threshold value. This process is susceptible to bias, as employing a low threshold results in loosened binarization where the steam plume characteristics could be artificially increased. For each experiment the appropriate threshold is selected enabling the correct determination of the plume characteristics.

Figure 5 shows a sensitivity plot for the penetration length for various thresholds. Applying a low value (i.e., a wide threshold) captures more gray pixels and increases the binarization area. A high value (i.e., a narrow threshold) results in a lower penetration length. The penetration length decreases linearly by increasing the threshold.

After a specific threshold, all pixels are converted to black and no steam plume can be detected, which is indicated by the abrupt ending of the graph. The standard deviation however flattens out and reaches a minimum. This implies that the overall variance of the measurement is minimized at the minimum of the graph. At this threshold value, the binarization is expected to resemble the original image closest and should therefore be selected. This method is employed for each digital image analysis to ensure consistency among the experiments and obtain the best possible accuracy in determining the penetration length.

3 | MACHINE LEARNING MODELS

The goal of this work is to investigate the critical parameters that determine the boundaries of the condensation regimes. In our study, 299 experiments were conducted with 79 different variables being logged whereas 2000 frames were recorded and saved per

FIGURE 4 Steps of image preprocessing for chugging regime.

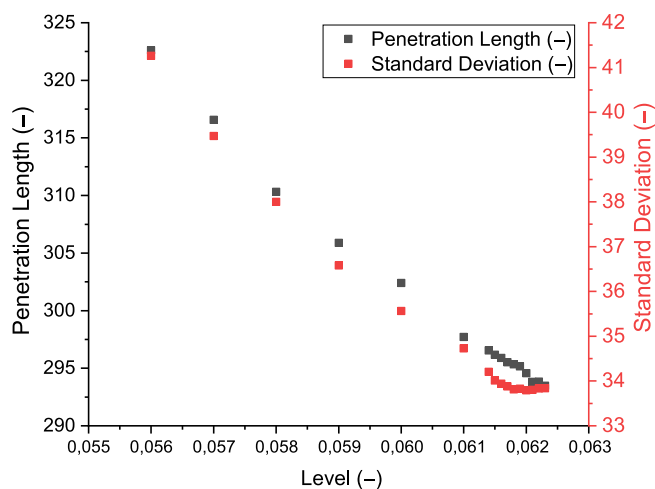
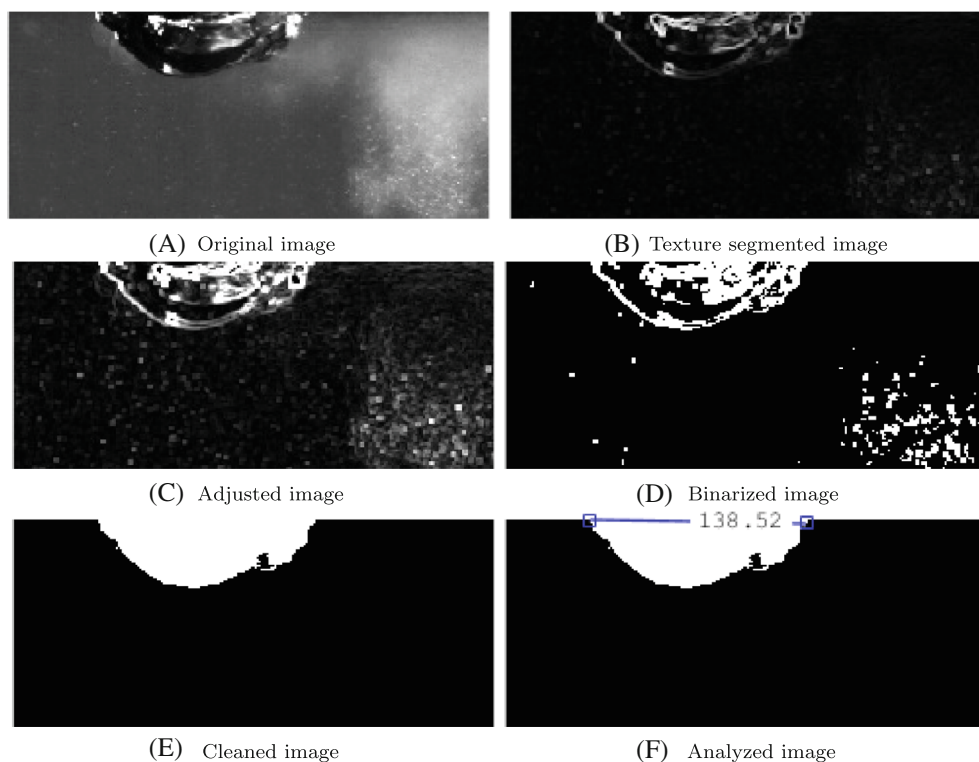


FIGURE 5 Sensitivity of the penetration length and standard deviation for a range of thresholds levels from wide to narrow.

experiment. These data can be used to find correlations between various system parameters and the condensation process. Machine learning (ML) is implemented to facilitate data processing and utilize algorithms to build a model that can classify samples based on the input data.

