

Titre: Identification des configurations d'un écoulement diphasique dans un faisceau de tubes à l'aide d'une machine à vecteurs de support
Title:

Auteur: Inès Benito
Author:

Date: 2016

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Benito, I. (2016). Identification des configurations d'un écoulement diphasique dans un faisceau de tubes à l'aide d'une machine à vecteurs de support [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/2057/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/2057/>
PolyPublie URL:

Directeurs de recherche: Njuki W. Mureithi
Advisors:

Programme: Génie mécanique
Program:

UNIVERSITÉ DE MONTRÉAL

IDENTIFICATION DES CONFIGURATIONS D'UN ÉCOULEMENT DIPHASIQUE
DANS UN FAISCEAU DE TUBES À L'AIDE D'UNE MACHINE À VECTEURS DE
SUPPORT

INÈS BENITO

DÉPARTEMENT DE GÉNIE MÉCANIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE MÉCANIQUE)

JANVIER 2016

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

IDENTIFICATION DES CONFIGURATIONS D'UN ÉCOULEMENT DIPHASIQUE
DANS UN FAISCEAU DE TUBES À L'AIDE D'UNE MACHINE À VECTEURS DE
SUPPORT

présenté par : BENITO Inès

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. BALAZINSKI Marek, Docteur ès sciences, président

M. MUREITHI Njuki William, Ph. D., membre et directeur de recherche

M. PETTIGREW Michel, M. Sc., membre

DÉDICACE

*À mon grand-père,
avec toute ma reconnaissance
pour son soutien dans mes études
et tout ce qu'il m'a apporté.*

REMERCIEMENTS

Je tiens à remercier tous ceux qui m'ont accompagnée dans cette maîtrise recherche au sein de la Chaire industrielle de recherche BWC/EACL/CRSNG en Interaction Fluide-Structure de l'École Polytechnique de Montréal.

En premier lieu, je voudrais remercier chaleureusement mon directeur de recherche, Njuki W. Mureithi, pour sa grande disponibilité, son attention et ses conseils, qui m'ont permis de mener à bien ce projet.

Je remercie tout spécialement François de Kerret pour sa collaboration fructueuse, notamment en ce qui concerne l'utilisation de notre boucle d'essai commune, et nos multiples échanges qui ont bien contribué à la progression de ma maîtrise.

Je tiens à remercier particulièrement Benedict Besner pour sa disponibilité permanente et son expertise technique, qui m'ont permis d'avancer tout au long de ce projet.

Je remercie aussi Cédric Béguin pour les conseils avisés et l'expérience qu'il a volontiers partagés avec moi.

Enfin, je remercie chaque membre de la Chaire et les étudiants gradués ou doctorants, qui m'ont soutenue dans cette maîtrise, par leurs conseils et leurs encouragements, tout particulièrement Bastien Cucuel, ainsi que Stephen Olala, Hao Li, Arnaud Dubourg, Elsa Piollet et Anne Sophie De Almeida, et enfin mes proches et spécialement mes colocos,

ainsi que le Conseil de Recherches en Sciences Naturelles et en Génie du Canada pour son soutien financier au long de ce projet.

RÉSUMÉ

Les écoulements diphasiques sont présents dans de nombreux équipements industriels comme les échangeurs de chaleur. On s'intéresse ici en particulier aux générateurs de vapeur, où il existe un écoulement transverse au niveau de la partie supérieure des faisceaux de tubes en U. Il est important de connaître les configurations que l'écoulement peut prendre à cet endroit, car dans certaines conditions, l'interaction fluide-structure peut engendrer des vibrations des tubes et de leurs supports, et donc une usure par frottement, qui peut aller jusqu'à causer des défaillances.

Alors que les configurations d'écoulement sont bien connues pour les écoulements diphasiques à l'intérieur de tubes, en comparaison, peu d'études ont porté sur l'identification des configurations en faisceaux de tubes. Parmi les études réalisées, un certain nombre s'appuie seulement ou majoritairement sur des observations visuelles, qui sont source de subjectivité dans l'identification de ces configurations. Par ailleurs, le développement récent des connaissances dans le domaine du *machine learning* (apprentissage automatique) a ouvert de vastes possibilités pour des applications comme la classification de données et l'identification de configurations. Des « machines » d'apprentissage supervisé comme les réseaux de neurones ou les machines à vecteurs de support (SVM) ont été utilisées avec succès pour assister l'identification des configurations d'écoulement dans des tubes, permettant ainsi d'obtenir des cartes d'écoulement basées sur des classifications plus objectives. Cependant, cette approche ne semble pas encore avoir été mise à profit pour les faisceaux de tubes.

L'objectif de cette étude est donc de développer une méthode d'identification des configurations d'écoulement diphasique dans les faisceaux de tubes, basée sur l'analyse de signaux obtenus par sonde optique, et l'utilisation d'une machine à vecteurs de support pour en obtenir une classification plus objective.

