scientific reports

Check for updates

OPEN

Machine learning based multi-parameter droplet optimisation model study

Ting Li, Likun Lu✉, Qingtao Zeng & Kexin Liao

Continuous inkjet technology, as a key technology in the field of industrial printing, is favoured for its excellent printing speed, precision and versatility. In order to achieve the accurate generation of ideal droplets in continuous inkjet devices, this paper proposes a new parameter optimisation method, BO-GP, which combines the Bayesian optimisation algorithm with computer vision, and after 50 rounds of iterations, it can converge to the optimal values of the control parameters, and successfully constructs the Pareto frontier of the control parameters. In this paper, experiments were conducted on two different device droplet image datasets, a millimetre-scale inkjet device and a microfluidic device, respectively. Compared with the original BO in Loop method, the optimised minimum objective function value is reduced from 0.378 to 0.331 in the millimetre-scale device, and from 0.073 to 0.046 in the microfluidic device. Moreover, the Pareto solution of the 10 sets of predicted parameters output using the BO-GP method tends to be stable with fluctuations around 0.1, and it takes only 1 h to derive the control conditions for achieving high roundness, high yield and uniform size droplets.

Keywords Bayesian optimisation, Objective function, Continuous inkjet printing, Multiparameter optimisation, Computer vision

Inkjet printing is the most widespread technological application of microfluidics, which is characterised by high droplet productivity, small size as well as extreme reproducibility. The ultimate goal of the inkjet printing process is to precisely control droplet generation[1]. There are two main methods for generating droplets, continuous inkjet (CIJ) printing and drop-on-demand (DOD) techniques[2,3]. DOD printing is widely used in desktop printing as it can eject smaller droplets and print detailed patterns, while CIJ can print at a much faster speed and is therefore mainly used in industrial printing, where traceability information is mainly printed onto packaging for food, beverages, pharmaceuticals, etc[4]. Both CIJ and DOD printing technologies face problems such as satellite droplet generation, too large droplet generation and too low droplet generation speed[5,6]. These problems reduce the stability and accuracy of inkjet printing techniques.

Accurate droplet generation from a CIJ device requires fine-tuning of the device control parameters to convert a continuously flowing liquid column into separated droplets. The two CIJ devices addressed in this paper are the millimetre-scale inkjet device and the microfluidic device. As shown in Fig. 1(a), the millimetre-scale inkjet device ejects ink from a nozzle under fluid pressure to form a continuous column of liquid. In order to produce uniform droplets at a specified frequency, a piezoelectric actuator is typically used to apply periodic oscillations to the liquid column, which splits to form droplets under surface tension due to Rayleigh instability[7–9]. A 'charging electrode' is fixed at the point where the column splits to form droplets, and an electric field is selectively applied to each droplet to control the amount of charging for each droplet. The droplets pass through a 'fixed electric field' which deflects the direction of flight of the droplets according to the amount of charge, and the deflected droplets fall onto the substrate, while the uncharged droplets are not deflected and are collected for later printing[4]. In the microfluidic device of Fig. 1(b), the aqueous-phase channel is responsible for transporting the aqueous-phase liquid and the oil-phase channel is responsible for transporting the oil-phase liquid, and the droplets are generated in the cross-channel at the intersection of the two channels. The core principle of droplet generation is based on capillary instability, the action of surface tension between the aqueous and oil phases, and the shearing and separation of droplets by precisely regulating the water and oil pressures[10,11].

Whether the liquid column separates to form droplets, the size of the generated droplets, and whether satellite droplets are generated during the process depends not only on the fluid pressure, piezoelectric drive frequency, nozzle translation speed, water pressure, and oil pressure, but also on the material properties of the ink, and all of

Beijing Key Laboratory of Signal and Information Processing for High-End Printing Equipment, Beijing Institute of Graphic Communication, Beijing 102600, China. ✉email: lklu@bigc.edu.cn

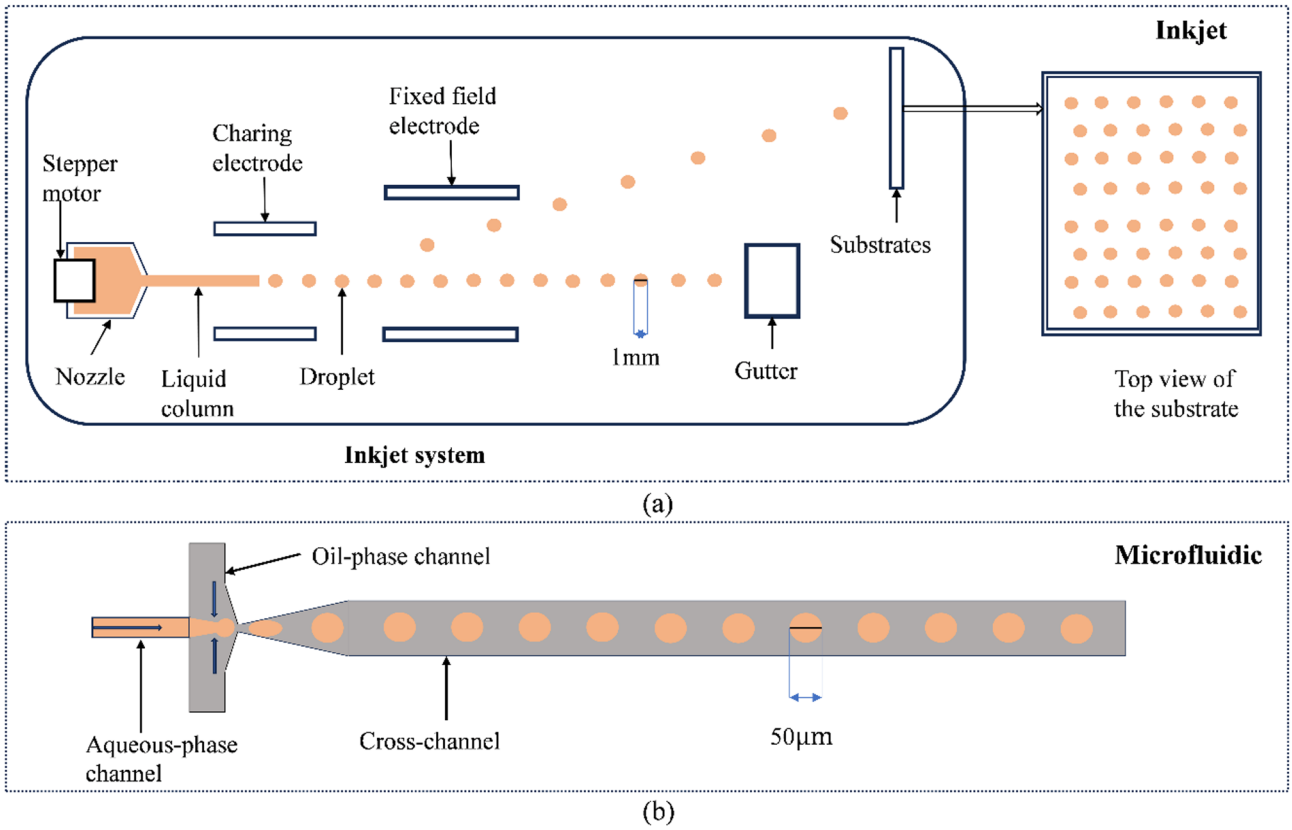

Fig. 1. Schematic diagram of droplet production for CIJ device. (**a**) Inkjet device droplet production schematic; (**b**) Microfluidic device droplet generation schematic.

these factors affect the accuracy of droplet formation. How to precisely tune the control parameters of the device to produce high roundness, high yield and uniform droplet size is the focus of this study.

In different CIJ devices, the distance from the nozzle to the substrate varies, and it is difficult to obtain an analytical relationship between the control parameters and droplet characteristics due to the complex nonlinear relationship between the forces acting on the fluid at different length scales. Typically, researchers conduct a large number of experiments to adjust experimental parameters to observe droplet morphology, develop proportional-integral-derivative controllers to predict the effect of parameter combinations, or perform hydrodynamic simulations of droplets to predict droplet behaviour[12,13]. Chu et al. proposed an automated monitoring and sorting system for droplet generating equipment, which reduces the labour cost, but when applied to different equipment it requires the real-time adjustment of process parameters[14]. Chu et al. proposed a deep learning based automated monitoring and detection system to classify droplet images for training, but this method uses a huge number of datasets and takes about 70 h of training time[12]. Rodriguez-Rivero et al. solved the problem of satellite droplet buildup on the print head deflector by quantifying the charge-to-mass ratio of the droplets, but did not eliminate satellite droplet generation at the root cause[15]. Solanki et al. proposed a neural network based method to predict the features of droplet generation, but the method relies on nearly 20,000 a priori training set of simulation results on a fluid dynamics simulator (CFD)[13]. Siemenn et al. proposed BO in Loop, a machine learning method that can quickly optimise droplets at different length scales, which requires only four iterations to generate the required droplet optimisation parameters, but the method needs to be validated by inputting the new predicted parameters generated in each iteration into the device[16].