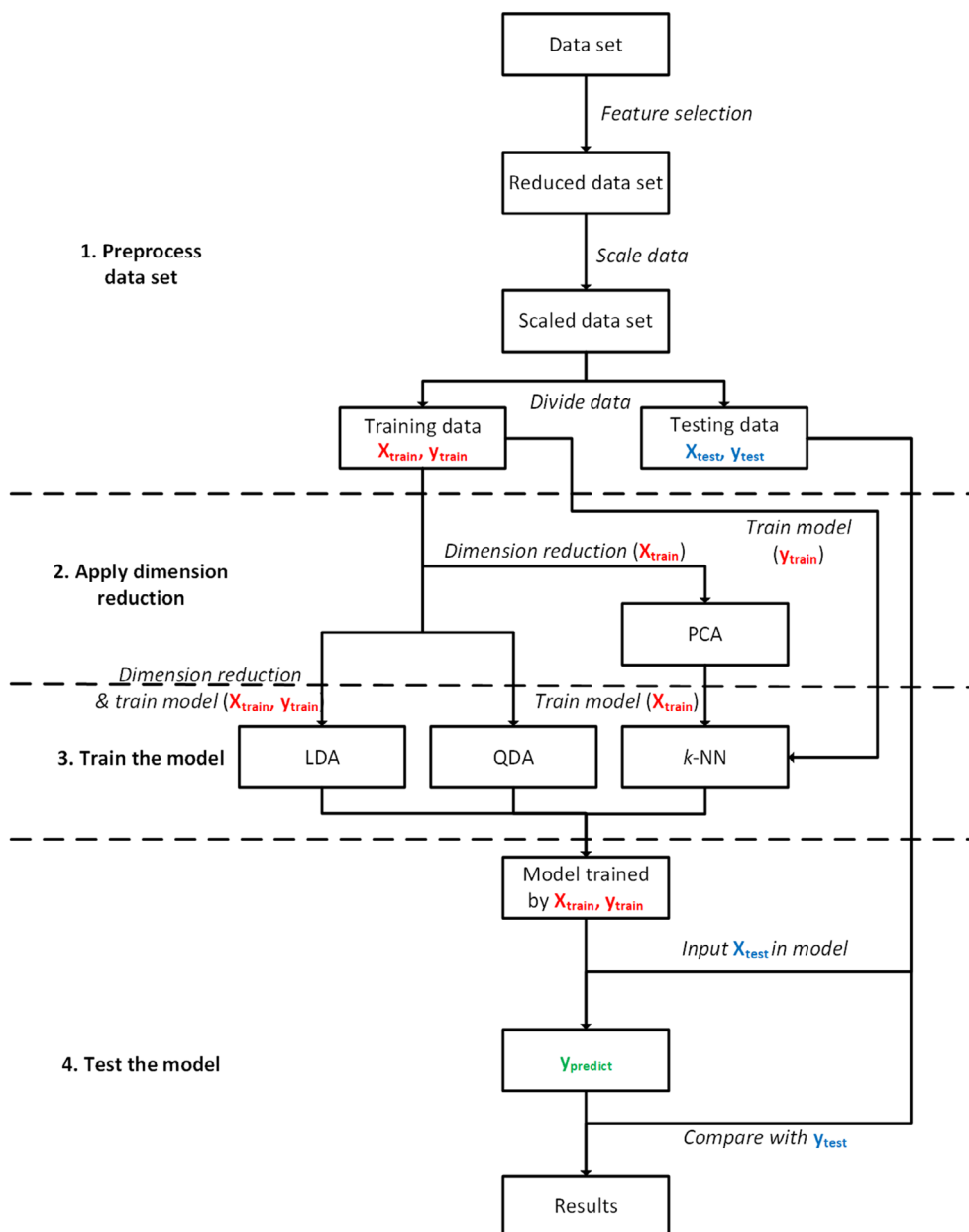
The machine learning methods used in this work focus on statistical classification and used to assign a specific class (regime) to experimental samples using existing data on explanatory variables termed features. In this work, three classification models will be implemented and compared; the k -NN algorithm preceded by principal component analysis (PCA), linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA).

The general implementation structure is visualized in Figure 6. The first step in this approach is data preprocessing (1). In this step a feature selection method is applied to identify the situation where two or more parameters are highly linearly related, (multi)collinearity. Additionally, in this step a separation is made in which data are divided into a training and testing set using a 80/20 ratio.

The next step is to reduce the dimensionality (2) of the data set. There are several benefits of this step. First, dimensionality reduction vastly reduces computation time because computation in high-dimensional data are costly.¹⁵ Second, it can remove random noise from the data and finally it removes unwanted degrees of freedom and thereby reduces the impact of properties that do not hold much variance and therefore reduce predictive power of the algorithm.^{16,17} Feature selection as dimension reduction is applied on the data set for all models.

Additionally, prior to application of the k -NN model, PCA is implemented, which is one of the most commonly applied dimension reduction methods in general and for k -NN specifically.^{18,19} PCA is an unsupervised model that applies a statistical procedure that computes linearly uncorrelated variables (principal components [PCs]) as new features to replace the existing features. The main idea of PCA is to simplify a high-dimensional data set by computing and selecting the two PCs as new variables that retain maximum information from the dimension reduction. The PCs are a linear combination of all features in the data set with weight coefficients termed loadings, and are all orthogonal to each other. The PCs have directions that represent the spread of the data (variance) and have magnitudes that represent how much variance is captured by the PC. Ideally, all variance is retained by the first two PCs to ensure no information is lost in the dimension reduction to two variables. The discriminants in LDA and QDA are

FIGURE 6 Visual flow diagram of the machine learning implementation procedure.



similar in form, but while PCA focuses the PCs to capture maximum variance in the data set, LDA and QDA compute the discriminants to have maximum separability between classes.