La section d'essai utilisée est une conduite transparente à section carrée dont les parois sont constituées de demi-tubes de façon à représenter le chemin d'écoulement d'un faisceau de tubes en configuration triangulaire tournée. Les demi-tubes ont un diamètre de 38 mm et le rapport entre le pas du faisceau et ce diamètre est de 1.5. 110 points de mesure ont été réalisés pour des titres volumiques variant entre 12% et 97%, et des vitesses superficielles de liquide comprises entre 0.1 m/s et 0.75 m/s.

Pour chaque point de mesure, des photographies par caméra haute vitesse ont été prises et le signal de la sonde optique, qui permet de détecter les changements de phase, a été enregistré. L'analyse des photos et leur montage en vidéos a permis dans un premier temps d'identifier visuellement les configurations observées, pour des cas où elles étaient facilement différenciables. On a ainsi défini les configurations à bulles, intermittente, et semi-annulaire. L'analyse des signaux de sonde optique, a confirmé dans un deuxième temps l'emplacement approximatif des transitions, et la nécessité d'une méthode de classification plus objective.

Afin d'obtenir des données quantitatives reliées intrinsèquement aux configurations, et permettant de les différencier, les fonctions de densité de probabilité (PDF) du taux de vide et les valeurs des moments statistiques du signal donné par la sonde ont été calculées. Les moments statistiques correspondent à l'allure des graphes de PDF, mais présentent l'avantage de nécessiter moins de traitement préalable du signal, ne dépendent d'aucun autre paramètre, et permettent de réduire le nombre de grandeurs d'entrée de la SVM. L'analyse finale a donc été basée sur les moments statistiques.

Le modèle de classification des configurations d'écoulement est basé sur un apprentissage supervisé et la sélection de données d'entraînement parmi des points de configuration connue. La capacité de généralisation de la SVM a été confirmée et a permis de prédire la configuration des données plus difficiles à classer.

On a ainsi pu obtenir une classification de tous les points de mesure par la SVM basée sur des paramètres optimisés, et en déduire des frontières plus objectives entre les trois configurations d'écoulement observées visuellement. La carte d'écoulement obtenue a finalement été comparée avec des cartes de la littérature pour des écoulements internes et en faisceaux de tubes.

ABSTRACT

Two-phase flows are commonly found in a variety of industrial equipment such as heat exchangers. In steam generators, two-phase cross-flow occurs in the upper section of the U-bend tube bundles. Depending on flow velocity and on the flow pattern, fluid-structure interaction can lead to tube vibrations impacting with the support plates, thus creating fretting wear and even tube failures.

Two-phase flow inside tubes are well understood, whereas few studies have been conducted to identify flow patterns inside tube arrays. Some studies rely only or mainly on visual observations and are therefore quite subjective. Recent insights in the machine learning field provide new opportunities for data classification and pattern identification. Supervised learning machines such as neural networks or support vector machines (SVM) have been successfully used to help in identifying flow patterns inside tubes, leading to more objective flow pattern maps. However, supervised learning has not yet been used in tube bundles.

Hence, this study aims at developing a two-phase flow-pattern identification method for tube bundles, based on optical probe signal analysis and the use of a support vector machine to obtain a more objective classification.

Experiments have been conducted in a transparent square section with half-tubes at the inner walls to simulate the open lane within a rotated triangular array. Half-tube diameter is 38 mm and the pitch-to-diameter ratio is 1.5. 110 conditions have been measured for volumetric qualities varying from 12% to 97% and superficial liquid velocities from 0.1 m/s to 0.75 m/s.

For each condition, photographs were taken by a high speed camera and the optical probe signals (*ie* phase changes) were acquired. Analysis of the images and videos of the flow allowed first to visually identify configurations when the flow pattern was clear. Bubbly, intermittent and semi-annular patterns have thus been defined. Analysis of the probe signals then confirmed the transition locations and the need for a more objective method to classify patterns.

To obtain quantitative data related to flow patterns, and allow their differentiation, the probability density functions (PDF) of the void fraction and statistical moments of the signals were calculated. In the final analysis, statistical moments were selected. The moments express the trend

of the PDF graphs, but need less signal processing, intermediate parameters, and minimize the number of features in the SVM.

The flow pattern classification model is based on a supervised learning approach and pre-selected training data. The generalization ability of the resulting SVM was confirmed using hard to classify unknown data.

The SVM with optimal parameters thus gave a predicted classification for all conditions, allowing an objective definition of boundaries between the three observed flow patterns. Finally, the map resulting from the classification was compared to those from literature for tubes and tube arrays.

