

ESTIMATING THE PARAMETERS OF A STOCHASTIC GEOMETRICAL MODEL FOR MULTIPHASE FLOW IMAGES USING LOCAL MEASURES

LÉO THÉODON^{✉,1}, TATYANA EREMINA¹, KASSEM DIA^{1,2}, FABRICE LAMADIE²,
JEAN-CHARLES PINOLI¹ AND JOHAN DEBAYLE¹

¹MINES Saint-Etienne, CNRS, UMR 5307 LGF, Centre SPIN, 158 cours Fauriel, 42023 Saint-Etienne Cedex 2, France, ²CEA, DES, ISEC, DMRC, Univ Montpellier, 30207 Bagnols-sur-Ceze, Marcoule, France
e-mail: l.theodon@emse.fr, tatyana.eremina@mines-stetienne.fr, Kassem.DIA@cea.fr, fabrice.lamadie@cea.fr, pinoli@emse.fr, debayle@emse.fr

(Received November 5, 2021; accepted November 27, 2021)

ABSTRACT

This paper presents a new method for estimating the parameters of a stochastic geometric model for multiphase flow image processing using local measures. Local measures differ from global measures in that they are only based on a small part of a binary image and consequently provide different information of certain properties such as area and perimeter. Since local measures have been shown to be helpful in estimating the typical grain elongation ratio of a homogeneous Boolean model, the objective of this study was to use these local measures to statistically infer the parameters of a more complex non-Boolean model from a sample of observations. An optimization algorithm is used to minimize a cost function based on the likelihood of a probability density of local measurements. The performance of the model is analysed using numerical experiments and real observations. The errors relative to real images of most of the properties of the model-generated images are less than 2%. The covariance and particle size distribution are also calculated and compared.

Keywords: Local Measures, Maximum Likelihood, Minkowski Functionals, Statistical Inference, Stochastic Geometry.

INTRODUCTION

Most physical and chemical recycling processes are based on multiphase flow (liquid-liquid extraction, dissolution, leaching, etc.), and require digital and experimental fluid mechanics studies to understand the possible effects of the flow on the physicochemical efficiency of the processes. Experimentally measuring flow properties, especially of the dispersed phase (concentration, size and shape distributions, etc.) has become an important issue in many energy industries, from hydrocarbon refining to nuclear fuel recycling or the recovery of valuable or polluting metals from electric batteries. As a common form of multiphase flow in physicochemical engineering, bubble flows have been extensively studied numerically and experimentally using image acquisition and image processing methods (Bello *et al.* (1985), Bai *et al.* (2012), Buffo *et al.* (2017)).

The various imaging techniques that can be used to characterize 2D images of bubbly flows (Paolinelli *et al.* (2018); Zhou *et al.* (2020); Torisaki *et al.* (2020); Haas *et al.* (2020); Cerqueira *et al.* (2018)) can be classified in two main categories. Deterministic methods mainly rely on segmentation, pattern recognition and geometrical methods to characterize the images (Cerqueira *et al.* (2021); Vinnett *et al.* (2020); De Langlard *et al.* (2018)). They

are usually fast but are less efficient for more complex flows (higher particle densities, larger clusters, more complex shapes). Stochastic methods, on the other hand, usually involve models able to generate digital images, the idea being to generate of large number of representations from the stochastic model and then determine the characteristics of the digital images obtained for a given set of parameter values. A cost function is then used to compare these characteristics with the values observed in the real images, and the best set of parameter values for the stochastic model is estimated using an optimization process. The main problem with this approach is that although these models are usually able to capture some of the complexity of real flows (De Langlard *et al.* (2021); Kracht *et al.* (2013)), they can be difficult to optimize, especially for more complex flows requiring a greater number of parameters.

Parametric models are often optimised using the covariance function (Serra (1988)), granulometry (De Langlard *et al.* (2018)) and Choquet capacity functional, along with global properties such as the Minkowski functionals (Molchanov (1997); Chiu *et al.* (2013)). This paper investigates the use of local measures to estimate the parameters of a geometrical model. This method is less time consuming than those based on the covariance and granulometry and more effective because it can discriminate binary

images with the same global measures of surface and perimeter. It has recently shown indeed (Eremina *et al.* (2021)), that different geometrical models may generate binary images with the same global measures of area and perimeter when their local geometric measures are different (Eremina *et al.* (2021)). Moreover, images with the same covariance (Rataj (2004); Benassi *et al.* (2010)) or granulometry (Lefèvre (2009)) can be obtained from very different random closed sets.

This article explores the possibility of using local measures to estimate the parameter values of a stochastic geometrical model for the geometrical characterization of multiphase flow images. The first section describes the way the real images used in the study were acquired and presents the baseline characterization using deterministic approaches. The STORMFLOW (STOchastic geometRical model for Multiphase FLOW) model is then introduced and the way local geometric measures can be used to retrieve its parameters is explained. The realizations of the STORMFLOW model presented and evaluated in the last section demonstrate the efficiency of the local measures approach. Future prospects are discussed in conclusion.

IMAGE ACQUISITION

The images used in this study come from an experimental mock-up designed for R&D fluid mechanics studies of a recycling process involving a bubbly flow. This section details how the images were acquired and discusses the information that can be extracted to characterize them geometrically.

EXPERIMENTAL IMAGING SETUP

The experimental optical imaging setup used to capture the images runs in "a transmitted light" configuration (also called backlight illumination in the literature). A light source and an imaging sensor are placed on either side of the flow, which in this case was a static water phase with air bubbles rising inside (Fig. 1).

In this setup, the size, number and shape of the bubbles depend on the gas injection rate. Furthermore, since the geometric characteristics of the bubbles depend directly on the injection rate, the bubbles are classified into families in terms of the gas flow rate used.

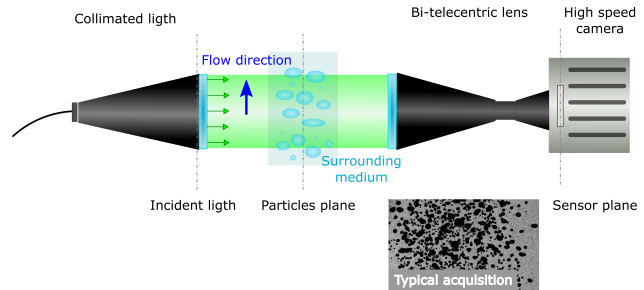


Fig. 1: Experimental setup used to acquire bubbly flow images.