After the dimension reduction is applied the ML models can be trained (3). The model performance is analyzed for all three classification models using scores for precision, recall, and F1 for each class together with the amount of samples in the testing set of this class.

4 | RESULTS AND DISCUSSION

4.1 | Machine learning results

Before the models are used, feature selection was performed by inspection of Pearson and Spearman correlation matrices to reduce

multicollinearity and improve predictive power. From the 79 logged variables, 26 features were chosen to remain in the data set, which are shown in Table 4. In the models, the PCs and linear discriminants (LDs) are comprised of a linear combination of all these features.

After the feature selection the machine learning models are able to identify the key parameters for determining the condensation regimes. The main features are listed in Table 5. Three classification models were implemented and their performance is evaluated. The parameters for k-NN and QDA were not further investigated for visualizing regime maps due to the poor performance of these models as will be explained in section 4.3. In addition, a Buckingham Pi analysis was performed and the outcome of this methodology is also listed in Table 5.

The first step in evaluating the performance of the k-NN model is choosing the right method and the optimal value for k. There are two

TABLE 4 Features used in data set for models.

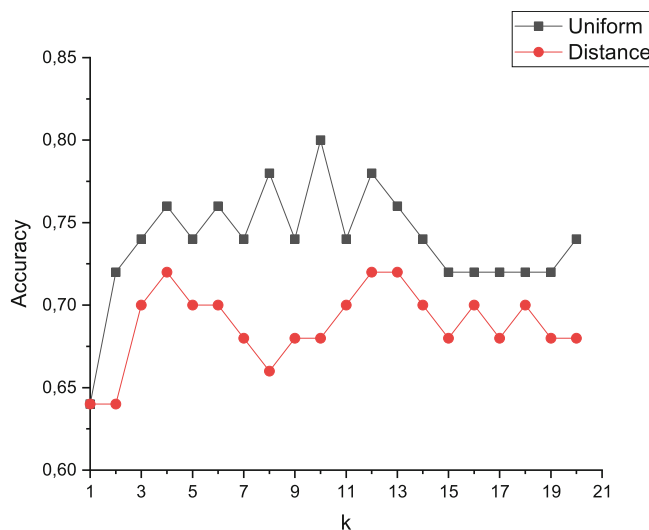
Parameter	Symbol	Unit
Liquid inlet temperature	$T_{L,in}$	°C
Subcooling	$\Delta T_{subcool} = T_{L,in} - T_{G,in}$	°C
Steam mass flux	G	kg/(m ² s)
Liquid inlet Reynolds number	$Re_{L,in}$	-
Steam inlet pressure	$P_{G,in}$	bar
Channel pressure	$P_{channel}$	bar
Channel pressure coefficient of variation	$CV_{P,channel} = \frac{\sigma_{P,channel}}{P_{channel}}$	-
Normalized penetration length	PL	-
Normalized steam plume area	A_G	-
Penetration length coefficient of variation	CV_{PL}	-
Steam plume area coefficient of variation	$CV_{A,G}$	-
Liquid temperature near the injector	$T_{L,inj}$	°C
Liquid outlet temperature	$T_{L,out}$	°C
Temperature increase	$\Delta T_{out-in} = T_{L,out} - T_{L,in}$	°C
Inlet liquid specific heat capacity	$c_{p,L,in}$	J/(kg K)
Outlet density	$\rho_{L,out}$	kg/m ³
Outlet viscosity	$\mu_{L,out}$	kg/(m s)
Outlet enthalpy	$h_{L,out}$	J
Outlet entropy of phase i	$S_{L,out}$	J/K
Outlet internal energy of phase i	$U_{L,out}$	J
Outlet specific heat capacity	$c_{p,L,out}$	J/(kg K)
Outlet speed of sound of liquid	$c_{L,out}$	m/s
Outlet thermal conductivity	$k_{L,out}$	J/(m K s)
Outlet surface tension	$\sigma_{L,out}$	J/m ²
Outlet liquid reduced pressure	$P_{R,out}$	-
Liquid outlet mass flow rate	$\dot{m}_{L,out}$	kg/s

TABLE 5 Overview of the identified defining features per origin.

Method	Identified features
k -NN with PCA	$\sigma_{out}, P_{G,in}, T_{out}$
LDA	$\mu_{out}, c_{p,out}, k_{out}, P_{G,in}, \Delta T_{subcooling}$
QDA	-
Buckingham-pi	$\frac{V_G}{V_L}$

Note: These features have received the largest coefficient from the ML models and therefore identified as most important.

main methods applied in this work for classifying a new sample; by majority of neighbors or by also taking their distance into account. Generally, large k 's are less affected by noise and result in smoother class boundaries.²⁰ Figure 7 shows the total accuracy of the model based on the method used and the amount of neighbors selected. As can be observed, the unweighted method yields better accuracy, with

**FIGURE 7** Accuracy of the k -NN model by method and amount of neighbors.**TABLE 6** Confusion matrix for the k -NN model.