TABLE DES MATIÈRES

| | |
|--|------|
| DÉDICACE..... | III |
| REMERCIEMENTS | IV |
| RÉSUMÉ..... | V |
| ABSTRACT | VII |
| TABLE DES MATIÈRES..... | IX |
| LISTE DES TABLEAUX | XII |
| LISTE DES FIGURES | XIII |
| LISTE DES SIGLES ET ABRÉVIATIONS..... | XVII |
| LISTE DES ANNEXES | XIX |
| CHAPITRE 1 INTRODUCTION..... | 1 |
| 1.1 Contexte de l'étude..... | 1 |
| 1.2 Problématique..... | 3 |
| 1.3 Objectifs de la recherche | 5 |
| 1.3.1 Objectif général | 5 |
| 1.3.2 Objectifs spécifiques | 5 |
| CHAPITRE 2 REVUE DE LITTÉRATURE | 7 |
| 2.1 Modèles d'analyse des écoulements diphasiques..... | 7 |
| 2.1.1 Définitions et modèles d'analyse | 7 |
| 2.1.2 Modèle homogène | 10 |
| 2.2 Configurations d'écoulement diphasique..... | 11 |
| 2.2.1 Configurations d'un écoulement vertical ascendant dans un tube | 11 |
| 2.2.2 Configurations d'écoulement en faisceau de tubes | 16 |
| CHAPITRE 3 MÉTHODE EXPÉRIMENTALE | 28 |

| | | |
|---|--|----|
| 3.1 | Montage expérimental..... | 28 |
| 3.1.1 | Boucle d'essai | 28 |
| 3.1.2 | Section d'essai..... | 30 |
| 3.1.3 | Points de mesure et paramètres mesurés | 35 |
| 3.1.4 | Prise d'images par caméra haute vitesse | 39 |
| 3.1.5 | Ajustement des débits..... | 39 |
| 3.2 | Instrument de mesure principal : sonde optique..... | 40 |
| CHAPITRE 4 RÉSULTATS BRUTS : FORME DE L'ÉCOULEMENT ET DONNÉES DE SONDE OPTIQUE | | 44 |
| 4.1 | Images de l'écoulement et formes observées | 44 |
| 4.2 | Définition des configurations observées | 47 |
| 4.3 | Signaux bruts de sonde optique..... | 55 |
| CHAPITRE 5 ANALYSE DES DONNÉES EXPÉRIMENTALES..... | | 60 |
| 5.1 | Obtention des PDF du taux de vide..... | 60 |
| 5.1.1 | Filtrage du signal de la sonde | 60 |
| 5.1.2 | Principe de calcul des PDF..... | 62 |
| 5.1.3 | Choix de l'intervalle de temps Δt | 66 |
| 5.1.4 | Choix de l'intervalle $\Delta \varepsilon$ | 70 |
| 5.2 | Calcul des moments statistiques des signaux de la sonde..... | 73 |
| 5.3 | Comparaison entre PDF et moments statistiques..... | 78 |
| CHAPITRE 6 UTILISATION D'UNE MACHINE À VECTEURS DE SUPPORT POUR OBTENIR UNE CARTE D'ÉCOULEMENT PLUS OBJECTIVE..... | | 82 |
| 6.1 | Machine learning..... | 82 |
| 6.1.1 | Principe général de l'apprentissage supervisé..... | 82 |
| 6.1.2 | Machine à Vecteurs de Support (SVM) | 83 |

| | | |
|--|---|-----|
| 6.2 | Application à l'identification des configurations d'écoulement | 88 |
| 6.2.1 | Choix des paramètres généraux de la SVM | 89 |
| 6.2.2 | Pré-classification des données..... | 93 |
| 6.2.3 | Implémentation d'une SVM..... | 99 |
| 6.3 | Analyse des classifications obtenues et choix final des paramètres..... | 110 |
| 6.3.1 | Classification obtenue pour $n_{feat}=4$ | 110 |
| 6.3.2 | Classification obtenue pour $n_{feat}=2$ | 111 |
| CHAPITRE 7 CARTE D'ÉCOULEMENT OBTENUE ET COMPARAISON AVEC LA LITTÉRATURE | | 115 |
| 7.1 | Carte d'écoulement obtenue..... | 115 |
| 7.2 | Comparaison avec la littérature..... | 116 |
| 7.2.1 | Comparaison avec les cartes obtenues en faisceaux de tubes | 116 |
| 7.2.2 | Comparaison avec la carte de Taitel et al..... | 121 |
| 7.2.3 | Synthèse des comparaisons | 123 |
| CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS | | 125 |
| BIBLIOGRAPHIE..... | | 128 |
| ANNEXES..... | | 132 |

LISTE DES TABLEAUX

| | |
|---|----|
| Tableau 2.1 : Caractéristiques des sections d’essai correspondant aux cartes de la Figure 2.16... | 27 |
| Tableau 3.1 : Éléments numérotés de la boucle d’essai (Figure 3.1)..... | 29 |
| Tableau 3.2 : Éléments de la section d’essai..... | 31 |
| Tableau 3.3 : Couples (β , j_i) choisis et numéros des points correspondants | 37 |
| Tableau 4.1 : Caractéristiques visuelles de la configuration à bulles..... | 49 |
| Tableau 4.2 : Caractéristiques visuelles de la configuration intermittente | 50 |
| Tableau 4.3 : Caractéristiques visuelles de la configuration semi-annulaire | 54 |
| Tableau 6.1 : Valeurs initiales des moments (minimales et maximales) | 90 |
| Tableau 6.2 : Récapitulatif des paramètres de la SVM..... | 92 |
| Tableau 6.3 : Conditions respectées par les points de référence..... | 94 |
| Tableau 6.4 : Types de profils de moments statistiques..... | 96 |
| Tableau 6.5 : Points de référence pour les trois configurations | 98 |