IMAGES

The lens used to capture the optical images is bi-telecentric, which suppresses perspective effects in the field of view and allows the projected area of the particles to be displayed on the imaging sensor. The bubbles appear as opaque 2D ellipses. Projection overlap leads to the formation of clusters of increasing complexity as the flow rate increases (Fig. 2). Separating images of individual particles within these clusters is a challenging problem.

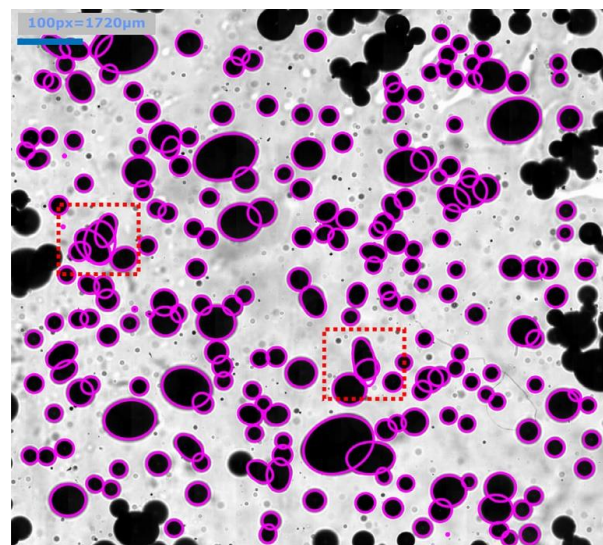


Fig. 4: Performance of the deterministic ellipse recognition algorithm. The particle images intersecting the border of the observation window are ignored. The red boxes highlight two incorrect fits due to the shape of the clusters. Picture from De Langlard *et al.* (2018).

PARTICLE CHARACTERIZATION

For the sake of simplicity, gas bubbles (the black shapes on the 2D images) are referred to as particles from now on. The characteristic geometrical properties of the particles are their shape (elongation ratio), size (major and minor axis) and orientation distribution,

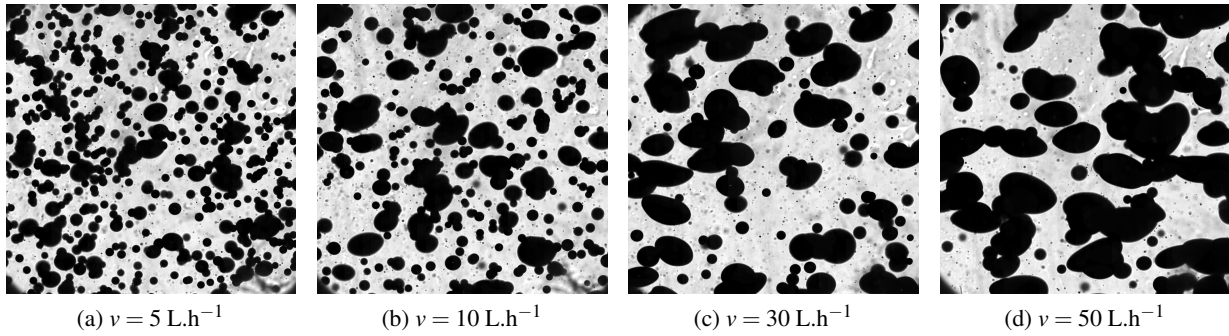
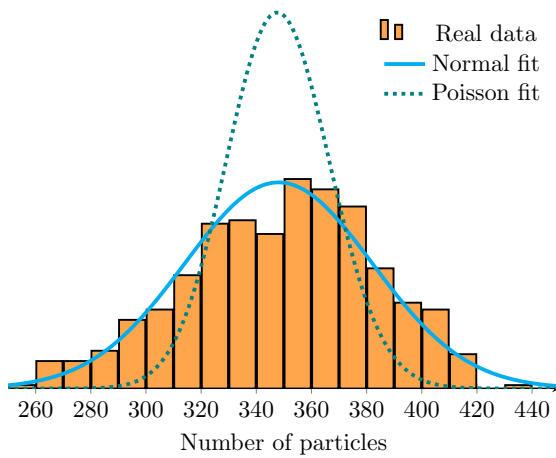
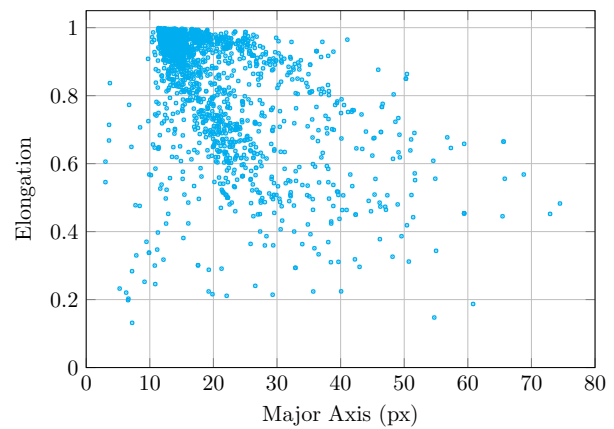


Fig. 2: Typical images captured with the experimental imaging setup. The flow rate v (L.h^{-1}) clearly affects the shape of isolated particles and of particle clusters.



(a) Distribution of the number of particles.



(b) Distribution of the major axis and elongation ratio of the particles.

Fig. 3: Data extracted from the images at a flow rate $v = 10 \text{ L.h}^{-1}$. (a) The average number of particles per image is best fit by a normal distribution. (b) The major axis and the elongation ratio of the particles are clearly correlated. (Each dot represents a particle.)

as well as their density and spatial distribution in the images.

After binarizing the images, as is facilitated by the fact that the particles appear as black objects on a homogeneous light-grey background with well-defined edges, De Langlard *et al.*'s (2018) ellipse recognition algorithm was applied to obtain the size, shape and orientation distribution of the particles.

Fig. 4 highlights the limitations of this algorithm. Large, complex particle clusters are incorrectly processed and particles or particle clusters that overlap with the border of the images cannot be correctly decomposed and are consequently ignored. Finally, while the algorithm performs well when the particles are perfectly elliptical, it is less robust when dealing with non-elliptical shapes such as caps. Other algorithms available, such as Bettaieb *et al.*'s (2020), have the same limitations. The measurements obtained

are therefore incomplete, especially for high flow rates, when particles become deformed and complex clusters form.

The results of this characterization can nevertheless be used as input for a geometrical model. The spatial distribution of particles is assumed to be uniform and the average number of particles follows an approximately normal distribution of mean μ_v and standard deviation $\sigma_v \sim \tilde{\sigma}_v = 0.1\mu_v$, regardless of the flow rate. Finally, the size (major axis), shape (elongation ratio) and orientation distribution of the particles are strongly correlated.