	Predicted		
	Bubbling	Chugging	Jetting
Actual Bubbling	28	0	1
Actual Chugging	6	7	0
Actual Jetting	3	0	5

TABLE 7 Accuracy metrics scores for the k -NN model.

	Precision	Recall	F1-score	Support
Bubbling	0.76	0.97	0.85	29
Chugging	1.00	0.54	0.70	13
Jetting	0.83	0.62	0.71	8
Accuracy			0.80	50
Macro average	0.86	0.71	0.75	50
Weighted average	0.83	0.80	0.79	50

an optimum at $k = 10$. Generally, the accuracy is largely unaffected by selecting a different k , as the variation in k only yields an improvement of a few percent for each method. The general trend and variations are very common and in line with existing works in literature.^{20,21}

After optimizing the model, the k -NN algorithm preceded by PCA is still a mediocre model for predicting the regimes. Table 6 shows the confusion matrix for the model which indicates correct predictions for each class, leading to the accuracy report shown in Table 7. In this table, the macro average per score is the regular average of each score. The weighted average takes into account the support of each class and averages the value attributed to each score for each class accordingly.

As can be seen in the diagonal of the confusion matrix, only 40 out of 50 testing samples were correctly predicted, resulting in an accuracy of 80.0%. Overall, the accuracy of predicting the bubbling

TABLE 8 Confusion matrix for the LDA model.

		Predicted		
		Bubbling	Chugging	Jetting
Actual	Bubbling	29	0	0
	Chugging	2	11	0
	Jetting	0	0	8

TABLE 9 Accuracy metrics scores for the LDA model.

	Precision	Recall	F1-score	Support
Bubbling	0.94	1.00	0.97	29
Chugging	1.00	0.85	0.92	13
Jetting	1.00	1.00	1.00	8
Accuracy			0.96	50
Macro average	0.98	0.95	0.96	50
Weighted average	0.96	0.96	0.96	50

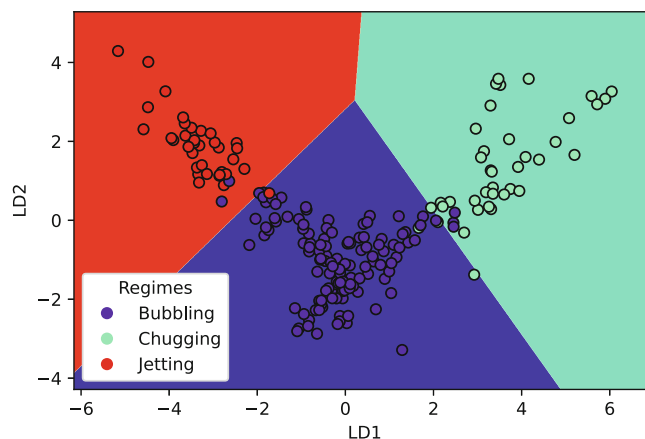
regime is good, as 28 of 29 samples are correctly classified as shown with the recall score of 0.97.

While the bubbling regime can be predicted well with this model, the chugging and jetting samples, however, are often mislabeled, mostly as bubbling. One explanation for this specific mislabeling could be the slightly unbalanced training set, as half of the samples belong to the bubbling regime. This may lead to a bias toward the bubbling regime. For the three classification models, this bias is especially problematic for the *k*-NN model, as it classifies a new sample based on the majority of votes. When one class is represented more than the others, it will become the majority more quickly. Introducing a more balanced data set could improve the performance of the *k*-NN model.

An explanation for the general mediocre performance of this model is the relatively low amount of variance captured by the PCs. PC1 accounts for 55.5% for the total variance, while PC2 captures 17.5% of variance. This means that 27% of the information from the variables is lost in the dimension reduction step of PCA.

4.2 | Linear discriminant analysis

LDA performs very well as a classification model for the data set. The confusion matrix and accuracy scores can be observed in Tables 8 and 9, respectively. As can be observed in the confusion matrix, only two samples are misclassified, leading to an accuracy of 96%. The jetting regime scores perfectly, as can also be seen in Table 9. The lowest score in this table is the recall of the chugging regime with 0.85, which is still a relatively good score. This implies that LDA is very well suited as a classification model. An explanation for the improved performance compared to the *k*-NN model can be found in the dimensionality reduction step. While PCA aims to achieve maximum variance between its axes, LDA tries to maximize separation between the classes. The explained variance of the LDs can also be compared to the

**FIGURE 8** Linear discriminant analysis decision boundaries.**TABLE 10** Confusion matrix for the QDA model.

		Predicted		
		Bubbling	Chugging	Jetting
Actual	Bubbling	28	1	0
	Chugging	4	9	0
	Jetting	2	0	6

TABLE 11 Accuracy metrics scores for the QDA model.

	Precision	Recall	F1-score	Support
Bubbling	0.82	0.97	0.89	29
Chugging	0.90	0.69	0.78	13
Jetting	1.00	0.75	0.86	8
Accuracy			0.86	50
Macro average	0.91	0.80	0.84	50
Weighted average	0.87	0.86	0.86	50

PCs. LD1 contains approximately 73.1% of total variance and LD2 captures 26.9%, compared to 55.5% and 17.5% of PC1 and PC2, respectively. Extremely little information is lost in the dimension reduction step of LDA compared to PCA, which contributes to the high accuracy of the LDA model. A visual of the high accuracy of the LDA model can be observed in Figure 8.