Although existing droplet generation monitoring and optimisation methods can improve the automation level, they generally have problems such as poor equipment adaptability, high computational cost, high data demand or dependence on experimental validation. In order to achieve accurate droplet generation in the CIJ device, this paper proposes a droplet optimisation method GP-BO that combines Bayesian optimisation (BO) and computer vision. The computer vision module performs watershed segmentation on the input image to separate the droplet pixels from the background, and quantifies the separated droplet pixels in combination with an objective function, and the BO module iterates according to the size of the objective function until convergence, and finally outputs the minimum objective function value corresponding to the predicted parameter values. The GP-BO method is novel in that it does not require pre-processing or repeated experiments on the device, and can converge to the optimal control parameter values after 50 rounds of Bayesian optimisation with a small number of droplet image datasets. The method does not require a specific physical or mathematical model to perform the optimisation, so it is much more generalisable, and the experimental results were validated on droplet image datasets of two inkjet devices, a millimetre-scale inkjet device and a microfluidic device. The

average experimental time of the GP-BO method is 1 h, which is a saving of nearly 1.3 h of optimisation time compared to the BO in Loop method proposed by Siemenn et al. and the outputs of the control parameters converge to smaller intervals and the results are more stable.

Machine learning processes such as BO can greatly improve the efficiency of exploring the parameter space in a short period of time, and have significant advantages when dealing with complex, high-dimensional optimisation problems[17]. BO can efficiently deal with high-dimensional parameter space by constructing agent models and collection functions to efficiently explore the high-dimensional parameter space, reducing the number of objective function evaluations while balancing exploration and exploitation, avoiding the trapping of into a local optimum, and can be scaled up to multi-objective optimisation to construct a Pareto frontier and achieve multi-objective trade-offs[18,19].

Results

Experimental data set

In this paper, the experiments are validated on two droplet image datasets, millimetre-scale inkjet device and microfluidic device droplet image datasets, this dataset is derived from[16].

The inkjet device has three control parameters: fluid pressure, piezoelectric drive frequency, and nozzle translation speed, which affect the shape and yield of droplets on the substrate. The experimental dataset for the millimetre-scale inkjet device consists of 140 droplet images covering the typical cases of no droplets, large droplets, satellite droplets and ideal droplets.

The microfluidic device has two control parameters: water pressure and oil pressure. The experimental dataset of the microfluidic device consists of 140 images of droplets formed during contraction without contacting the contraction wall, droplets formed when contacting the contraction wall under two droplet generation mechanisms, and the case of no droplet generation[20–22].

Demonstration of GP model fitting effect

In BO, Gaussian process (GP) is used to fit a proxy model of the objective function. The predictive effects of the GP (predicted mean and variance) are used to construct the acquisition function EI. The EI determines the choice of the next sampling point, thus guiding the optimization process towards the globally optimal solution (Fig. 2).

In order to assess the effect of GP on the fitting of the objective function and to ensure the accuracy of the predicted objective function values, the objective function was fitted to the initial image, and the predicted objective function values derived from the GP model were compared with the true objective function values (Figs. 3 and 4). Two assessment indicators, namely mean square error (MSE) and coefficient of determination (R^2), were used to test the model fitting effect, and when the MSE was lower and the R^2 was higher, it indicated that the GP model was fitted well. The formulas are detailed in the supplementary Methods "Validation of model fit section" (Fig. 5).

The effect of GP model fitting for the droplet images of the inkjet device as well as the microfluidic device is shown in Fig. 6(D), where the horizontal coordinates are the true objective function values and the vertical coordinates are the predicted objective function values. It can be seen that the true objective function value and the predicted objective function value are equal, at this time the mean square error $MSE = 0.0$, the coefficient of determination $R^2 = 1.0$. It shows that the GP agent model fits the objective function perfectly and the predicted objective function values derived with this agent model are consistent with the true values.

Bayesian optimisation results

The BO optimization flowchart is shown in Fig. 6 in the "Methods" section, where the droplet image is subjected to watershed segmentation as well as BO. The BO module uses the GP as a proxy model to fit the objective function and updates the GP model according to the objective function, and then iteratively optimises to approximate the global optimal solution step by step. Each round of BO outputs two minimum objective function values and the corresponding prediction parameter values, and a total of 50 rounds of iterative optimisation are performed. 100 groups of prediction parameters and their corresponding objective function values are generated during the optimisation process, and the 10 groups with the smallest objective function values are selected as the Pareto front.

Figure 6(D) clearly shows the 100 sets of predicted parameters generated during the BO iteration process, with the inkjet device corresponding to the 3D scatterplot with fitted surface and the microfluidic device corresponding to the 2D scatterplot. During the iteration process, the EI function gradually selects points with smaller objective function values to optimise the model until convergence. The denser the distribution of points in the graph, the region where the optimal control parameter values are located.

The specific data of the 10 groups of Pareto frontiers with the smallest objective function values output by BO are shown in supplementary Table S1 and supplementary Table S2. The order of ranking is in the order of the objective function values, with the top group having the smallest objective function value representing the optimal solution of the optimisation.

Supplementary Table S1 shows the output of the Pareto front after BO of the inkjet device. EI is chosen as the acquisition function, and the inkjet device has three control parameters: fluid pressure, piezoelectric drive frequency, and nozzle translation speed. From supplementary Table S1, it can be observed that when the value of the objective function is small, the fluid pressure ranges from [0.000-0.750], the piezoelectric drive frequency ranges from [0.106–0.165], and the nozzle translation speed ranges from [0.964-1.000], which corresponds to the value of the objective function between [0.331–0.388]. In other words, the inkjet device is insensitive to changes in fluid pressure, and it is the ideal control condition when the piezoelectric drive frequency is low and the nozzle travelling speed is fast, at which time high roundness, high yield as well as droplets of uniform size