2D binary images can thus be characterized at any flow rate, despite the limitations of the ellipse recognition algorithm, and the values obtained can be used to define a stochastic geometric model whose validity and robustness can be validated by comparison with real observations.

Flow rate ν	μ_ν	σ_ν	$\tilde{\sigma}_\nu$	Error
5 L.h ⁻¹	390.7	40.2	39.1	-2.7%
10 L.h ⁻¹	348.1	34.0	34.8	+2.3%
30 L.h ⁻¹	125.9	12.3	12.6	+1.9%
50 L.h ⁻¹	120.7	12.2	12.1	-0.9%

Table 1: Comparison of σ_ν and $\tilde{\sigma}_\nu$, respectively the standard deviation of the normal distribution that best fits the probability distribution of the number of particles per image, and its approximation.

Note finally that the average number of particles does not follow a Poisson probability distribution, which means that the Boolean model (Chiu *et al.* (2013)) and Minkowski functionals cannot be used. Moreover, even if the spatial distribution of the particles were considered uniform, the images show the presence of aggregates, which are not well taken into account by the Boolean model.

GEOMETRICAL MODEL USING LOCAL MEASURES

GERM-GRAIN MODEL

While copula-based approaches (Schmidt (2007)) are too complex to implement (they involve more than 60 parameters), the much simpler STORMFLOW model is built around just two parameters: μ , the mean number of particles, and S , the joint probability distribution of the major axis and elongation ratio of the particles. The particles are assumed to be uniformly distributed in space and the number of particles is assumed to follow a normal distribution $\mathcal{N}(\mu, \sigma)$, with $\sigma = 0.1\mu$. Moreover, the orientations of the particles are assumed to follow a von Mises distribution (Kurz *et al.* (1999)), which is an approximation of a wrapped normal distribution (Mardia *et al.* (1999)), i.e. a normal distribution "wrapped" around a unit circle. The variance of this distribution depends on the elongation ratio of the particles, tending to infinity when the elongation ratio tends towards 1, i.e. when the particles are close to spherical, and decreasing when the elongation ratio decreases, i.e. when the particles have a preferential orientation.

The model can therefore be expressed as a type of germ-grain model:

$$\Xi = \bigcup_{i=1}^{\infty} (z_i + Z_i), \quad (1)$$

where the points (*germs*), z_i , belong to a stationary process in \mathbb{R}^2 , wherein the average number of points

in a given observation window W follows a normal probability distribution $\mathcal{N}(\mu, 0.1\mu)$. As for the *grains*, Z_i , they are i.i.d. random compact sets in $\mathbb{R}_+ \times [0; 1] \times [0; 2\pi[$ and independent from the germs, z_i . The grains are typically elliptical and follow the same probability distribution as Z_i . A typical grain, Z_0 , with a random major axis, M , a random elongation ration, e , and a random major axis orientation, θ , would be parametrized as $Z_0 = Z_0(M, e, \theta)$. As mentioned above, the random variables M and e follow a joint probability distribution S , and the orientations of the particles follow a von Mises distribution:

$$f_{\theta,e}(x) = \frac{\exp(\kappa \cos(x))}{2\pi I_0(\kappa)} \quad (2)$$

where I_0 is the modified Bessel function of order 0 and $1/\kappa$, which is analogous to the variance of a normal distribution wrapped around a unit circle, depends directly on the random elongation ratio e of the grains:

$$\kappa = 32 \times (1 - e) \quad (3)$$

meaning that $f_{\theta,e}$ is constant if $e = 1$, i.e. if the particles are perfectly spherical. The latter equation was determined empirically using the data from the ellipse recognition algorithm. Fig. 6 shows examples of the binary images generated with this model.

LOCAL MEASURES

A cost function is required to fit the STORMFLOW model parameters to real data, with the covariance function and the granulometry being common choices (Matheron (1972); Diggle (1981); Dereudre *et al.* (2014)). The global area, perimeter or Euler-Poincaré number can also be used. As mentioned in the introduction however, the intent here is to use the local Minkowski measures proposed by Eremina *et al.* (2021).

The principle of this approach is to place a geometric probe (basically a disk D with a radius usually very close to the average equivalent radius of the grains) at a point in the random set Ξ of particle positions. The local area, local perimeter and/or local Euler-Poincaré number of the subset $D \cap \Xi$ are then calculated. To avoid edge effects, the geometric probe D must remain in the observation window W , i.e. $D \subset W$. Here, only the local perimeter will be used since Eremina *et al.* (2021) have shown that this is the most discriminating measure.

Fig. 5 shows the histogram of local perimeter measurements obtained from the real images. Note that there is always a maximum for $P = 0$, where the geometric probe verifies $D \subset \Xi$. The second

maximum occurs at local perimeter values close to the diameter of the disk D used as a probe. The local measurement histograms corresponding to different flow rates are themselves very different, validating the strategy of relying on local measures to optimize the STORMFLOW model.

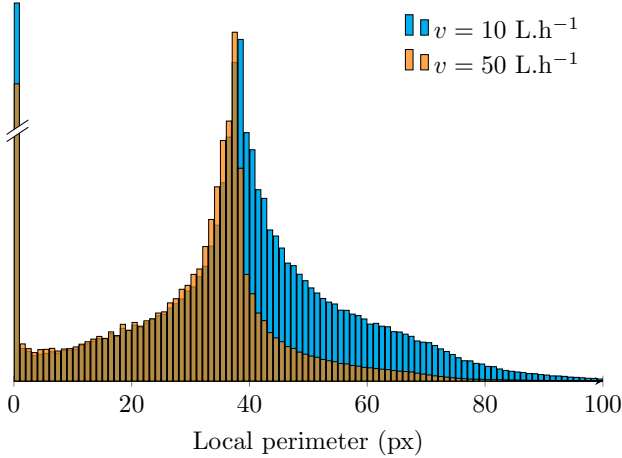


Fig. 5: Histogram of local perimeter measurements with a geometric probe of radius $r = 19$ px performed on real flow images at two flow rates ($v = 10$ L.h $^{-1}$ and $v = 50$ L.h $^{-1}$).

More precisely, to obtain n local perimeter measurements P_1, P_2, \dots, P_n , n_{loc} measurements are performed per image on a set of N images so that $n = N \times n_{loc}$. Thus, for each of the n_{loc} observations on every image I within the set of N images, the local perimeter is calculated for $D_i \cap \Xi_{v,I}$, where D_i , $i = 1 \dots n_{loc}$ is the geometric probe placed at a randomly chosen position in $\Xi_{v,I}$, the set of particle positions in image I of a flow of particles with flow rate v . Hence, $\Xi_{v,I}$ can be seen as the restriction of Ξ_v to an observation window the size of image I .