4.3 | Quadratic discriminant analysis

The last model, quadratic discriminant analysis, was applied to the data set and yields average results. Tables 10 and 11, respectively show the confusion matrix and the accuracy scores for the QDA model. As shown in the diagonal of the confusion matrix, the model is fairly good in predicting the classes of the testing set with 86% accuracy. The precision for predicting the bubbling regime is the lowest of the three with 82% (28 out of 34 correct predictions). This result is

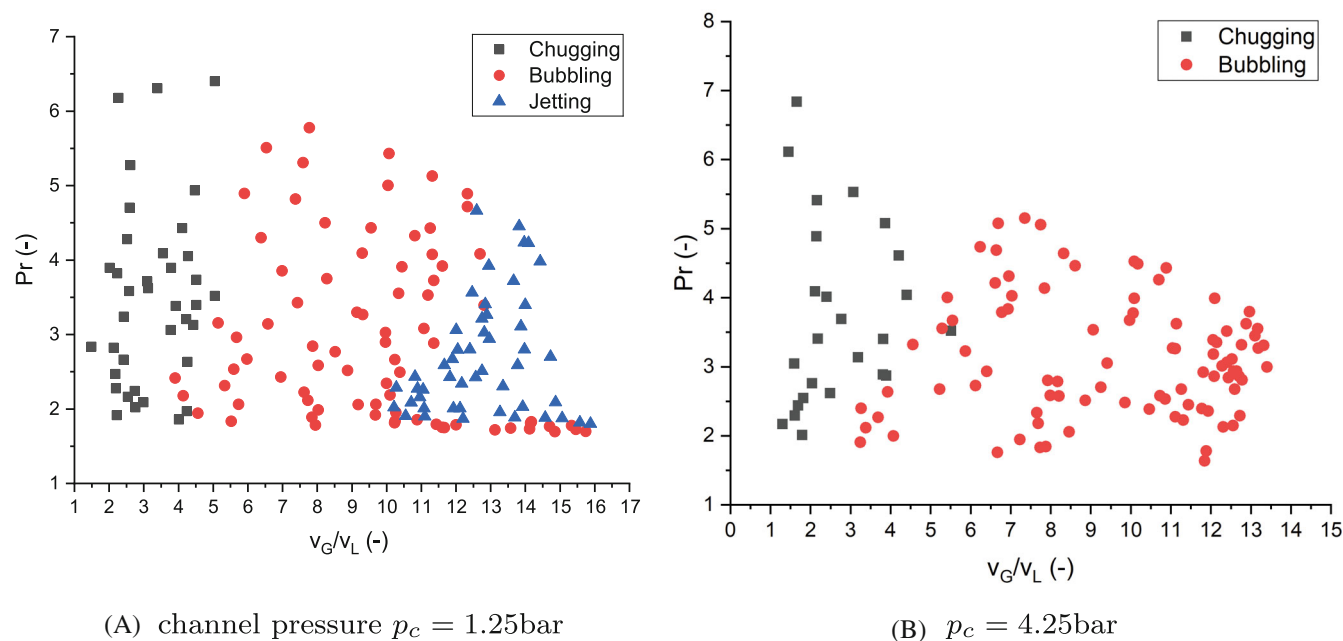


FIGURE 9 Regime map based on the Prandtl number and dimensionless velocity. $Re = 12,500$ at different back pressures.

similar to LDA, but quite different from the relatively high precision score for bubbling in the k -NN model, which was reasonably good at predicting the bubbling regime compared to the others. Another surprising detail can be found in the recall of the chugging regime with a score of 0.69. This means that the model does not perform well in predicting the chugging regime.

QDA performs worse than LDA, which can be explained in several ways. One of the main differences between LDA and QDA is the way they treat covariance matrices. LDA assumes that all classes share a common covariance matrix due to the assumption of a multivariate Gaussian distribution.²² QDA, however, assumes that each class has its own covariance matrix. This difference between the methods leads to a bias-variance trade-off. For QDA, a separate covariance matrix is calculated for every class, tripling the amount of parameter estimations in this work. The LDA model becomes linear in x (amount of measurements) by assuming every class shares the same covariance matrix. Therefore, LDA becomes a less flexible classifier and thus has considerably lower variance, which can lead to improved prediction performance.²³ The trade-off is that if the assumption of a shared covariance matrix is off, the LDA model may suffer from high bias as it is less flexible. LDA is typically better than QDA with small data sets and so reducing variance is critical.²⁴

4.4 | Regime maps

The key outputs of the machine learning models, next to the predictive power, are the key process variables that influence these regimes. The added value of this approach is that it allows us to redefine the regime maps using objectively chosen process variables. The variables of interest are the steam pressure (P_s),

subcooling ($T_{sc} = T_s - T_w$), and the water Prandtl number (Equation 1) at the exit of the injector and the dimensionless velocity ratio between gas and liquid (Equation 2).