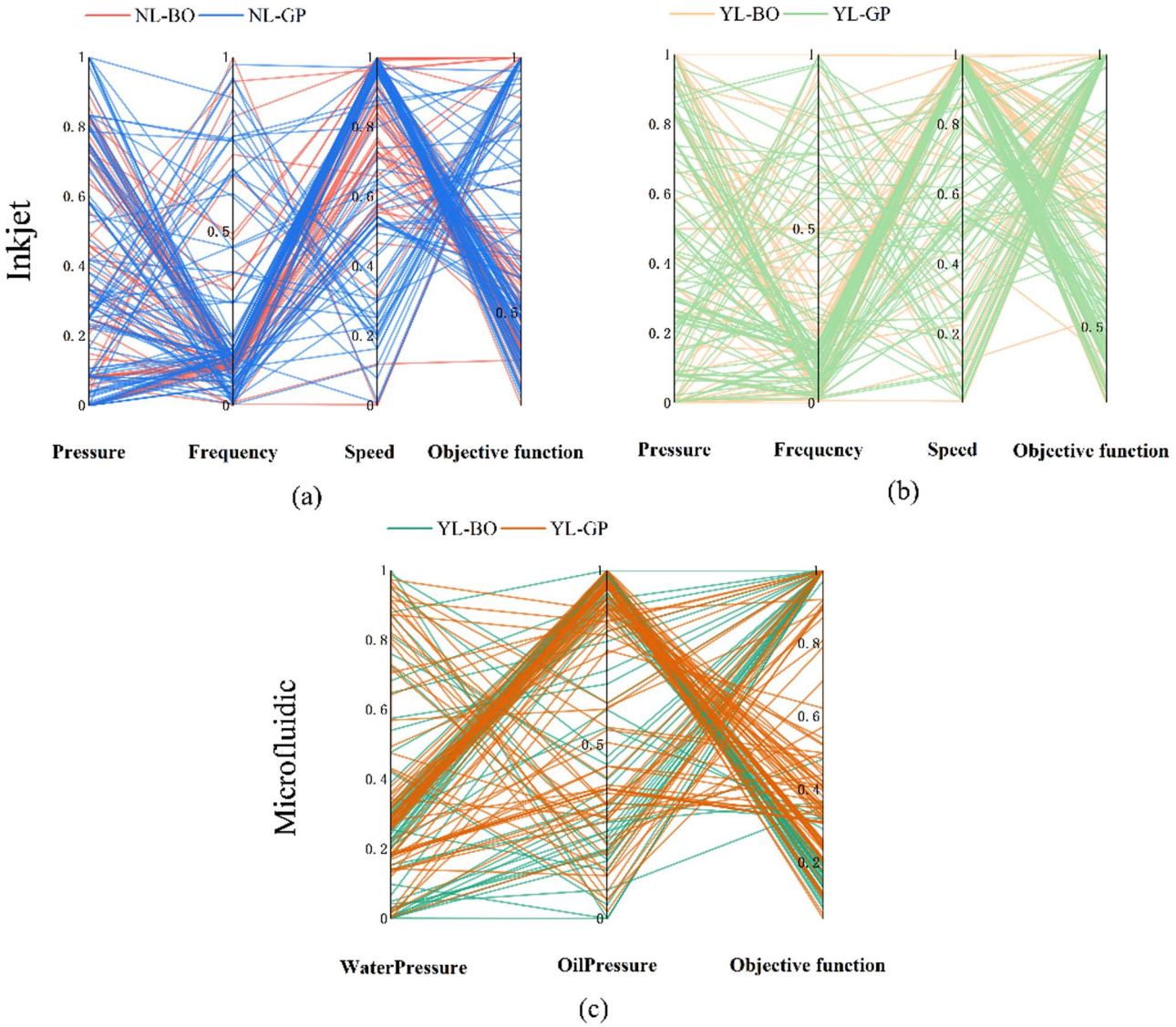

Fig. 2. Parallel coordinate plots of predicted parameter values as well as objective function values are generated for the CIJ devices using GP-BO and BO in Loop methods in the optimisation process, respectively. (**a**)-(**b**) Inkjet device generates parallel coordinate plots of predicted parameter values as well as objective function values using two optimisation methods under objective functions $l\,(x)$ and $l'\,(x)$, respectively; (**c**) Microfluidic device generates parallel coordinate plots of predicted parameter values as well as objective function values under $l'\,(x)$ using two optimisation methods.

can be formed. In supplementary Fig. S3, the insensitivity of the inkjet device to fluid pressure values is described in terms of the magnitude of the SHAP eigenvalues, and the main reason for this result is the CIJ device droplet generation mechanism. Fluid pressure provides the main energy input in CIJ to maintain the jet flow, but has less influence on the size and spacing of droplet generation, with piezoelectrically-driven perturbations and nozzle travelling speed dominating the dynamic generation of droplets.

Supplementary Table S2 shows the Pareto front output of the microfluidic device after BO. The microfluidic device has two control parameters: water pressure and oil pressure. The objective function of the microfluidic device consists only of the geometric loss L_{geom} and the yield loss L_{yield}. The objective function $l'\,(x)$ is the mean value of the geometric loss and the yield loss $l'\,(x) = 0.5L_{geom} + 0.5L_{yield}$. It can be observed from supplementary Table S2 that when the objective function value is small, the water pressure takes values in the range of [0.260–0.312] and the oil pressure takes values in the range of [0.951-1.000]. This shows that the microfluidic device is sensitive to both control parameters, water and oil pressure, and produces desirable droplets under low water and high oil pressure conditions.

Comparison test

In order to verify the accuracy of the output predicted parameters and objective function values of the GP-BO optimization method in this paper, the output results of the GP-BO method in this paper are compared with

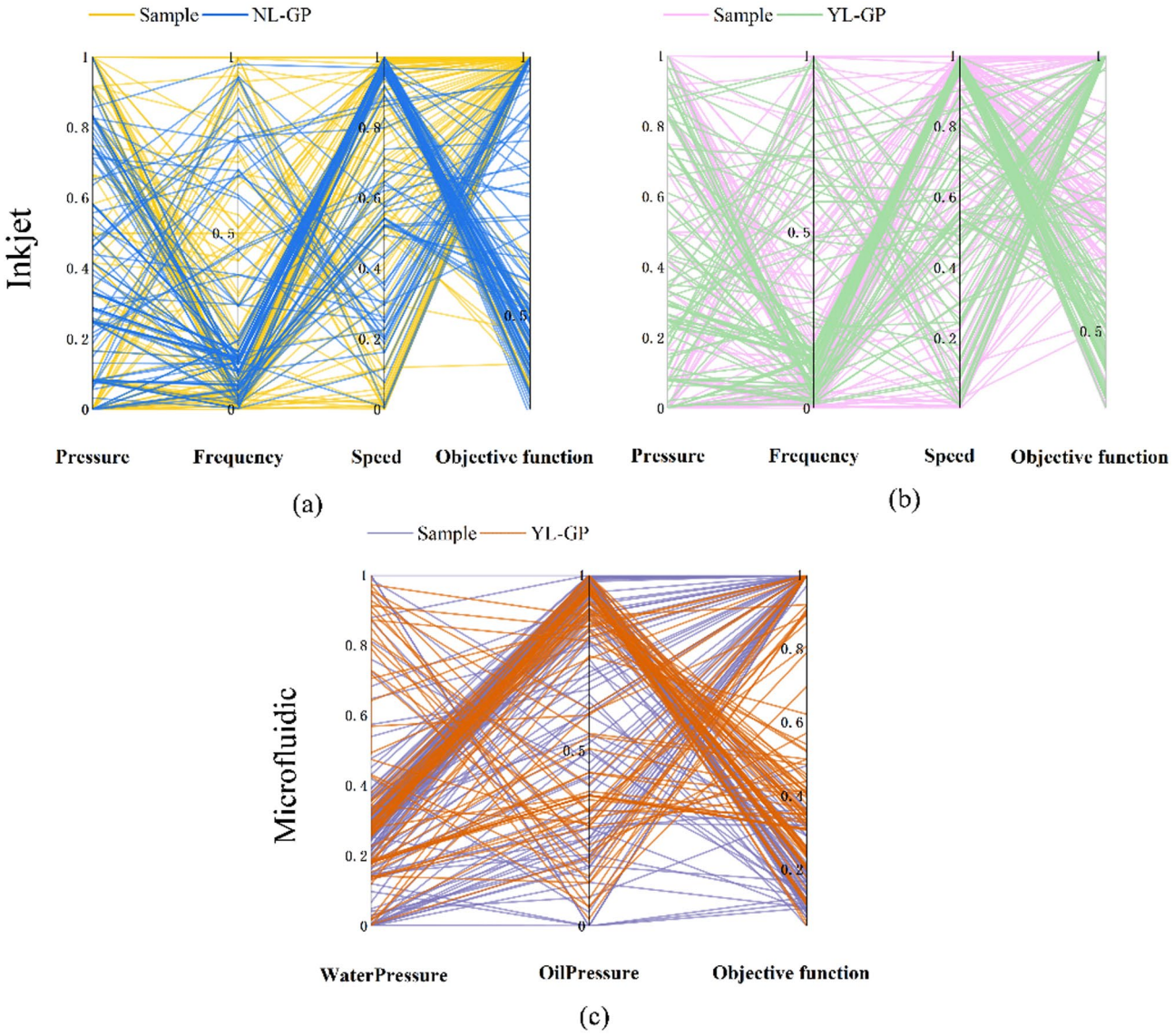

Fig. 3. Parallel coordinate plots of the CIJ devices based on the initial sample parameter values as well as the objective function values versus the optimised generation of predicted parameter values as well as the objective function values using the GP-BO method. (**a**,**b**) Parallel coordinate plots of the initial sample values of the inkjet device versus the generation of predicted parameter values as well as objective function values using the GP-BO method under objective functions $l(x)$ and $l'(x)$, respectively; (**c**) Parallel coordinate plots of the initial sample values of the microfluidic device versus the generation of predicted parameter values as well as objective function values using the GP-BO method under $l'(x)$.

those of the previous BO in Loop method, and the experimental results of the BO in Loop method are shown in the literature[16]. The BO in Loop method takes $l'(x)$ as the objective function to optimise the values of the device control parameters. In order to eliminate the influence of satellite droplets in the inkjet device, this paper improves on the basis of $l'(x)$ by adding the loss of dimensional uniformity to form the objective function $l(x)$, which is defined in Eq. (7) in the "Methods" section. Since the results before and after the change in the loss function cannot be compared directly due to the change in the loss function, the original loss function $l'(x)$ is used to verify the feasibility of the GP-BO method, which in turn generates new predicted parameter values under the new loss function $l(x)$.