LISTE DES FIGURES

| | |
|--|----|
| Figure 1.1 : Générateur de vapeur et partie transverse du faisceau de tubes en U..... | 1 |
| Figure 2.1 : Configurations d'un écoulement vertical ascendant dans un tube d'après Taitel et al. (1980) | 11 |
| Figure 2.2 : Carte des configurations d'un écoulement vertical ascendant dans un tube d'après Taitel et al. (1980) | 13 |
| Figure 2.3 : Configurations d'écoulement d'après Giraudeau (2011) | 15 |
| Figure 2.4 : Photographies de l'écoulement d'après Rosa et al. (2010) | 15 |
| Figure 2.5 : Configurations d'un faisceau triangulaire | 16 |
| Figure 2.6 : Carte de Grant et Murray (1972) citée par Ulbrich et Mewes (1994) | 17 |
| Figure 2.7 : Schémas de certaines configurations observées par Kondo et Nakajima (1980) | 18 |
| Figure 2.8 : Carte obtenue par Ulbrich et Mewes (1994)..... | 19 |
| Figure 2.9 : Carte d'écoulement obtenue par Xu et al. (1998)..... | 21 |
| Figure 2.10 : Schémas des configurations décrites par Xu et al. (1998)..... | 21 |
| Figure 2.11 : Cartes des configurations de Noghrehkar et al. (1999) | 22 |
| Figure 2.12 : Points de mesure de Chung et Chu (2006) sur la carte de Ulbrich et Mewes (1994) | 23 |
| Figure 2.13 : Photographies caractéristiques obtenues par Hong et Liu (2010) | 24 |
| Figure 2.14 : Carte de Hong et Liu (2010) et comparaison avec la littérature pour $P/D=1.3$ | 25 |
| Figure 2.15 : Carte de Hong et Liu (2010) et comparaison avec la littérature pour $P/D=1.8$ | 25 |
| Figure 2.16 : Comparaison par Ribatski et Thome (2007)..... | 26 |
| Figure 3.1 : Configuration initiale de la boucle d'essai | 29 |
| Figure 3.2 : Caméra et projecteurs | 30 |
| Figure 3.3 : Fixation de la section d'essai (support inférieur) | 32 |

| | |
|---|----|
| Figure 3.4 : Chemin d'écoulement d'un faisceau de tubes triangulaire tourné et portion de la section d'essai..... | 33 |
| Figure 3.5 : Champ des vitesses moyennes de la phase gazeuse | 34 |
| Figure 3.6 : Carte de Noghrehkar et al. pour le faisceau triangulaire normal..... | 36 |
| Figure 3.7 : Position des points de mesure effectués | 37 |
| Figure 3.8 : Position des points de mesure par rapport aux frontières de Noghrehkar et al. | 38 |
| Figure 3.9 : Principe de fonctionnement de la fibre optique :..... | 40 |
| Figure 3.10 : Exemple de signal échelon donné par la sonde optique | 41 |
| Figure 3.11 : Correspondance entre l'écoulement diphasique et le signal donné par la sonde optique | 42 |
| Figure 4.1 : Zone observée avec la caméra | 45 |
| Figure 4.2 : Exemples d'images obtenues à différents instants et β , pour $j_l=0.3$ m/s | 46 |
| Figure 4.3 : Exemples d'images obtenues à différents instants pour $j_l=0.3$ m/s et $\beta=96\%$ | 47 |
| Figure 4.4 : Configurations intermittente-dispersée (ID) et annulaire dispersée (AD)..... | 51 |
| Figure 4.5 : Exemple typique de la configuration semi-annulaire | 53 |
| Figure 4.6 : Exemple d'évolution temporelle du signal brut..... | 55 |
| Figure 4.7 : Évolution du tracé du signal brut sur 1s, pour $j_l=0.3$ m/s et différentes valeurs de β | 56 |
| Figure 4.8 : Longueur L_0 | 57 |
| Figure 4.9 : Évolution du tracé du signal brut sur une durée variable= L_0 / j pour $j_l=0.3$ m/s et différentes valeurs de β | 58 |
| Figure 5.1 : Différences entre le signal brut (ligne claire) et le signal filtré (échelon) | 61 |
| Figure 5.2 : Évolution temporelle du signal brut de la sonde ($\beta=12\%$, $j_l=0.3$ m/s) | 61 |
| Figure 5.3 : Évolution temporelle du signal filtré correspondant ($\beta=12\%$, $j_l=0.3$ m/s) | 62 |
| Figure 5.4 : Division du signal temporel en intervalles Δt | 62 |