$$Pr = \frac{c_p \mu}{k} = \frac{\text{momentum diffusivity}}{\text{thermal diffusivity}} \quad (1)$$

$$\Pi_v = \frac{v_G}{v_L} \quad (2)$$

In this research, it has been chosen to work as much as possible with dimensionless numbers to depict the regime maps as it makes it less dependent on our specific experimental set-up and operating conditions. The proposed regime map is a combination of the Prandtl number and the velocity ratio of gas and liquid. The Prandtl number was identified by the LDA analysis as a critical parameter for defining regimes. Figure 9 shows the regime map for two different back pressure settings and shows good separation between the regimes. In Figure 9A, all three regimes can be discovered. For higher water inlet temperatures (and as a consequence, lower Prandtl numbers) the regime switched again from jetting to bubbling. When the back pressure was increased the jetting regime was no longer observed as shown in Figure 9B. At higher back pressure, the steam pressure also increased. There are then several combined physical effects that occur. Firstly at elevated channel pressure, the saturation temperature is increased. At the same time a higher steam pressure is required, resulting in a higher degree of subcooling. These effects, combined in Equation (3), increase the rate of condensation and as a consequence the steam cavity penetrates less into the channel. Finally, due to the elevated pressure the gas bubble size distribution decreases, resulting in a higher interfacial area as shown in Equation (4).

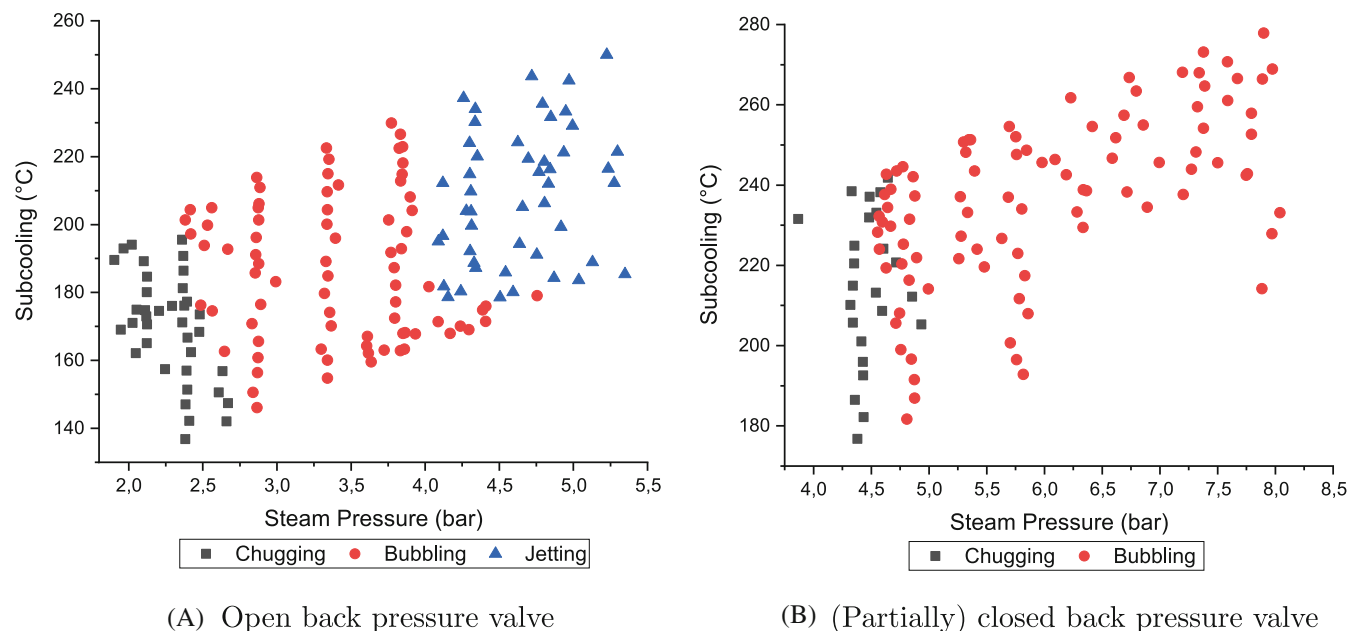


FIGURE 10 Regime map based on the steam pressure and degree of subcooling. $Re = 12,500$ at different back pressures.

$$m_{g \rightarrow l} = a_{if} \frac{\alpha_l (T_{sat} - T_l) + \alpha_g (T_{sat} - T_g)}{h_{g,if} - h_{l,if}} \quad (3)$$

$$a_{if} = \frac{6\epsilon_d}{d_{32}} \quad (4)$$

The influence of the channel pressure is also found in the interphase heat transfer coefficient (α_l). This quantity can be estimated using a Nusselt correlation as given by Equations (5) and (6). The impact of the channel pressure is found in the Reynolds number (Equation 7) of the dispersed gas phase. The Reynolds number is not only influenced by the gas bubble size (d_{32}), but also by the relative vapor velocity which is influenced by the density of steam.

$$Nu_l = \frac{\alpha_l \cdot d_b}{k} \quad (5)$$

$$Nu_l = 2 + 0.4 Re_g^{1/2} Pr_l^{1/3} \quad (6)$$

$$Re_g = \frac{\rho_l |v_g - v_l| d_{32}}{\mu_l} \quad (7)$$

The main drawback of working with the Prandtl number and dimensionless velocity is that these parameters are less intuitive for the use of regime maps. Another set of parameters identified by the machine learning are the steam pressure and degree of subcooling, Figure 10. The benefit of this regime map is that the variables used are intuitive for operational decision making as well as that it shows clear distinction between regimes.