Figure 2 shows the parallel coordinate plots of the predicted parameters and objective function values generated by the optimisation of the inkjet device and the microfluidic device using the GP-BO and BO in Loop methods, respectively. There are 160 sets of predicted parameters as well as objective function values in each parallel coordinate plot, 100 sets of data are generated during the GP-BO optimisation method iteration and 60 sets of data are involved in the BO in Loop method iteration.

Figure 2(a,b) show the parallel coordinate plots of the predicted parameters as well as the objective function values generated by the inkjet device using the objective function $l(x)$ and $l'(x)$ on the two optimisation

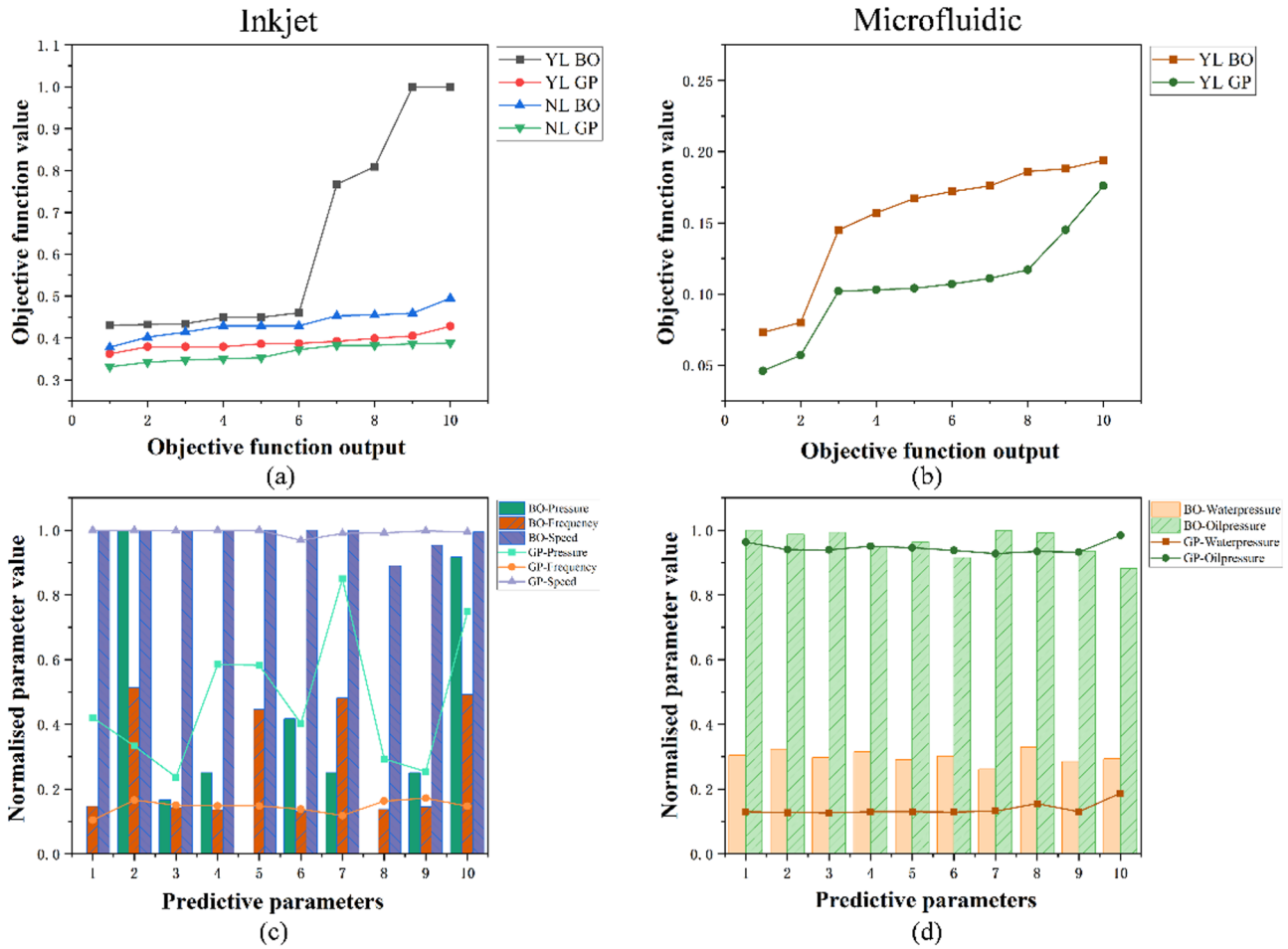

Fig. 4. Comparison plots of the 10 sets of minimum objective function values and corresponding predicted parameters output by the CIJ device using the GP-BO and BO in Loop methods. (**a**), (**b**) Comparison plots of the objective function values output by the inkjet device under objective functions $l(x)$ and $l'(x)$, and by the microfluidic device under $l'(x)$ using the two optimisation methods, respectively; (**c**)-(**d**) Comparison plots of the predicted parameter values output by the inkjet device and by the microfluidic device under $l'(x)$ using the two optimisation methods, respectively.

methods, respectively. The lines in the graphs are named, NL for using $l(x)$, YL for using $l'(x)$, GP and BO stand for the use of GP-BO versus BO in Loop optimisation methods, NL-GP for GP-BO optimisation of inputs using $l(x)$, and so on. It can be observed that the results obtained by either method are more or less the same, with the fluid pressure being distributed over the entire area, and a lower piezoelectric drive frequency along with a higher nozzle travelling speed corresponds to a smaller value of the objective function.

Figure 2(c) shows the parallel coordinate plots of the microfluidic device using the two-optimisation method to generate the predicted parameter values as well as the objective function values under $l'(x)$. Observation shows that it is easier to generate ideal droplets when the control parameters are at low water pressure as well as high oil pressure.

Figure 3(a,b) show parallel coordinate plots of the parameter values versus the objective function values before and after optimisation using the GP-BO method for the inkjet device at objective functions $l(x)$ and $l'(x)$, respectively, and for the microfluidic device at $l'(x)$ in Fig. 3(c). Each parallel coordinate plot involves 240 sets of parameter values as well as objective function values, 140 sets of initial sample data for the device, and 100 sets of data generated during the iterative process of the GP-BO optimisation method. The naming of the lines in the graph is the same as in Fig. 2, and Sample represents the initial sample data. Whether it is an inkjet device or microfluidic device, the data before optimisation is more chaotic, GP-BO method optimisation data iteration towards a clear, gradually approaching the objective function value of the smaller region, which can be seen in the region where the convergence of the control parameters. The above can verify the reliability of the GP-BO optimisation method in this paper.

In order to verify that GP-BO is superior to the BO in Loop optimisation method in terms of output objective function values and their corresponding predicted parameter values, 10 sets of results with the smallest objective function values of the two methods are taken for comparison.

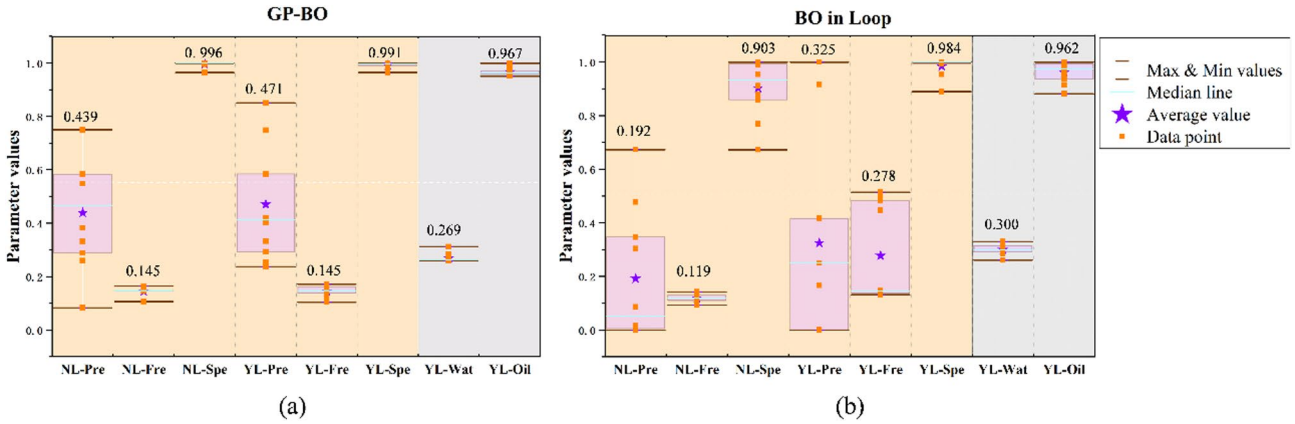

Fig. 5. Boxplots of the 10 sets of minimum objective function values corresponding to the predicted parameter values generated by the CIJ device using the GP-BO and BO in Loop optimisation methods, respectively. (**a**) Boxplots of predicted parameter outputs optimised using the GP-BO method for the inkjet device with two objective functions $l(x)$ and $l'(x)$, and for the microfluidic device with $l'(x)$; (**b**) Boxplots of predicted parameter outputs optimised using the BO in Loop method for the inkjet device with two objective functions $l(x)$ and $l'(x)$, and for the microfluidic device with $l'(x)$.