COST FUNCTION

Let $f_{v,\mu,S}$ be the probability distribution function of a random variable P , such as the local perimeter, with (μ, S) defined as above. The likelihood of (μ, S) and f_v with respect to the sample of n observations P_1, P_2, \dots, P_n is defined by:

$$L_{\Xi_v}(\mu, S) = \prod_{i=1}^n f_{v,\mu,S}(P_i). \quad (4)$$

As the $N \times n_{loc}$ locations are independent and identically distributed, so are the local perimeter measures. The maximum likelihood of the average

number of particles μ and the joint probability distribution S can then be estimated as:

$$\left(\hat{\mu}, \hat{S}\right) = \arg \max_{(\mu, S)} L_{\Xi_v}(\mu, S). \quad (5)$$

The cost function of the STORMFLOW model, M_v , is defined as the opposite of the log-likelihood, namely:

$$M_v = -\log(L_{\Xi_v}(\mu, S)) = -\sum_{i=1}^n \log(f_{v,\mu,S}(P_i)). \quad (6)$$

Indeed, since the logarithm function is increasing and real-valued, maximizing the likelihood function is equivalent to minimizing the opposite of the log-likelihood function, which is easier to calculate.

The probability density function $f_{v,\mu,S}$ can be estimated from the histograms of local perimeter measurements performed on the large set of images generated with different parameter pairs, (μ, S) . The cost function of the STORMFLOW model at a flow rate v will therefore be

$$M_v(\mu, S) = -\sum_{i=1}^n \log(\hat{f}_{v,\mu,S}(P_i)) \quad (7)$$

where $\hat{f}_{v,\mu,S}$ is the estimator of $f_{v,\mu,S}$.

However, to allow the use of an efficient optimization algorithm, the cost function has to be defined over a spatial domain much smaller. The STORMFLOW model was therefore adjusted to take two real numbers $(\mu, t) \in \mathbb{R}_+ \times [0, 1]$ as input instead of (μ, S) .

PARAMETRIC MODEL DEFINITION

To simplify the input of the first iteration of the STORMFLOW model and strengthen the optimization and validation steps, the data obtained with the ellipse recognition algorithm (Fig. 3b) were used to build a set of probability density surfaces, defined for each flow rate $v \in \{5; 10; 30; 50\}$ by:

$$S_v(d) : \mathbb{R}_+ \rightarrow \mathcal{S} \quad (8)$$

where \mathcal{S} is the spatial domain of the joint probability distribution of the elongation ratio and major axis, and d is a kernel parameter (Epanechnikov (1969)). Fig. 6 shows the joint probability distributions obtained from the data shown in Fig. 3b using this function. The idea is to normalize the real data at all flow rates, to avoid scaling issues and define a fixed grid in \mathbb{R}_+^2 . The number of points inside a disk of radius d centered on a given point of the grid are then counted and the result is normalized to obtain a discrete joint probability distribution.

Although four sets of 2D binary images were available, one for each of the four different flow rates, only the two extreme values $v = 5$ and $v = 50$

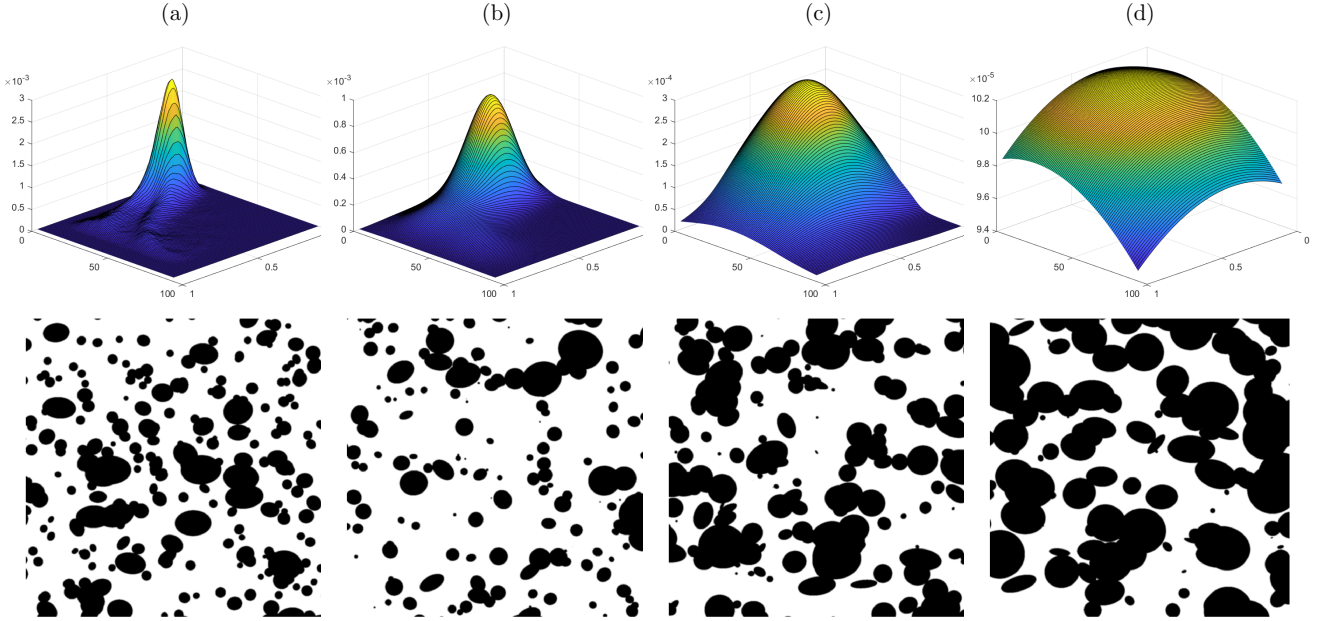


Fig. 6: Examples of the images generated using the geometrical model with the four probability density surfaces shown, and different mean numbers of particles, μ : (a) 400, (b) 350, (c) 300, and (d) 250. The surfaces were reconstructed using the data shown in Fig. 3b and an increasing kernel parameter d .

were considered in determining the estimators of the joint probability density distributions \widehat{S}_5 and \widehat{S}_{50} . The optimal values of (μ, d) that minimize the cost function and maximize the likelihood, given the local perimeter measurements, can be found by minimizing the cost function, $M_v(\mu, S)$ defined by equation (7),

$$\widetilde{M}_v(\mu, d) = M_v(\mu, S_v(d)), \quad (9)$$

Fig. 7 shows the joint probability distributions \widehat{S}_5 and \widehat{S}_{50} obtained using an implementation of the Nelder–Mead method in MATLABTM (Lagarias *et al.* (1998)), along with the images generated by the STORMFLOW model using these joint probability distributions as first input.