| | |
|--|----|
| Figure 5.5 : Calcul du taux de vide moyen sur un intervalle Δt | 63 |
| Figure 5.6 : Exemple d'évolution du taux de vide calculé sur chaque intervalle Δt , sur toute la durée d'acquisition | 64 |
| Figure 5.7 : Exemple de graphe de PDF | 65 |
| Figure 5.8 : Illustration du calcul de la PDF associée à un $\Delta \varepsilon$ donné | 65 |
| Figure 5.9 : Illustrations de mauvais choix de Δt pour un même point de mesure | 66 |
| Figure 5.10 : Exemple de visualisation d'un Δt approprié..... | 68 |
| Figure 5.11 : Visualisation de $L=5 L_c$ sur les images de l'écoulement | 69 |
| Figure 5.12 : Comparaison de différents choix pour $\Delta \varepsilon$ | 71 |
| Figure 5.13 : Exemples de tracés finaux des PDF du taux de vide, pour $j_l=0.3$ m/s et différentes valeurs de β | 72 |
| Figure 5.14 : Asymétrie d'une courbe..... | 74 |
| Figure 5.15 : Comparaison PDF/moments statistiques pour $\beta=12\%$, $j_l=0.6$ m/s | 75 |
| Figure 5.16 : Comparaison PDF/moments statistiques pour $\beta=55\%$, $j_l=0.5$ m/s | 76 |
| Figure 5.17 : Comparaison PDF/moments statistique pour $\beta=97\%$, $j_l=0.15$ m/s..... | 76 |
| Figure 5.18 : Exemples de tracés finaux des 4 premiers moments statistiques du signal brut, pour $j_l=0.3$ m/s et différentes valeurs de β | 77 |
| Figure 5.19 : Différences entre le signal brut (ligne claire) et le signal filtré (échelon) | 78 |
| Figure 5.20 : Comparaison de l'asymétrie pour le signal brut (foncé) et le signal filtré (clair)..... | 79 |
| Figure 5.21 : Comparaison du kurtosis pour le signal brut (foncé) et le signal filtré (clair)..... | 80 |
| Figure 6.1 : Deux exemples d'hyperplans séparateurs, pas optimaux | 84 |
| Figure 6.2 : Hyperplan optimal H définissant une marge maximale | 84 |
| Figure 6.3 : Définition de l'hyperplan..... | 85 |
| Figure 6.4 : Exemple de répartition en <i>training</i> et <i>test sets</i> , et valeurs des composantes (<i>features</i>) correspondantes | 91 |

| | |
|--|-----|
| Figure 6.5 : Ensemble des profils de moments statistiques et regroupements manuels en trois groupes | 97 |
| Figure 6.6 : Répartition aléatoire en <i>training</i> et <i>test sets</i> , exécution 1 | 100 |
| Figure 6.7 : Illustration de l' <i>overfitting</i> dans l'espace de départ..... | 102 |
| Figure 6.8 : Illustration de l' <i>underfitting</i> dans l'espace de départ | 102 |
| Figure 6.9 : Exemple de grille de CV obtenue pour l'exécution 1..... | 104 |
| Figure 6.10 : Illustration du choix de C et γ pour une exécution donnée (#1)..... | 106 |
| Figure 6.11 : Prédiction des configurations de l'ensemble des points pour l'exécution 1, $C=1, \gamma=16$ | 108 |
| Figure 6.12 : Classification finale moyenne sur les 100 premières exécutions ($n_{feat}=4$)..... | 109 |
| Figure 6.13 : Classification finale, moyenne sur les 100 premières exécutions ($n_{feat}=2$)..... | 112 |
| Figure 6.14 : Classification finale, moyenne sur les 100 premières exécutions telles que $\gamma_{max}=16$ ($n_{feat}=2$)..... | 114 |
| Figure 7.1 : Carte d'écoulement finale et frontières des configurations proposées | 115 |
| Figure 7.2 : Comparaison avec les frontières de Grant & Chisholm et Xu et al..... | 117 |
| Figure 7.3 : Comparaison avec les frontières de Ulbrich & Mewes, et Noghrehkar et al. | 118 |
| Figure 7.4 : Schéma de la configuration intermittente de Grant et Chisholm..... | 119 |
| Figure 7.5 : Schémas des configurations décrites par Xu et al. | 119 |
| Figure 7.6 : Comparaison des frontières proposées pour la configuration annulaire..... | 121 |
| Figure 7.7 : Comparaison avec la carte de Taitel et al. (frontières A, E et D) | 122 |

LISTE DES SIGLES ET ABRÉVIATIONS

| | |
|-----------------|---|
| A | Aire transverse de passage de l'écoulement entre les demi-tubes |
| C | Paramètre de la SVM : importance accordée aux erreurs de classification |
| CFM | Pieds cube par minute (<i>Cubic Feet per Minute</i>) |
| D | Diamètre des tubes ou demi-tubes |
| HV | Haute vitesse (caméra) |
| IA | Intelligence Artificielle |
| j | Vitesse du mélange homogène |
| j_g | Vitesse superficielle du gaz (air) |
| j_l | Vitesse superficielle du liquide (eau) |
| j_p | Vitesse inter-tubes |
| K | Kurtosis du signal de la sonde |
| L_c | Longueur caractéristique d'une bulle |
| L_p | Longueur de passage de l'écoulement entre deux tubes |
| L_{pr} | Profondeur de la section d'essai |
| LPM | Litres Par Minute |
| M | Moyenne du signal de la sonde |
| n_A | Indice de réfraction du milieu A |
| N_ε | Nombre total d'intervalles sur lesquels on calcule les PDF |
| P | Pas du faisceau de tubes |
| P/D | Ratio pas sur diamètre d'un faisceau de tubes |
| P_0 | Pression dans les conditions standard (NIST) |
| P_{rel} | Pression relative entre la pression à l'intérieur du faisceau et P_0 |
| PDF | Fonction de densité de probabilité (<i>Probability Density Function</i>) |
| PIV | Technique de mesure de champ de vitesse à l'aide de particules (<i>Particle Image Velocimetry</i>) |
| Q_d | Débit d'air affiché par le débitmètre |