Figure 4(a,b) show line plots of 10 sets of minimum objective function values output by the inkjet device under the objective function $l(x)$ and $l'(x)$, and the microfluidic device under $l'(x)$ using the GP-BO and BO in Loop methods, respectively, with the same line nomenclature as in Fig. 2. Observing Fig. 4(a), it can be seen that when using the objective function $l'(x)$, the GP-BO method outputs objective function values between [0.362–0.428] and the BO in Loop method outputs objective function values between [0.431-1.000]. When the objective function $l(x)$ is used, the GP-BO method outputs objective function values between [0.331–0.388] and the BO in Loop method outputs objective function values between [0.378–0.428].

In Fig. 4(b), when the BO in Loop method uses EI as the acquisition function, the final output is instead not as good as the first iteration, when replacing the acquisition function with LCB as well as MPI, the final optimised output objective function value is only 0.026 different from the minimum objective function value outputted by one iteration of the EI as the acquisition function, so the method selects the results of one round of iterations to compare with the output of the GP-BO method. When using the objective function $l'(x)$, the GP-BO method outputs the objective function value between [0.046–0.176], and the BO in Loop method outputs the objective function value between [0.073–0.196].

In summary, in inkjet as well as microfluidic device, the GP-BO method outputs overall lower objective function values than the BO in Loop optimisation method, regardless of whether the objective function $l'(x)$ or $l(x)$ is used.

Figures 4(c)-(d) show the 10 sets of predicted parameter values output by the BO in Loop method and the optimisation method GP-BO in this paper for the inkjet device and the microfluidic device under the same objective function $l'(x)$, respectively. The bar and line graphs in the Fig. 4 represent the outputs of the BO in Loop method and the GP-BO method, respectively, and the bar and line graph shapes of the same colour scheme represent the same control parameter. In Fig. 4(c), the green bars and squares represent the fluid pressure, the red bars and dots represent the piezoelectric drive frequency, and the purple bars and triangles represent the nozzle translation speed. The BO in Loop method outputs that the fluid pressure values in the 1st, 5th, and 8th sets of predicted parameters are 0, so they are not shown on the figure. In the inkjet device, the BO in Loop and GP-BO methods yielded roughly the same conclusions, forming ideal droplets at lower piezoelectric drive frequencies and faster nozzle movement. In Fig. 4(d), the pink bars and squares represent water pressure, and the green bars and dots represent oil pressure. In the microfluidic device, the optimisation method in this paper yields similar conclusions as the BO in Loop method, where ideal droplets are generated under low water pressure as well as high oil pressure conditions. The specific details of the predicted parameter values are shown in Fig. 5.

Figure 5 shows the boxplots of 10 sets of predicted parameter values generated using GP-BO and BO in Loop optimisation methods for the inkjet device as well as the microfluidic device, respectively. The horizontal coordinates in Fig. 5 are named, NL and YL represent the use of the objective function $l(x)$ and $l'(x)$, and the suffixes are abbreviations for different control parameters, Pre, Fre, and Spe represent the three control parameters fluid pressure, piezoelectric drive frequency, and nozzle moving speed of the inkjet device, respectively, and Wat and Oil represent the two control parameters water and oil pressure of the microfluidic device. For example, the column NL-Pre represents the value of fluid pressure output under the $l(x)$.

In Fig. 5, the purple rectangle is the box part of the box plot, the blue line is the median line, the star markers are the mean values of each predicted parameter value, and the specific values of the mean values are shown above the box. Compared with the BO in Loop method, the box part of the predicted parameter box plot of the GP-BO optimisation method is smaller, which indicates that the output data are more concentrated and the model prediction ability is more stable, removing the fluid pressure, the remaining predicted parameter values fluctuate up and down above and below the median line no more than 0.1.

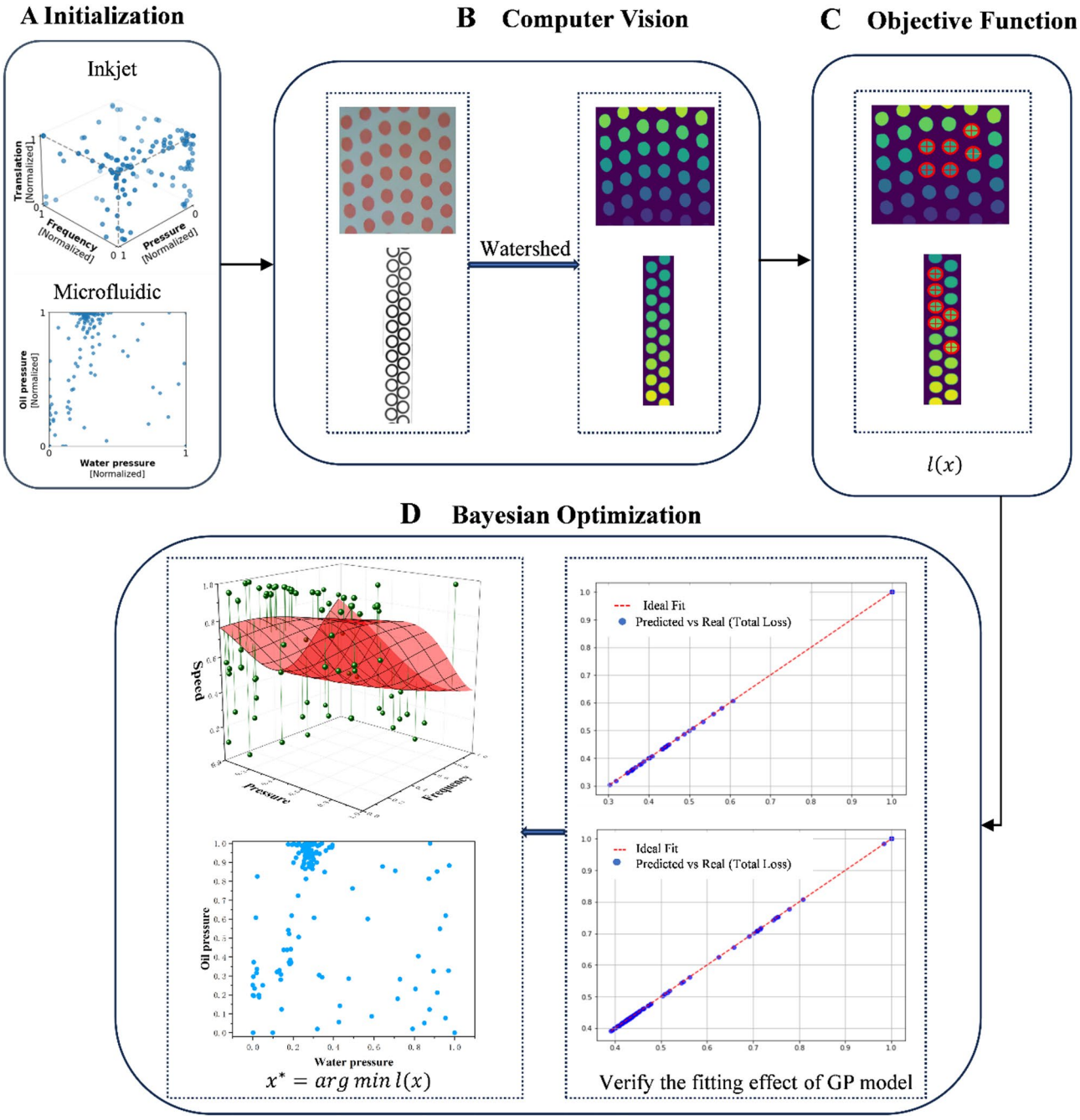

Fig. 6. GP-BO optimisation flow chart.(**A**) Initialisation of the control parameters is performed to normalise the parameter values to between [0–1]. Shown above and below are scatter plots of the initial control parameter distributions for the millimetre-scale inkjet device as well as the microfluidic device, respectively. (**B**) 140 droplet images generated under different parameter controls for each of the millimetre-scale inkjet device as well as the microfluidic device were used as the dataset. The droplet images corresponding to the initialised control parameter values are input into the computer vision module, and the computer vision uses the watershed segmentation technique to separate the droplets from the background in the droplet images to obtain all the droplet pixels in the images. (**C**) The separated droplets are quantified by the objective function, and the separated droplets are scored for geometric roundness, yield, and size uniformity detection and input into the BO module. (**D**) The control parameters are modelled using a GP, and the fit of the GP to the objective function is first examined, and when the fit is perfect, the input objective function values and control parameter values are BO'd and the predicted parameter values with the smallest objective function values, which consist of the loss of roundness, the loss of yield, and the loss of uniformity of size, are output.