The last step is to define an interpolation function:

$$s : [0; 1] \rightarrow \mathcal{S} \quad (10)$$

so that

$$s(t) = (1-t) \times \widehat{S}_5 + t \times \widehat{S}_{50}. \quad (11)$$

This function allows probability density surfaces to be defined from an interpolation parameter $t \in [0; 1]$. Given the estimators for the two extreme flow rates (\widehat{S}_5 and \widehat{S}_{50}), the values of μ and t are optimized to minimize the cost function $M_v(\mu, s(t))$. Note that by definition, $s(0) = \widehat{S}_5$ and $s(1) = \widehat{S}_{50}$.

A model (Fig. 8) can then be built using parameter pairs $(\mu, t) \in \mathbb{R}_+ \times [0; 1]$ derived from a first iteration

of the STORMFLOW model using $(\mu, s(t))$ as input. This second iteration is much easier to optimize, and since it uses only two of the four sets of images ($v = 5$ and $v = 50$), the two remaining sets ($v = 10$ and $v = 30$) can be used for validation.

RESULTS

VALIDATION AND SIMULATED DATA

The optimal parameter pairs (μ, t) can be estimated by minimizing the cost function $M_v(\mu, s(t))$ for $v \in \{5, 10, 30, 50\}$. A sufficient number of images can then be generated to determine all the relevant characteristics, such as the global area \mathcal{A} , perimeter \mathcal{P} or Euler-Poincaré number χ . It is also possible to estimate the mean elongation ratio $\mathbb{E}(e)$ of the particles, the mean major axis $\mathbb{E}(M)$, the equivalent radius $\mathbb{E}(R_m)$, the covariance, the particle size distribution (PSD), and the cumulative PSD (Fig. 9). The latter can be calculated by generating m images and applying morphological openings (with reconstruction) to every single observation with the use of homothetic structuring elements, i.e. disks, of increasing radius: $1 < \dots < m$.

The covariance C_{Ξ} of a random closed set Ξ in \mathbb{R}^2 is defined as the following probability:

$$\forall \vec{r} \in \mathbb{R}^2, C_{\Xi}(\vec{r}) = \mathbb{P}(z \in \Xi, z + \vec{r} \in \Xi), \quad (12)$$

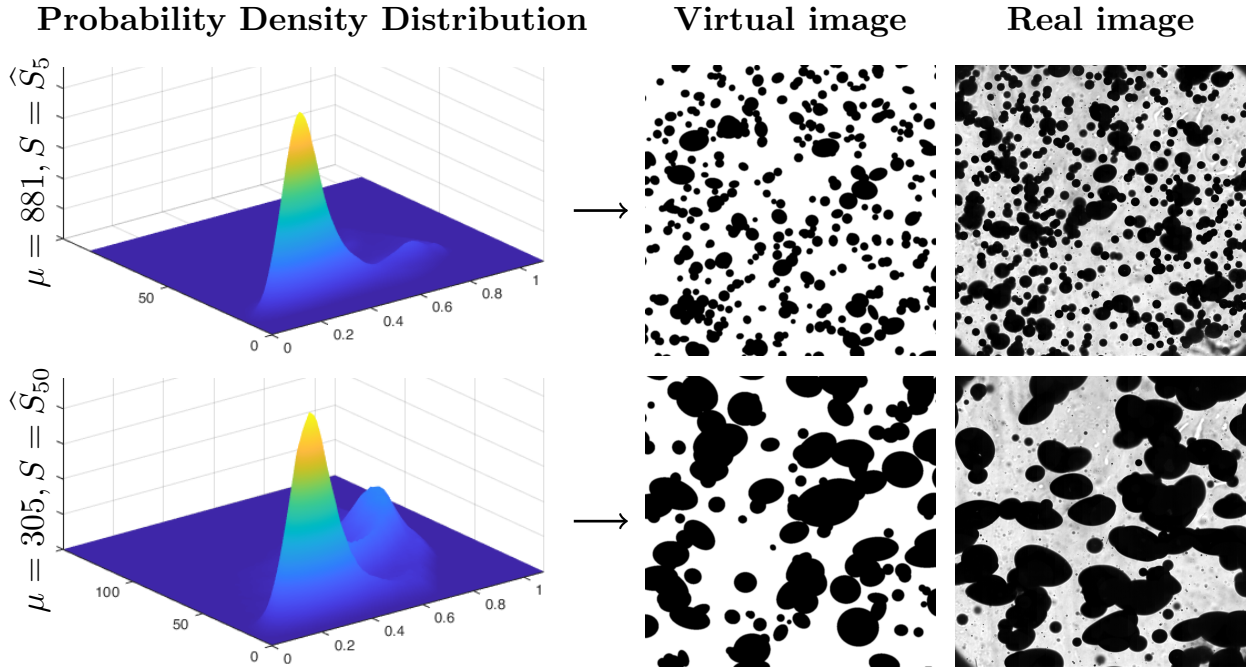


Fig. 7: Probability density functions reconstructed with the data obtained from the ellipse detection algorithm for $\nu = 5$ (top) and $\nu = 50$ (bottom) and comparison of the virtual and real images.

where $z \in \mathbb{R}^2$ and $\Xi_{\vec{r}}$ is the translation of Ξ by the vector \vec{r} . Thus, for a homogeneous isotropic random set, as is the case with the STORMFLOW model, C_{Ξ} only depends on the distance $r = \|\vec{r}\|$. The covariance is known to be rich in morphological information (Serra (1988)), for instance:

$$C_{\Xi}(0) = \mathcal{A}(\Xi), \quad \text{and} \quad \lim_{r \rightarrow +\infty} C_{\Xi}(r) = \mathcal{A}(\Xi)^2. \quad (13)$$

Although the STORMFLOW model can be used to calculate various average grain properties, its main function is to provide precise estimates of many of the characteristics of digital images. Indeed, the joint and marginal distributions of every single random variable including the orientation, major and minor axes, and the elongation ratio of the particles are fully described.

Fig. 10 compares the probability density distributions of the major axis and elongation ratio produced by the STORMFLOW model for $\nu = 10$ to the ones obtained from the ellipse detection algorithm applied to real images. The agreement is clearly very good, demonstrating the model's ability to fully and accurately characterize the images.