| | |
|---------------------|---|
| Q_g | Débit volumique de la phase gazeuse |
| Q_l | Débit volumique de la phase liquide |
| R | Rayon des tubes ou demi-tubes |
| RPM | Nombre de tours par minute (<i>Revolutions per minute</i>) |
| S | Taux de glissement |
| SCFM | Pieds cube standards par minute (<i>Standard Cubic Feet per Minute</i>) |
| Skw | Asymétrie (<i>Skewness</i>) du signal de la sonde |
| Std | Ecart-type (<i>Standard deviation</i>) du signal de la sonde |
| SVM | Machine à vecteurs de support (<i>Support Vector Machine</i>) |
| T | Durée d'acquisition du signal de la sonde optique |
| T_0 | Température dans les conditions standard (NIST) |
| T_g | Temps passé par la sonde optique dans la phase gazeuse |
| T_l | Temps passé par la sonde optique dans la phase liquide |
| U_g | Vitesse de phase pour la phase gazeuse |
| U_{hom} | Vitesse du mélange homogène |
| U_l | Vitesse de phase pour la phase liquide |
| V_g | Volume occupé par la phase gazeuse |
| V_l | Volume occupé par la phase liquide |
| β | Titre volumique |
| γ | Paramètre de la fonction-noyau gaussienne de la SVM |
| $\Delta\varepsilon$ | Précision du tracé du graphe de PDF : « pas » de l'axe des abscisses |
| Δt | Intervalle de temps pour les PDF |
| ε | Taux de vide |

LISTE DES ANNEXES

| | |
|---|-----|
| Annexe A – Détails du montage expérimental..... | 132 |
| Annexe B – Tracé des PDF du taux de vide | 147 |
| Annexe C – Moments statistiques du signal brut de la sonde | 158 |
| Annexe D – Liens vers les ressources utilisées pour la SVM..... | 171 |
| Annexe E – Résultats obtenus avec les valeurs par défaut de C et γ : mauvaise classification finale | 172 |

CHAPITRE 1 INTRODUCTION

1.1 Contexte de l'étude

On retrouve des écoulements diphasiques dans de nombreux domaines industriels, parmi lesquels l'industrie pétrochimique et l'industrie nucléaire. Dans le cas de cette dernière, il s'agit par exemple des réacteurs à eau bouillante ou des générateurs de vapeur.

Dans cette étude, on s'intéresse au cas des écoulements transverses, qu'on retrouve notamment dans la partie supérieure d'un générateur de vapeur, où les milliers de tubes en U du faisceau sont disposés perpendiculairement à l'écoulement diphasique comme illustré à la Figure 1.1.

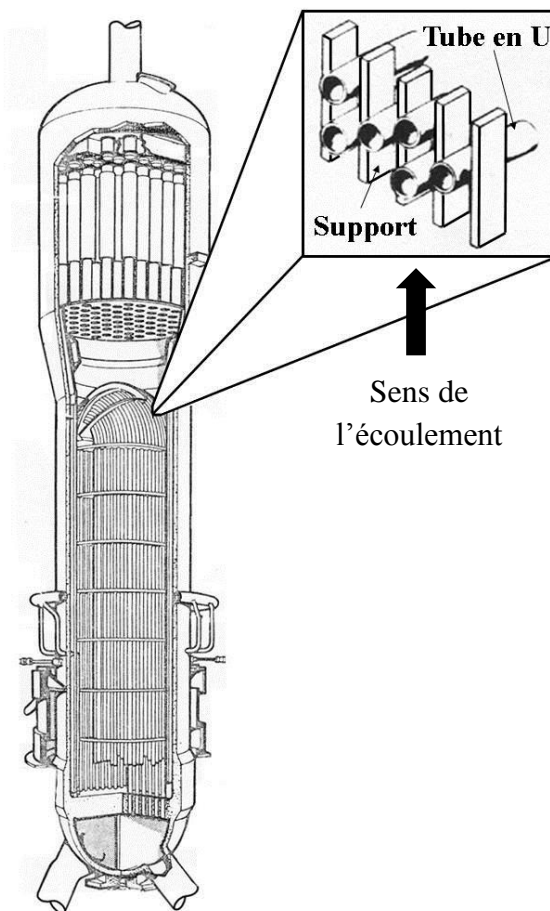


Figure 1.1 : Générateur de vapeur et partie transverse du faisceau de tubes en U

Ces tubes transportent le fluide caloporteur du circuit primaire et transfèrent ainsi la chaleur dégagée par la réaction nucléaire à l'eau du circuit secondaire, qui se transforme alors progressivement en vapeur. C'est donc un mélange diphasique composé d'eau et de vapeur qui constitue l'écoulement transverse étudié.