A clear comparison between the GP-BO and GP-BO optimisation methods in terms of the number of optimisation iterations, the initial sample size, the number of samples involved in the optimisation process, and the optimisation duration is presented in supplementary Table S4, which shows the mean and standard deviation of the 12 model optimisation runs. The BO in Loop method samples the initial dataset and selects 20 groups of samples for optimisation, each optimisation generates 10 groups of predictive parameters, and the four iterations involve a total of 60 groups of sample data. The GP-BO method iterates the optimisation of the initial 140 groups of control parameters and generates 2 groups of new predictive parameters in each iteration, and the optimisation process involves a total of 240 groups of sample data.

The two optimisation methods have the same initial sample size, and the GP-BO method applies different loss functions to the initial dataset, with all the experimental durations adding up to about one hour. As a whole, the GP-BO method has more overall iterations than the BO in Loop method, but reduces the time needed to run experiments on the device for each round of iterations, saving about 1.3 h in total optimisation time.

Discussion

The ultimate goal of the inkjet printing process is to control the precise generation of droplets, but due to the difficulty of resolving the relationship between the control parameters and droplet characteristics, previous researchers have done a large number of experiments or used fluid dynamics simulation software or machine learning methods to predict the relationship between the control parameters and the droplets, but these methods generally have poor equipment adaptability, high computational cost, or rely on experimental validation. However, these methods generally suffer from poor equipment adaptation, high computational cost, or dependence on experimental verification. In this paper, a multi-parameter droplet optimisation model, which combines the BO algorithm with computer vision to optimise the droplet control parameters, is designed and validated on two different device datasets, a millimetre-scale inkjet device and a microfluidic device.

The model is iteratively optimised for 50 rounds, with a total optimisation time of 1 h, to optimise the output of the best control parameter values, which will generate high roundness and high yield droplets for the inkjet device under full fluid pressure, low drive frequency and high nozzle travel speed, and for the microfluidic device under low water pressure and high oil pressure. With the same amount of initial data, compared with the previous BO in Loop method, the GP-BO method saves the time required to run experiments on the hardware device for each iteration although the number of optimisations is increased, and the overall optimisation time is saved by nearly 1.3 h, which greatly improves the optimisation efficiency. The optimisation method in this paper is compared with the BO in Loop method in terms of the output objective function value and the predicted parameter value, and the output minimum objective function value of the method in this paper is reduced by 0.047 in the inkjet device, and reduced by 0.027 in the microfluidic device, and the output predicted parameter value is more stable.

The multi-parameter droplet optimisation model proposed in this paper has a good performance in control parameter prediction, as it does not require a mathematical or physical model to perform the optimisation, and purely relies on the objective function relationship between the control parameters and the droplet image to guide the optimisation model, which means that the model can be applied to a wider range of devices. The parameters are adjusted according to the actual demand, and can be converged to the optimal parameter values in a short time, which reduces the cost of experimental trial and error.

Methods

Figure 6 illustrates the workflow diagram of the workflow proposed in this paper for multi-parameter droplet optimisation. Processes (B)-(D) perform a total of 50 rounds of BO iterative optimisation, recording the combinations of predicted parameter values and corresponding objective function values in each round, and progressively improving the BO prediction of the objective function. Each round of optimisation generates two minimum sets of objective function values and their corresponding predicted parameter values, and a total of 100 sets of objective function values as well as predicted parameter values are generated. The 10 smallest sets of objective function values and their corresponding predicted parameter values are selected as the output of the Pareto front, where the predicted parameter values produce droplets with high roundness, high yield, and uniform size.

Parameter space initialisation

In this paper, we use the method of BO to learn the topology of the parameter space, which needs to be searched globally in order to find the best control parameters. To make BO more efficient and stable, the control parameter values are normalised.

Computer vision

Image segmentation is one of the key image processing techniques in computer vision[23,24]. Watershed transform is an efficient image segmentation algorithm that has been widely used in image preprocessing[19,25]. The image is fed into the computer vision module, and the computer vision performs watershed segmentation processing on the droplet image to separate the droplets from the background. In Fig. 6(B), the left side shows the original image and the right side shows the watershed segmentation processing to separate the droplets from the background.

The watershed algorithm relies heavily on distance transformations and morphological operations. Otsu threshold segmentation converts the original image to a binary image and this method automatically calculates the binarisation threshold T to minimise the within-class variance:

$$\sigma_w^2(T) = \omega_0(T)\,\sigma_0^2(T) + \omega_1(T)\,\sigma_1^2(T) \tag{1}$$

.

In Eq. (1), w_0 and w_1 are the proportions of background and foreground pixels (droplets), respectively, and σ_0^2 and σ_1^2 are the variances of the droplets and the background pixels. The Otsu method iterates over all the possible thresholds T and selects the value that minimises $\sigma_w^2(T)$.

The distance transform is used to calculate the shortest distance from each droplet to the nearest background pixel in the binary image. It is calculated as:

$$D(x,y) = \min_{(x',y')\in B} d\left((x,y),\left(x',y'\right)\right) \tag{2}$$

$$d\left((x,y),\left(x',y'\right)\right) = \sqrt{(x-x')^2+(y-y')^2} \tag{3}$$

.

In Eqs. (2- 3), (x,y) is the position of the droplet, B is the set of background pixels, and $d((x,y),(x',y'))$ denotes the Euclidean distance. The purpose of the distance transformation is to generate a gradient map with high intensity at the centre of the droplet and low intensity at the edges.

The watershed algorithm treats the gradient map as a topographic height map and divides the area according to the watershed filling water:

$$\frac{\partial I}{\partial x} = \frac{I(x+1,y) - I(x-1,y)}{2} \tag{4}$$

$$\frac{\partial I}{\partial y} = \frac{I(x,y+1) - I(x,y-1)}{2} \tag{5}$$

.

Gradient mode calculation:

$$|\nabla I(x,y)| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \tag{6}$$

In Eqs. (4- 6), $I(x,y)$ represents the pixel value at position (x,y) in the image. Eq. 6 represents the magnitude of the 'slope' of the intensity change in the image at point (x,y). Where the gradient is large is the 'ridge' between the droplet and the background, and the watershed algorithm is based on the 'ridge' to divide different regions. The watershed transform constructs an initial seed region based on the minima of the gradient map and then performs region growth until it fills the entire image[25]. The watershed algorithm is executed for final image segmentation.

Objective function

As shown in Fig. 7, droplets are separated from the background using watershed segmentation, and the separated droplets form a set of indexed pixels P_{drop}[16]. The generated droplet indexes facilitate the calculation of the size of the objective function in each image, with the aim of minimising the value of the objective function in order to achieve the ideal droplet with high roundness, high yield, and uniform size.

In the inkjet device, if only the droplet geometry roundness loss as well as the yield loss are considered, the total objective function value of the image with the presence of highly rounded satellite droplets in the initial image data set is instead small, which gives a wrong guidance to the optimisation process and makes the result deviate from the expected one. To avoid this, a size uniformity loss is added to the objective function. The optimisation process focuses on the effect of geometric loss and yield loss on droplets, so these two items account for a large proportion of the objective function, while the addition of dimensional uniformity loss to increase the objective function value of the image of the presence of satellite droplets can make the results in the production of high roundness, high yield droplets at the same time, to avoid the generation of satellite droplets. In Eq. (7), the objective function $l(x)$ is composed of L_{geom}, L_{yield} and L_{size}, and the ratio of these three loss functions is 0.425:0.425:0.15, $l(x) \in [0,1]$. The formulas for the three loss functions are detailed in the "Objective function" section of the supplementary methods. The optimisation process focuses on the effect of geometric loss and yield loss on the droplets, so these two items account for a larger proportion in the objective function.

$$l(x) = 0.425L_{geom} + 0.425L_{yield} + 0.15L_{size} \tag{7}$$

Bayesian optimisation

BO is an effective global optimisation strategy that has two components, a GP and an acquisition function[19,26,27].