Flow rate $\nu = 5 \text{ L.h}^{-1}$						
Properties	\mathcal{A}	\mathcal{P}	χ	$\mathbb{E}(e)$	$\mathbb{E}(R_m)$	$\mathbb{E}(M)$
Real images	0.38	21.86	46.38	0.38	17.29	18.64
Virtual images	0.38	21.49	40.66	0.38	17.58	18.64
Error (%)	$\sim 0\%$	-2%	-12%	$\sim 0\%$	$+2\%$	$\sim 0\%$

Flow rate $\nu = 10 \text{ L.h}^{-1}$						
Properties	\mathcal{A}	\mathcal{P}	χ	$\mathbb{E}(e)$	$\mathbb{E}(R_m)$	$\mathbb{E}(M)$
Real images	0.41	18.55	68.44	0.37	19.98	21.52
Virtual images	0.4	18.36	32.44	0.39	19.57	21.12
Error (%)	-1%	-1%	-53%	$+5\%$	-2%	-2%
Flow rate $\nu = 30 \text{ L.h}^{-1}$						
Properties	\mathcal{A}	\mathcal{P}	χ	$\mathbb{E}(e)$	$\mathbb{E}(R_m)$	$\mathbb{E}(M)$
Real images	0.42	12.87	18.19	0.46	27.57	31.42
Virtual images	0.41	13.35	17.69	0.46	27.84	31.42
Error (%)	-2%	$+4\%$	-3%	$\sim 0\%$	$+1\%$	$\sim 0\%$
Flow rate $\nu = 50 \text{ L.h}^{-1}$						
Properties	\mathcal{A}	\mathcal{P}	χ	$\mathbb{E}(e)$	$\mathbb{E}(R_m)$	$\mathbb{E}(M)$
Real images	0.5	12.15	30.87	0.49	33.94	39.41
Virtual images	0.5	12.72	11.19	0.5	34.03	39.16
Error (%)	$\sim 0\%$	$+5\%$	-64%	$+1\%$	$\sim 0\%$	-1%

Table 2: Model performance. The global area \mathcal{A} , perimeter \mathcal{P} and Euler-Poincaré number χ were calculated directly from the real and virtual images. The expected values of the elongation ratio, mean equivalent radius and major axis were given by the geometrical model for the virtual images and calculated for the real images using the ellipse recognition algorithm.

DISCUSSION

As shown in Table 2 and Fig. 9, the values obtained are in good agreement with those determined directly on real images for all the characteristics except the Euler-Poincaré number. The larger errors for this property can be explained by the limitations of the STORMFLOW model, especially the assumption

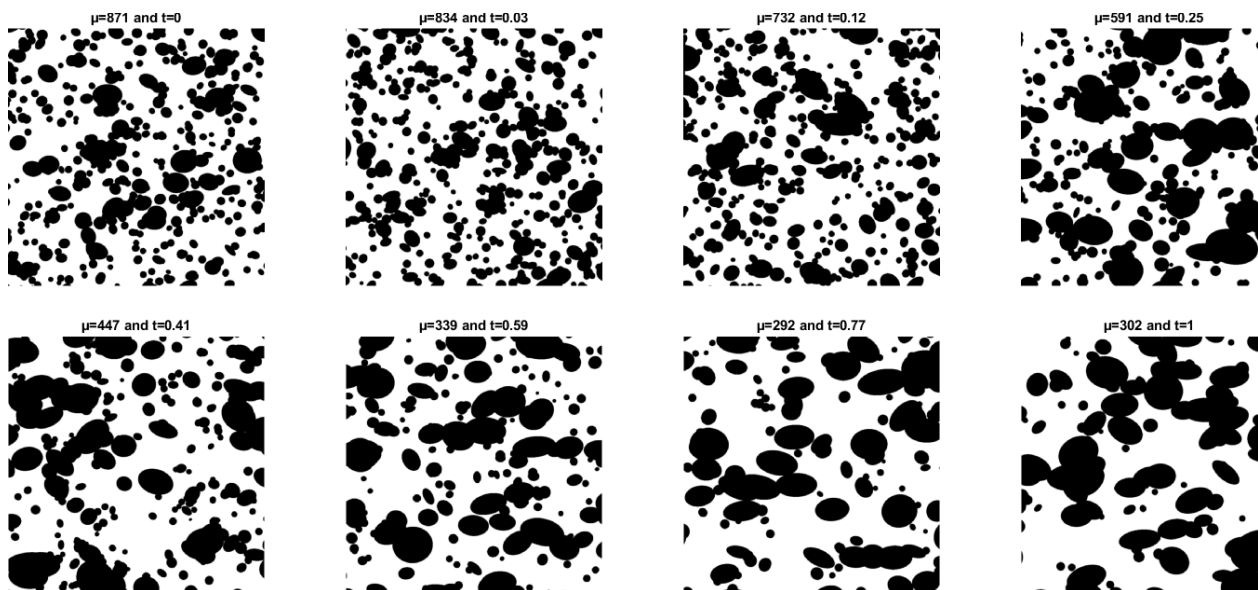


Fig. 8: Examples of realizations of the STORMFLOW model for different parameter pairs (μ, t) . The values of μ for the extreme cases $t = 0$ and $t = 1$, are obtained by minimizing the cost function $M_v(\mu, s(t))$ defined by (7), and the intermediate values are obtained by interpolation.

that the particles have a uniform spatial distribution. Moreover, the error on the fit of the orientation distribution of the grains affects the estimated number of connected components and holes, and is thereby magnified in the Euler-Poincaré number, which is the difference of the two. The Euler-Poincaré numbers of the virtual images are too low, which means that there are too many holes and/or too few connected components. In this respect, a more regular spatial distribution of germs might be helpful since it would increase the number of connected components. Note also that working close to the percolation threshold may have caused instabilities.

CONCLUSION AND PROSPECTS

CONCLUSION

This paper shows that local geometric measures can be used to infer the parameters of a stochastic geometrical model for multiphase flow images. Overall, minimizing a cost function based on the likelihood of a probability density function of local measurements fitting a sample of real data seems to provide good estimates of the model's parameters. The results obtained for the global characteristics of the images, the characteristics of typical grains, and for other indicators such as the covariance function and the cumulative PSD are clear evidence of the relevance of this approach.

The relative errors are almost always less than 2% when comparing the area and perimeter of the model images with those of real ones. The relative errors on the expected elongation ratio, mean equivalent radius and major axis are also almost all under 2% when compared to the values obtained with a deterministic ellipse recognition algorithm. The values obtained for the Euler-Poincaré number are not as accurate, possibly because the image sample size used was too small, because the spatial distribution of the particles in the images was unfavourable (too irregular), and/or because of numerical instability close to the percolation threshold.