Comme dans la plupart des études, au sein de la Chaire industrielle de recherche BWC/EACL/CRSNG en Interaction Fluide-Structure de l'École Polytechnique de Montréal, cet écoulement diphasique est modélisé par un mélange air-eau, plus aisé à manipuler, et les faisceaux de tubes utilisés comportent plusieurs rangées de tubes horizontaux.

Interactions fluide-structure liées aux configurations d'écoulement

L'un des axes de recherche de la Chaire est de mieux connaître les phénomènes fluides qui se produisent dans les faisceaux de tubes, qui sont susceptibles de causer des vibrations néfastes des tubes. Les écoulements transverses peuvent en particulier générer des interactions fluide-structure auto-entretenues à cause de mécanismes comme l'instabilité fluidélastique (Pettigrew et al., 1998).

Ces interactions entre l'écoulement diphasique transverse et les tubes du faisceau entraînent ainsi un risque d'usure voire de rupture des tubes, par frottement au niveau de leurs supports (cf Figure 1.1), ce qui pose des problèmes évidents de sécurité et de coût d'arrêt des installations dans le contexte de l'industrie nucléaire.

Suivant les conditions expérimentales, notamment les vitesses des deux phases ou le taux de vide, un écoulement diphasique peut se trouver dans différentes configurations. Le terme *configuration* est ici préféré à l'appellation *régime* car il semble mieux traduire l'anglais *pattern* : il s'agit de décrire la distribution spatiale des interfaces qui séparent les phases.

Les configurations de l'écoulement diphasique transverse sont directement liées aux mécanismes d'interaction fluide-structure évoqués plus haut. Il est donc essentiel de pouvoir prédire quelles configurations peuvent être observées et dans quelles conditions expérimentales.

1.2 Problématique

Dans le cas des écoulements diphasiques, l'approche théorique de description du comportement des interfaces par des équations de transport se révèle à la fois complexe et insuffisante pour le moment car il reste encore beaucoup à faire à ce sujet.

Afin de progresser dans une connaissance pratique permettant d'assister la conception industrielle, une approche plus pragmatique s'est développée. Elle consiste à obtenir expérimentalement des indicateurs reliés à la configuration de l'écoulement, qui prennent des valeurs quantitatives ou caractéristiques qualitatives différentes selon les conditions expérimentales, ce qui permet de classer des mesures les unes par rapport aux autres. On définit ainsi des groupes dont les points présentent des indicateurs similaires, et des noms de configuration sont ensuite attribués à ces groupes selon les caractéristiques géométriques observées visuellement.

Pour ce qui est des écoulements internes, de nombreux travaux ont été menés depuis les années 1960 pour identifier de cette façon les différentes configurations possibles et représenter leurs frontières sur des cartes d'écoulement.

La carte proposée par Taitel et al. (1980) pour un écoulement diphasique gaz-liquide ascendant en conduite verticale, basée sur une analyse théorique des phénomènes physiques, reste une référence majeure car le fait d'extrapoler à d'autres montages les frontières obtenues expérimentalement pour une boucle d'essai donnée présente des limites. De nombreux auteurs ont par la suite confirmé ou complété le travail de Taitel et al., ce qui fait que les configurations d'écoulement en conduite sont à présent bien définies et des cartes d'écoulement permettent de prédire les conditions expérimentales correspondantes et les zones de transition.

Les écoulements diphasiques en faisceau de tubes ont été beaucoup moins étudiés, probablement à cause de leur complexité. Plusieurs cartes ont néanmoins été proposées, et seront présentées plus en détail dans la revue de littérature (Chapitre 2). Cependant, la plupart des frontières définies sont basées sur des observations visuelles depuis l'extérieur du faisceau, parfois couplées à des mesures quantitatives, mais analysées manuellement en fin de compte pour attribuer une configuration à chaque point de mesure et en déduire des frontières.

Le fait de définir des indicateurs visuels ou de classer les points de mesure manuellement, bien qu'ayant permis des avancées importantes, présente un haut degré de subjectivité, qui fait que la fiabilité des frontières proposées peut être partiellement remise en question.

Premièrement, les observations visuelles sont limitées et difficiles à interpréter, même dans le cas de photographies par caméra haute vitesse, à cause de la complexité des interfaces entre les phases (Wallis, 1969) et du fait qu'il s'agit d'images en 2D d'un écoulement 3D.

Deuxièmement, dans les zones de transition entre deux configurations, les caractéristiques des observations que l'on peut obtenir (images ou mesures) ne sont pas claires, et deux configurations différentes mais voisines peuvent sembler aussi probables l'une que l'autre. Dans le cas de l'écoulement en conduite, Taitel et al. (1980) ont montré qu'il y a parfois des divergences non négligeables entre les différentes frontières proposées basées sur des observations visuelles, à cause de la subjectivité de la personne qui effectue la classification.