As shown in Table 1, the objective of the BO is to find the minimum value l^* of the objective function $l(x)$ corresponding to the predicted parameter value x^*, which in this case controls the production of droplets with high roundness, high yield, and uniformity in size.

In Eq. (8), x^* is the predicted parameter value corresponding to the minimum objective function value $l(x)$, $x^* \in X^{(N)}$, $X^{(N)}$ represents the search space of x, and N is the space dimension.

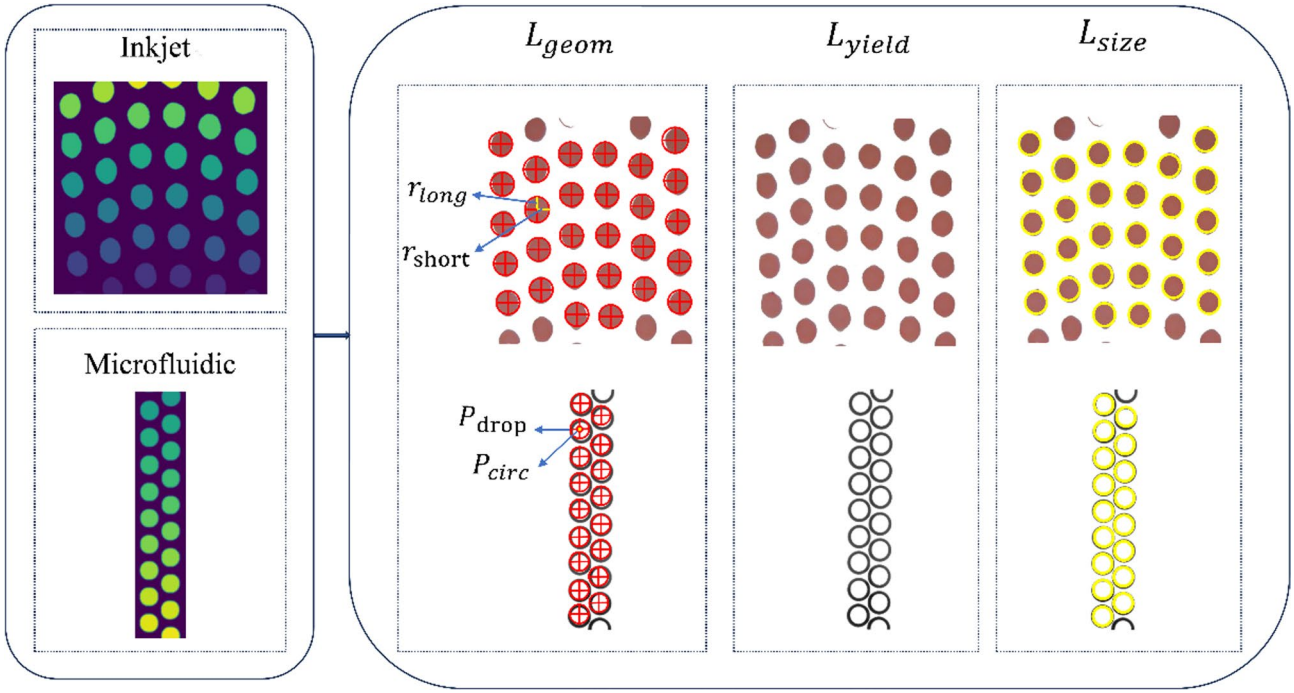

Fig. 7. Droplet pixels segmented out of the inkjet device as well as the microfluidic device droplet image watershed are quantised by an objective function $l(x)$. The objective function consists of a geometric loss L_{geom}, a yield loss L_{yield}, and a size uniformity loss L_{size}, based on which all droplet pixel objective function values are calculated for each image.

Algorithm 1 GP construct the BO droplet model GP-BO
Input: The number of Pareto solutions, batch_size: Control parameter values $x\, x \in X^{(N)}$ Droplet images corresponding to each set of control parameter values; BO model optimisation times, iterations.
Output: The mean square error MSE, and the coefficient of determination R^2; Select the $l(x)$ with the smallest objective function in the first batch_size and its corresponding prediction parameter value x_pred to form the Pareto front, $x^* = \arg\min l(x)$ $l^* = \min(l(x))\, x^* \in x_{pred}\, l^* \in l(x)$
While no convergence do **1. Initialise**: Build the GP model and choose EI as the acquisition function. **2. Objective function calculation**: $l(x) = 0.425L_{geom} + 0.425L_{yield} + 0.15L_{size}$ the pixels of the droplets segmented by computer vision are quantitatively analysed by the objective function $l(x)$, where $l(x) \in [0,1]$
3. Evaluation of GP model fit: The true loss value $l(x)$ is compared with the predicted loss value $l_{pred}(x)$ for the sampling point and MSE, R^2 is calculated. $l(x) = l_{pred}(x)$ if $MSE = 0.0$ and $R^2 = 1.0$
4. BO loops: A total of iterations of loops are executed, with the acquisition function EI guiding the selection of x to the point where the current optimal solution can be improved, finding $l(x)$ such that $l(x) < \min(l(x))$, where $\min(l(x))$ is the smallest currently known value of the objective function, and $l(x) \sim GP(\cdot)$. end while return x_{pred}, $l(x)$.

Table 1. Pseudo-code for the GP-BO algorithm.

GP is the most commonly used agent model in BO. In Eqs. (9- 11), the GP consists of a mean function $m(x)$ given by and a covariance function $k(x, x')$ which is used to predict how $l(x)$ will change when $x \in X^{(N)}$ changes.

$$l(x) \sim GP\left(m(x), k\left(x, x'\right)\right) \tag{9}$$

$$m(x) = E[l(x)] \tag{10}$$

$$k\left(x, x'\right) = E\left[\left(l(x) - m(x)\right)\left(l\left(x'\right) - m\left(x'\right)\right)\right] \tag{11}$$

The covariance function represents the correlation of the process, if two points x and x' are strongly correlated, then the correlation between $l(x)$ and $l(x')$ is also strong.

$$P(lx) = \mathcal{N}\left(\mu(x), \sigma^2(x)\right) \tag{12}$$

In Eq. (12), for a given input point x, the prediction of the objective function $l(x)$ by means of a GP model results in a normal distribution of the predicted mean $\mu(x)$ with uncertainty $\sigma(x)$.

The collection function selects the next most promising point based on the predicted mean $\mu(x)$ and the variance $\sigma(x)$[28–30]. In this paper, Expected Improvement (EI) is chosen to guide the selection of x. The EI function is not prone to falling into local optimal solutions, and its goal is to select the point that improves the current optimal solution, i.e., to find the $l(x)$ making $l(x) < \min(l(x))$[17]. The EI can be calculated by in Eq. (13)

$$EI(x) = (\mu(x) - \min(l(x)) - \xi) \cdot \Phi(Z) + \sigma(x) \cdot \phi(Z) \tag{13}$$

$$Z = \frac{\mu(x) - min(l(x)) - \xi}{\sigma(x)} \tag{14}$$

In Eq. (14), $\min(l(x))$ is the minimum currently known value of the objective function and ξ is a trade-off parameter that balances the acquisition function between exploitation ($\mu(x)$ sampling at higher levels) and exploration ($\sigma(x)$ sampling at higher levels). $\Phi(Z)$ is the cumulative distribution function of the standard normal distribution and $\varphi(z)$ is the probability density function of the standard normal distribution.

Experimental parameter setting

In millimetre-scale inkjet device, the nozzle translation speed controls the deposition position of the droplets on the substrate, and the fluid pressure determines the speed and flow rate of the ink from the nozzle. The inkjet device was operated in the pressure range of 0.03–0.15 MPa, the piezoelectric actuation frequency was limited to 1–600 Hz, and the nozzle translation speed was limited to 10–36 mm/s. The operation was carried out at a flow rate of 0.8–4 m/s for a given pressure range. The experiments were carried out in a controlled variable manner with the fluid (dyed water) and nozzle diameter (70 μm) controlled constant[16].

The control parameters for inkjet printing include the nozzle diameter, material properties of the ink such as surface tension, viscosity, density, etc., which are usually expressed as dimensionless numbers[2,31,32]. In inkjet printing, the Reynolds number (Re), Ohnesorge number (Oh), and Weber number (We) are involved[31,33]. In the experiment, $Oh \approx 0.01$, $We \approx$ 1-200, $Re \approx$ 10–50 are fixed. $Bond\ number \approx 0.001$, the gravity can be neglected.