PROSPECTS

Since the parameters of the STORMFLOW model were estimated using only local perimeter measurements, a logical next step would be to use the local area or a combination of the local area and perimeter (perhaps even the local Euler-Poincaré number) instead. Methods based on local measurements seem to be faster than those based on covariance or the particle size distribution (PSD). It would be interesting to quantify this gain in efficiency by benchmarking.

Several features of the STORMFLOW model can also be improved. First, improving the modeling of the spatial distribution of the germs and/or of the particle orientations would probably improve the accuracy of the Euler-Poincaré numbers obtained.

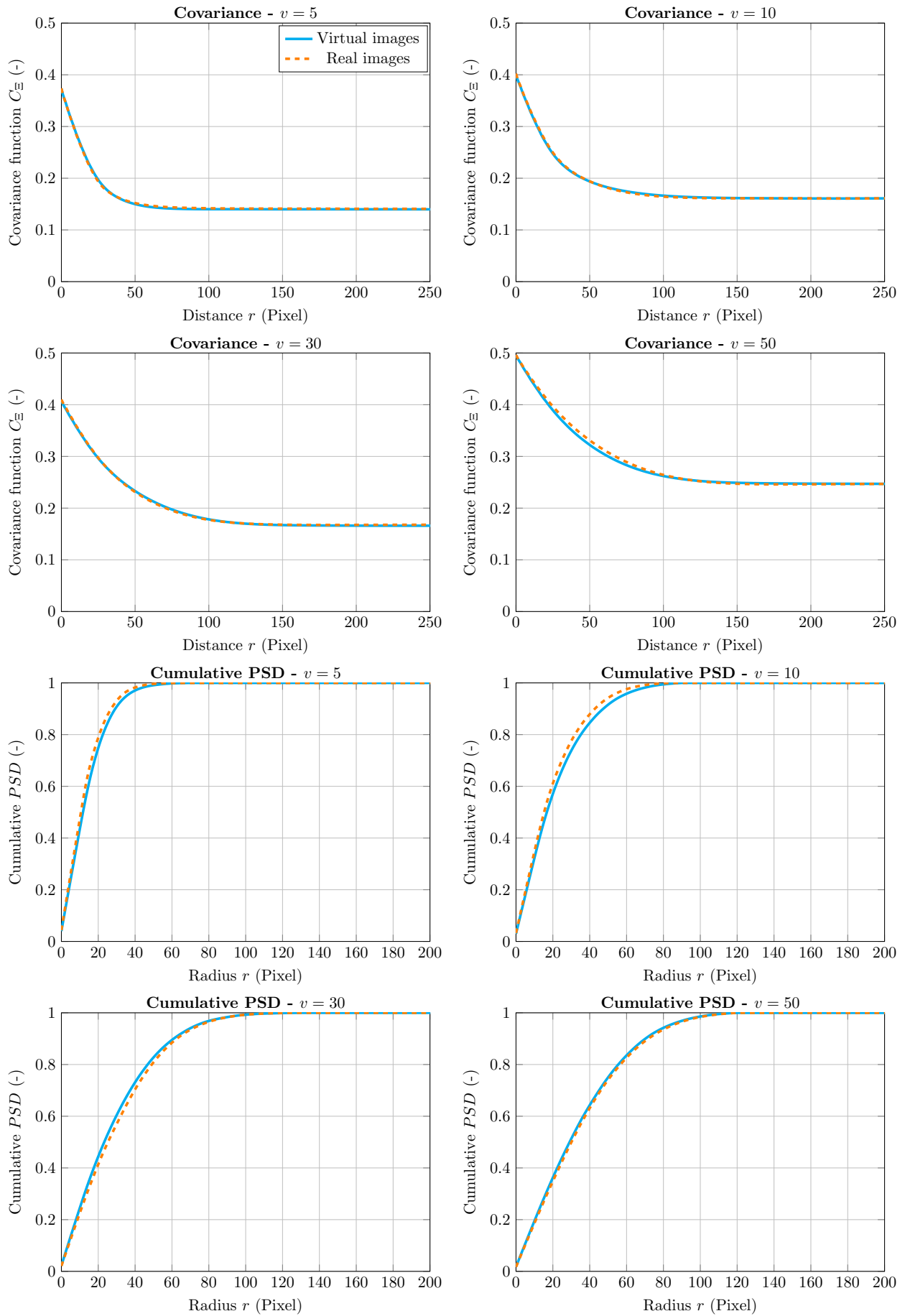


Fig. 9: Comparison of the covariance and cumulative particle size distribution (PSD) of real and virtual flow images at different flow rates v .

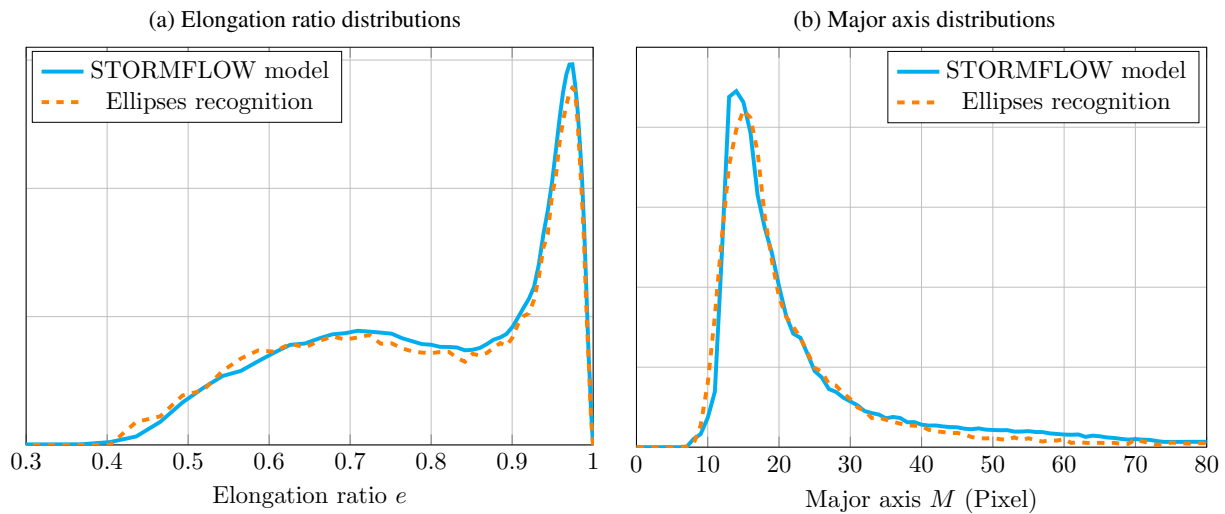


Fig. 10: Comparison of the probability density distributions produced by the STORMFLOW model with those obtained by analyzing real images with the ellipse recognition algorithm for (a) the elongation ratio and (b) the major axis of the particles, at a flow rate $v = 10 \text{ L.h}^{-1}$.