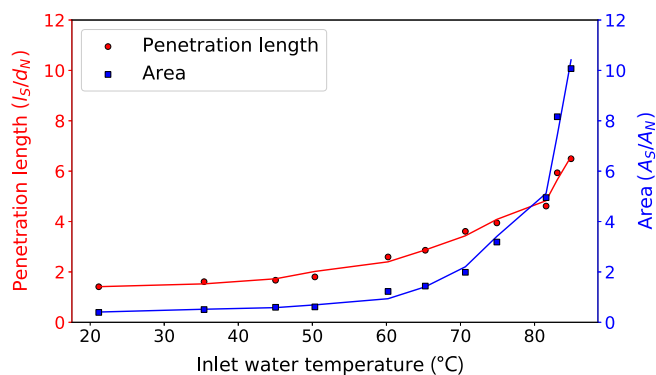


FIGURE 11 Penetration length and area as a function of liquid inlet temperature at a steam mass flux of $315 \text{ kg}/(\text{m}^2 \text{ s})$, channel pressure = 1.7 bar, steam pressure = 4.2 bar, and $Re = 12,500$.

4.5 | Impact of elevated liquid temperature

As discussed in section 4.4, when the Prandtl number is decreased the jetting regime is no longer observed. This is a direct consequence of increasing the liquid inlet temperature. When the liquid temperature is increased, the driving force for condensation decreases. This allows the steam cavity to penetrate deeper into the channel and occupy a larger area, as shown in Figure 11. Both the penetration length (l_s) and the area (A_s) are normalized by dividing the values by the injector diameter (d_N) and its area, respectively (A_N). As the steam cavity penetrates deeper and the surface area increases, more condensation can occur, in turn reducing the steam cavity size. This effect yields the bubbling regime instead of the jetting regime and is consistent with observations in literature.²⁵

5 | CONCLUSION AND FUTURE WORK

The objective of this work was to experimentally investigate which parameters play a role in the direct steam condensation process. Pilot plant scale experiments were conducted for a wide range of steam mass fluxes and inlet temperatures at high and low channel pressures. The results were processed using an image analysis protocol to obtain relevant steam cavity properties and further processed with machine learning models to provide objective classification and regime prediction possibilities.

The experimental campaign yielded 299 experiments with 79 different logged variables and 2000 frames recorded per experiment that were used to generate the regime maps. These large amount of data have not been observed in previous research. It was also found that it is important to properly note, besides steam pressure, the channel pressure as it has a significant effect on the condensation regime boundaries. The inlet temperature effect on condensation regime is small, but it does have a large effect on the steam cavity penetration length.

Regarding machine learning, the k -NN and QDA classification models perform reasonably well. Moreover, it can be concluded that LDA is the superior classification model with 96% accuracy, being able to both classify regimes in an objective fashion and predict the regime for new samples. A parameter study was performed via the inspection of the LDs, which led to the insight that the outlet Prandtl number is a good predictor for condensation regimes. Besides the Prandtl number, the steam pressure, channel pressure, subcooling, and velocity ratio are better variables for regime maps than steam mass flux and inlet temperature, respectively. It is recommended to use steam pressure and subcooling as key variables for regime maps and to report the channel pressure.

For future work it is recommended to firstly increase the liquid viscosity to match the profile of high protein dairy products. It can be expected that the viscosity of the liquid influences the regime map via the Prandtl number. It is important to map the regimes using representative viscosity profiles which will be subject to further research.

NOTATION

c_p	specific heat capacity (J/kgK)
μ	dynamic viscosity (Pa·s)
k	thermal conductivity (W/mK)
v_i	velocity of phase “ i ” (m/s)
T_{sat}	saturation temperature (K)
T_i	temperature of phase “ i ” (K)
h_i	enthalpy of phase “ i ” (J/kg)
d_{32}	gas bubble diameter (m)
a_{if}	interface area bubble–liquid (m ^{−1})
$m_{g \rightarrow l}$	condensation volumetric rate (kg/m ³ s)
α_i	heat transfer coefficient phase “ i ” (W/m ² K)
ε_g	volume fraction of gas (–)

AUTHOR CONTRIBUTIONS

Sherwin S. Safavi Nic: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – original draft (equal).

Tijmen H. P. van Veen: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); resources (equal); software (equal); validation (equal). **Kay A. Buist:** Conceptualization (equal); data curation (equal); supervision (equal); writing – review and editing (equal). **Ruud E. M. Verdurmen:** Conceptualization (equal); project administration (equal); supervision (equal). **J. A. M. Kuipers:** Conceptualization (equal); supervision (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

The authors want to acknowledge the support from the pilot plant of Danone Nutricia Research in the design of the experimental set-up.