The microfluidic device has two control parameters: water pressure and oil pressure. These two control parameters are controlled by a flow controller and the device operates over a pressure range of 0–2,000 mbar[16].

The droplet properties of the microfluidic system are described by two factorless numbers: the capillary number (Ca), and the Weber number (We).The dimensionless numbers are in the range of $Ca \approx$ 0.05–0.5 and $We \approx$ 0.0005–0.003, respectively.

Data availability

Data is provided within the manuscript or supplementary information files.

Received: 1 April 2025; Accepted: 27 June 2025
Published online: 17 July 2025

References

1. Hu, H. et al. Inkjet-Printed tungsten oxide memristor displaying Non-Volatile memory and neuromorphic Properties. *Adv. Funct. Mater.* **34** (20), 2302290 (2024).
2. Lohse, D. Fundamental fluid dynamics challenges in inkjet printing. *Annu. Rev. Fluid Mech.* **54** (1), 349–382 (2022).
3. Shin, D. Y., Moon, Y. J., Ju, B. K. & Kang, K. T. Measurement of inkjet droplet speed using interference fringe by diffracted light. *Sci. Rep.* **14** (1), 22364 (2024).
4. Otowa, T., Tsubouchi, S. & Suwa, Y. Analysis of the Ink-stream Break-Up phenomenon in continuous inkjet Printing. *ACS Omega*. **8** (38), 34442–34447 (2023).
5. Hou, M. et al. High-precision silver electrode based on PEN substrate with robust mechanical performance. *Surf. Interfaces*. **54**, 105158 (2024).
6. Zhang, S., Fang, K., Liu, X., Qiao, X. & Wang, J. Simplified and efficient inkjet printing of cotton fabrics using cationic colored nanoparticles. *Ind. Crops Prod.* **193**, 116217 (2023).
7. Maîtrejean, G. et al. Comprehensive experimental dataset on large-amplitude Rayleigh-Plateau instability in continuous inkjet printing regime. *Data Brief.* **52**, 109941 (2024).
8. Schilling, O. Self-similar Reynolds-averaged mechanical–scalar turbulence models for Rayleigh–Taylor, Richtmyer–Meshkov, and Kelvin–Helmholtz instability-induced mixing in the small Atwood number limit. *Phys. Fluids*, **33**(8). (2021).
9. Briard, A., Gréa, B. J. & Nguyen, F. Turbulent mixing in the vertical magnetic Rayleigh–Taylor instability. *J. Fluid Mech.* **979**, A8 (2024).
10. Wang, Z. et al. High-quality semiconductor fibres via mechanical design. *Nature* **626** (7997), 72–78 (2024).
11. Yao, P., Li, G., Li, Y. & Gao, J. Free surface tension modelling using particle-grid hybrid method without considering gas particles. *J. Comput. Phys.* **498**, 112674 (2024).
12. Chu, A. B., Nguyen, D., Kaplan, A. D. & Giera, B. Image classification and control of microfluidic systems[C]//Applications of machine learning. *SPIE* **11139**, 23–29 (2019).
13. Solanki, S. et al. Machine learning for predicting microfluidic droplet generation properties. *Comput. Fluids*. **247**, 105651 (2022).
14. Chu, A. et al. Automated detection and sorting of microencapsulation via machine learning. *Lab. Chip.*. **19** (10), 1808–1817 (2019).
15. Rodriguez-Rivero, M. C., Philpott, J. M., Hann, A. B., Harries, J. L. & Daly, R. Deflecting the Issue: The Origin of Nanoscale Material Build-up in Continuous Inkjet Printing[C]//NIP & Digital Fabrication Conference. Society for Imaging Science and Technology, 36: 44–53. (2020).
16. Siemenn, A. E. et al. A machine learning and computer vision approach to rapidly optimize multiscale droplet generation. *ACS Appl. Mater. Interf.*. **14** (3), 4668–4679 (2022).

17. Tran, Q. D., Shin, D. & Jang, G. W. Bayesian optimization-based topology optimization using moving morphable bars for flexible structure design problems. *Eng. Struct.* **300**, 117103 (2024).
18. Li, J., Lu, Y. & Lu, R. Detection of early decay in navel oranges by structured-illumination reflectance imaging combined with image enhancement and segmentation. *Postharvest Biol. Technol.* **196**, 112162 (2023).
19. Fu, Y. & Aldrich, C. Online particle size analysis on conveyor belts with dense convolutional neural networks. *Miner. Eng.* **193**, 108019 (2023).
20. Nguyen, M. D. et al. Generalized correlation for predicting the droplet size in a microfluidic flow-focusing device under the effect of surfactant. *Phys. Fluids*, **34**(3). (2022).
21. Liu, Z. et al. Effects on droplet generation in step-emulsification microfluidic devices. *Chem. Eng. Sci.* **246**, 116959 (2021).
22. Lee, J. M. et al. *Generation of Tumor Spheroids Using a droplet-based Microfluidic Device for Photothermal therapy*652 (Microsystems & Nanoengineering, 2020).
23. Oza, P., Sindagi, V. A., Sharmini, V. V. & Patel, V. M. *Unsupervised Domain Adaptation of Object Detectors: A survey* (IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023).
24. Khan, S. et al. Transformers in vision: A survey. ACM computing surveys (CSUR), **54**(10s): 1–41. (2022).
25. Kucharski, A. & Fabijańska, A. CNN-watershed: A watershed transform with predicted markers for corneal endothelium image segmentation. *Biomed. Signal Process. Control.*. **68**, 102805 (2021).
26. Teng, C. et al. Dual-level training of Gaussian processes with physically inspired priors for geometry optimizations. *J. Chem. Theory Comput.* **18** (9), 5739–5754 (2022).
27. Horak, J. et al. Reconstructing QCD spectral functions with Gaussian processes. *Phys. Rev. D.* **105** (3), 036014 (2022).
28. Zhang, S., Yang, F., Yan, C., Zhou, D. & Zeng, X. An efficient batch-constrained bayesian optimization approach for analog circuit synthesis via multiobjective acquisition ensemble. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **41** (1), 1–14 (2021).
29. Lourenço, M. P. et al. Automatic structural Elucidation of vacancies in materials by active learning. *Phys. Chem. Chem. Phys.* **24** (41), 25227–25239 (2022).
30. Tian, Y. et al. Air entrapment of a neutral drop impacting onto a flat solid surface in electric fields. *J. Fluid Mech.* **946**, A21 (2022).
31. Kang, S. H., Kim, S., Sohn, D. K. & Ko, H. S. Analysis of drop-on-demand piezo inkjet performance. *Phys. Fluids*, **32**(2). (2020).
32. Ray, S., Chi, Y., Zhang, P. & Cheng, S. Head-on collision of unequal-size droplets on a wetting surface. *Phys. Fluids*, **35**(2). (2023).
33. Li, Y. et al. Effect of pore structure and capillary number on gas-water flow patterns in carbonate rocks. *SPE J.* **27** (04), 1895–1904 (2022).

Acknowledgements

This work was supported by the Beijing Key Laboratory of Signal and Information Processing for High-End Printing Equipment (Project Name: Beijing Institute of Graphic Communication Research Platform Construction Project (Project No.: KYCPT202509)); The Beijing Municipal Commission of Education's Emerging Interdisciplinary Platform for Publishing - Digital Inkjet Printing Technology and Multifunctional Rotary Offset Printing Machine Key Technology Research and Development Platform (Project No.: 04190123001/003, Responsible Person: Likun Lu); Case Study on the Integration of Intelligent Printing Industry and Education--Taking the Course of Digital Inkjet Printing and Inspection as an Example (Project No.: MS2023168, Responsible Person: Qingtao Zeng); Research on Key Technology of Artificial Intelligence for Resource Management in Transcoding of Integrated Media Video (Project No.: E6202405, Responsible Person: Erqing Zhang); Beijing Institute of Graphic Communication Advantageous Discipline Construction Project (Project No.: 21090525004); Beijing Institute Of Graphic Communication Discipline Construction and Postgraduate Education Special (Project Name: Construction of a Comprehensive Platform for Deepening Reform and Quality Enhancement of Graduate Education in Drop-on-demand Publishing Inkjet Printing Technology (Project No.: 21090325003)).

Author contributions

T. Li conducted the experiment and wrote the main manuscript text. K. Liao conducted the experiment with T. Li. L. Lu and Q. Zeng supervised and guided this study. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-09435-8.

Correspondence and requests for materials should be addressed to L.L.