Second, while grains in STORMFLOW are assumed to be perfect ellipses, gas bubbles in fact become more and more irregular as the flow rate increases (Fig. 2c & 2d). Adding more complex shapes (caps, wobbling particles, etc.) by taking into account physical parameters (viscosity, density, etc.) in the STORMFLOW model would also no doubt improve its accuracy.

ACKNOWLEDGEMENTS

A special thanks to Mathieu De Langlard for helping with the set-up of the ellipse recognition algorithm and his insights on the matter.

REFERENCES

- Bai W., Deen N.G., Kuipers J.A.M. (2012) Numerical investigation of gas holdup and phase mixing in bubble column reactors. *IND ENG CHEM RES* 51(14):1949–61.
- Bello R. Ade, Robinson Campbell W., Moo-Young Murray (1985). Gas holdup and overall volumetric oxygen transfer coefficient in airlift contactors. *BIOTECHNOL BIOENG* 27(3):369–81
- Benassi C., Bianchi G., D’Ercole G. (2010). Covariogram of non-convex sets. *MATHEMATIKA* 56(2):267–84.
- Bettaieb A., Filali N., Filali T., Ben Aissia H. (2020). An efficient algorithm for overlapping bubbles segmentation. *Computer Optics* 44(3):363–74.
- Buffo A., Vanni V., Marchisio, D. (2017) Simulation of reacting gas-liquid bubbly flow with CFD and PBM: validation with experiments. *APPL MATH MODEL* 44:43–60.
- Cerqueira R.F.L., Paladino E.E., Ynumaru B.K., Maliska C.R. (2018). Image processing techniques for the measurement of two-phase bubbly pipe flows using particle image and tracking velocimetry (PIV/PTV). *CHEM ENG SCI* 189:1–23.
- Cerqueira R.F.L., Paladino E.E. (2021). Development of a deep learning-based image processing technique for bubble pattern recognition and shape reconstruction in dense bubbly flows. *CHEM ENG SCI* 230:116–63.
- Chiu S.N., Stoyan D., Kendall W.S., Mecke J. (2013). *Stochastic geometry and its applications*, 3rd Ed. Chichester: Wiley.
- De Langlard M., Al-Saddik H., Charton S., Debayle J., Lamadie F. (2018). An efficiency improved recognition algorithm for highly overlapping ellipses: Application to dense bubbly flows. *PATTERN RECOGN LETT* 101:88–95.
- De Langlard M., Lamadie F., Charton S., Debayle J. (2018). A 3D stochastic model for geometrical characterization of particles in two-phase flow applications. *IMAGE ANAL STEREOLOG* 37(3):233–47.
- De Langlard M., Lamadie F., Charton S., Debayle J. (2021). Bayesian Inference of a Parametric

- Random Spheroid from its Orthogonal Projections. *METHODOL COMPUT APPL* 23.
- De Langlard M. (2019). La géométrie aléatoire pour la caractérisation de populations denses de particules: application aux écoulements diphasiques. Autre. Université de Lyon. Français. NNT: 2019LYSEM001.
- Dereudre D., Lavancier F., Helisova K.S. (2014). Estimation of the intensity parameter of the germ-grain Quermass-interaction model when the number of germs is not observed. *SCAND J STAT* 41:809–29.
- Diggle P.J. (1981). Binary mosaics and the spatial pattern of heather. *BIOMETRICS* 37:531–9.
- Epanechnikov V.A. (1969). Non-Parametric Estimation of a Multivariate Probability Density. *THEOR PROBAB APPL+*, 1969, Vol. 14, No. 1 : pp. 153–58.
- Eremina T., Debayle J., Gruy F., Pinoli J.C. (2021). Numerical estimation of local Minkowski measures on Boolean models, 2020 10th International Symposium on Signal, Image, Video and Communications (ISIVC), 2021, pp. 1-5, doi: 10.1109/ISIVC49222.2021.9487537.
- Eremina T., Debayle J., Gruy F., Pinoli J.C. (2021). Local Measures Distribution for the Estimation of the Elongation Ratio of the Typical Grain in Homogeneous Boolean Models. *IMAGE ANAL STEREOLOG* 40(2):95–103.
- Haas T., Schubert C., Eickhoff M., Pfeifer H. (2020). BubCNN: Bubble detection using Faster RCNN and shape regression network. *CHEM ENG SCI* 216:1154–67.
- Kracht W., Emery X., Paredes C. (2013). A stochastic approach for measuring bubble size distribution via image analysis. *INT J MINER PROCESS* 121:6–11.
- Kurz G., Gilitschenski I., Siegwart R., Hanebeck U.D. (2016). Methods for Deterministic Approximation of Circular Densities. *Journal of Advances in Information Fusion* 11:138–56.
- Lagarias J.C., Reeds J.A., Wright H., Wright P.E. (1998). Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM J OPTIMIZ* 9(1):112–47.
- Lefèvre S. (2009). Beyond Morphological Size-distribution 18. *J ELECTRON IMAGING*.
- Mardia K.V., Jupp P.E. (1999). *Directional Statistics*. Wiley.
- Matheron G. (1972). *Ensembles aléatoires et géométrie intégrale*. Les cahiers du centre de morphologie mathématiques de Fontainebleau.
- Molchanov I.S. (1997). *Statistics of the Boolean Model for Practitioners and Mathematicians*. Wiley.
- Paolinelli L.D., Rashedi A., Yao J. (2018). Characterization of droplet sizes in large scale oil–water flow downstream from a globe valve. *INT J MULTIPHAS FLOW* 99:132–50.
- Rataj J. (2004). On set covariance and three-point test sets. *CZECH MATH J* 54.1 : 205-14.
- Schmidt T. (2007). *Coping with copulas. Copulas - From Theory to Application in Finance*.
- Serra J.A. (1988). *Image Analysis and Mathematical Morphology: Theoretical Advances*. Academic Press, Cornell University.
- Torisaki S., Miwa S. (2020). Robust bubble feature extraction in gas-liquid two-phase flow using object detection technique. *J NUCL SCI TECHNOL* 57(11):1231–44.
- Vinnett L., Yianatos J., Arismendi L., Waters K.E. (2020). Assessment of two automated image processing methods to estimate bubble size in industrial flotation machines. *MINER ENG* 159:106636.
- Zhou H., Niu X. (2020). An image processing algorithm for the measurement of multiphase bubbly flow using predictor-corrector method. *INT J MULTIPHAS FLOW* 128:103277.