DATA AVAILABILITY STATEMENT

The full set of experimental data as well as the scripts to analyse the data are available as an integral data package. The data package can be accessed via this DOI from 4TU.Research Data: <https://doi.org/10.4121/21541776.v1>

ORCID

Kay A. Buist  <https://orcid.org/0000-0003-1765-576X>

REFERENCES

- Takeya Y, Miwa S, Hibiki T, Mori M. Application of steam injector to improved safety of light water reactors. *Prog Nucl Energy*. 2015;78:80-100. <http://www.sciencedirect.com/science/article/pii/S0149197014002303>
- Innings F, Hamberg L. Steam condensation dynamics in annular gap and multi-hole steam injectors. *Procedia Food Sci*. 2011;1:1278-1284. <http://linkinghub.elsevier.com/retrieve/pii/S2211601X1101908>
- Van Boekel M. Effect of heating on Maillard reactions in milk. *Food Chem*. 1998;62(4):403-414.
- Malmgren B, Ardö Y, Langton M, et al. Changes in proteins, physical stability and structure in directly heated UHT milk during storage at different temperatures. *Int Dairy J*. 2017;71:60-75.
- Petrovic de With A, Calay RK, de With G. Three-dimensional condensation regime diagram for direct contact condensation of steam injected into water. *Int J Heat Mass Transf*. 2007;50(9-10):1762-1770.
- Liang KS. *Experimental and Analytical Study of Direct Contact Condensation of Steam in Water*. PhD thesis. Massachusetts Institute of Technology; 1991. <https://dspace.mit.edu/handle/1721.1/13298>
- Zhao Q, Hibiki T. Review: condensation regime maps of steam submerged jet condensation. *Prog Nucl Energy*. 2018;107:31-47.
- Xu Q, Guo L. Direct contact condensation of steam jet in crossflow of water in a vertical pipe. Experimental investigation on condensation regime diagram and jet penetration length. *Int J Heat Mass Transf*. 2016;94:528-538. <http://www.sciencedirect.com/science/article/pii/S0017931015001994>
- Zong X, Liu JP, Yang XP, Yan JJ. Experimental study on the direct contact condensation of steam jet in subcooled water flow in a rectangular mix chamber. *Int J Heat Mass Transf*. 2015;80:448-457.
- Heinze D, Schulenberg T, Behnke L. A physically based, one-dimensional three-fluid model for direct contact condensation of steam jets in flowing water. *Int J Heat Mass Transf*. 2017;106:1041-1051. doi:10.1016/j.ijheatmasstransfer.2016.10.076
- Xu Q, Guo L, Zou S, Chen J, Zhang X. Experimental study on direct contact condensation of stable steam jet in water flow in a vertical pipe. *Int J Heat Mass Transf*. 2013;66:808-817.

12. Liang KS, Griffith P. Experimental and analytical study of direct contact condensation of steam in water. *Nucl Eng Des.* 1994;147(3): 425-435.
13. Clerx N. *Experimental Study of Direct Contact Condensation of Steam in Turbulent Duct Flow.* PhD Thesis, Eindhoven University of Technology; 2010.
14. Karabağ C, Verhoeven J, Rachel Miller N, Reyes-Aldasoro CC. Texture segmentation: an objective comparison between five traditional algorithms and a deep-learning U-net architecture. *Applied Sciences.* 2019;9(18):3900 <https://www.mdpi.com/2076-3417/9/18/3900/html> <https://www.mdpi.com/2076-3417/9/18/3900>
15. Van Der Maaten L, Postma E, Van Den Herik J. Tilburg Centre for creative computing dimensionality reduction: a comparative review dimensionality reduction: a comparative review. Tilburg centre for Creative Computing, Tilburg University, Tech. 2009. <http://www.uvt.nl/ticc>
16. Wang W, Carreira-Perpiñán MÁ. The role of dimensionality reduction in linear classification. 2014. <https://arxiv.org/abs/1405.6444v1>
17. Giessen AV. Dimension Reduction Methods for Classification. Ph.D. thesis, Delft University of Technology. 2012. http://repository.tudelft.nl/assets/uuid:89bf4948-87b4-4a9a-853b-8b828ae7ed24/Thesis_AM_250312.pdf
18. Yin CX, Peng QK. A careful assessment of recommendation algorithms related to dimension reduction techniques. *Knowl Based Syst.* 2012;27:407-423.
19. Deegalla S, Boström H. Classification of microarrays with kNN: comparison of dimensionality reduction methods. In: Yin H, Tino P, Corchado E, Byrne W, Yao X, eds. *Intelligent Data Engineering and Automated Learning—IDEAL 2007.* IDEAL 2007. (Lecture Notes in Computer Science). Vol 4881. Springer; 2007:800-809. doi:[10.1007/978-3-540-77226-2_80](https://doi.org/10.1007/978-3-540-77226-2_80)
20. Starzacher A, Rinner B. Evaluating KNN, LDA and QDA classification for embedded online feature fusion. ISSNIP 2008—Proceedings of the 2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing; 2008: 85-90.
21. Gupta SC, Goel N. Performance enhancement of diabetes prediction by finding optimum K for KNN classifier with feature selection method. Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT; 2020: 980-986.
22. Härdle WK, Simar L. *Applied multivariate statistical analysis.* 3rd ed. Springer; 2013.
23. Schmidt P. Linear vs. quadratic discriminant analysis—comparison of algorithms. 2018. <https://thatdatatho.com/linear-vs-quadratic-discriminant-analysis/>
24. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning.* 2nd ed. Springer; 2021.
25. Kim HY, Bae YY, Song CH, Park JK, Choi SM. Experimental study on stable steam condensation in a quenching tank. *Int J Energy Res.* 2001;25(3):239-252.

How to cite this article: Safavi Nic SS, van Veen THP, Buist KA, Verdurmen REM, Kuipers JAM. An objective classification of condensation regimes in direct contact condensation. *AIChE J.* 2023;69(9):e18121. doi:[10.1002/aic.18121](https://doi.org/10.1002/aic.18121)