Enfin, Noghrehkar et al. (1999) ont montré qu'il peut y avoir contradiction entre la configuration observée *au niveau de la paroi* (visuellement) et la configuration qui serait attribuée selon des mesures effectuées simultanément à l'*intérieur* du faisceau.

Il y a donc un besoin réel d'une méthode plus objective permettant de définir les caractéristiques des configurations d'un écoulement diphasique transverse en faisceau de tubes et d'en définir les frontières.

Intelligence artificielle et cartes d'écoulement

Par ailleurs, depuis les années 1950, s'est développé un nouveau champ de connaissances : l'intelligence artificielle (IA). Une des branches de l'IA est le *Machine Learning* (qu'on peut traduire par apprentissage automatique), qui consiste à utiliser les possibilités de calcul des nouveaux outils informatiques pour construire des « machines » capables de déduire des groupes sous-jacents à un ensemble de données, de façon autonome ou plus ou moins guidée (on parle alors d'apprentissage supervisé).

Ce champ d'études a ouvert des possibilités très intéressantes, en particulier dans le domaine de la classification de données. Des algorithmes basés sur un apprentissage supervisé comme les machines à vecteur de support (*Support Vector Machines*) ou les réseaux de neurones ont été

utilisés avec succès pour classer de façon plus objective des mesures quantitatives effectuées pour des écoulements *internes* (en conduite). On peut citer notamment les travaux de Tan et al. (2007) avec une SVM ou encore Rosa et al. (2010), qui ont comparé la performance de plusieurs algorithmes d'apprentissage dont des réseaux de neurones (*Artificial Neural Networks*). Au sein de la Chaire, une étude réalisée par Komgom (2006) dans le cas d'une conduite a également démontré le potentiel des SVM pour l'identification des configurations d'écoulement.

Toutefois, ces algorithmes ne semblent pas avoir été encore mis à profit pour l'identification des configurations dans les faisceaux de tubes.

1.3 Objectifs de la recherche

Étant donné le contexte évoqué ci-dessus, les objectifs suivants ont été définis.

1.3.1 Objectif général

L'objectif principal de ce projet est de développer une méthode expérimentale d'identification des différentes configurations d'un écoulement diphasique présentes dans les faisceaux de tubes, basée sur l'analyse de signaux obtenus par sonde optique, à l'aide d'une machine à vecteurs de support (SVM).

1.3.2 Objectifs spécifiques

Le contexte de l'étude et les travaux de recherche consultés – présentés plus en détail dans le Chapitre 2 – permettent de décliner cet objectif de la façon suivante :

- 1) Balayer une large plage de conditions expérimentales lors des mesures par sonde optique et prise simultanée d'images de l'écoulement par caméra haute vitesse.
- 2) Analyser les signaux obtenus afin d'en extraire des indicateurs permettant de différencier les configurations d'écoulement, conjointement avec les observations visuelles et les données de la littérature, qui permettent de donner un sens physique aux groupes qui se dégagent de cette analyse.

- 3) Concevoir une machine à vecteurs de support adaptée aux données obtenues, puis réaliser l'entraînement de la SVM à partir de points de référence dont la configuration est facilement identifiable (c'est-à-dire loin des frontières définies par la littérature, avec vérification visuelle), afin d'effectuer ensuite la classification des autres points de mesure.
- 4) En déduire une carte d'écoulement pour la section d'essai utilisée, et comparer les résultats obtenus avec la littérature, afin de valider la méthode expérimentale d'identification ainsi développée.

NB : Le fait de s'appuyer sur des observations visuelles et donc subjectives dans une démarche qui se veut plus objective peut sembler paradoxal.

Premièrement, on peut remarquer que le fait d'utiliser comme données d'entrée pour l'algorithme de classification des données quantitatives, provenant directement de l'écoulement comme les signaux de sonde, et analysées de façon à obtenir des indicateurs pertinents reliés à la configuration, constitue déjà un progrès par rapport à des observations visuelles seules.

Il faut ensuite noter que dans le cas d'un algorithme d'apprentissage supervisé, il est impossible de se passer de points de référence, dont la configuration ne peut être définie que manuellement, avec une part de subjectivité. Mais en prenant des points de référence loin des frontières supposées, dont la configuration est clairement définie, le fait d'utiliser ensuite un tel algorithme pour classer les points des zones de transition paraît être une solution déjà plus objective qu'une classification manuelle de l'ensemble des points. De façon générale, pour que la classification proposée ait un sens physique, il est inévitable à un moment ou à un autre de relier les mesures quantitatives (quelles qu'elles soient) à l'observation de la structure réelle de l'écoulement, qui reste une donnée qualitative comportant donc une part de subjectivité.

CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre présente les définitions et modèles de base permettant d'étudier les écoulements diphasiques, ainsi que les travaux de recherche existants sur lesquels on s'appuiera par la suite.

2.1 Modèles d'analyse des écoulements diphasiques

2.1.1 Définitions et modèles d'analyse

Afin de décrire un écoulement diphasique, les paramètres suivants sont définis :

Taux de vide ε

Le taux de vide d'un écoulement diphasique est le ratio du volume occupé par la phase gazeuse sur le volume total :

$$\varepsilon = \frac{V_g}{V_g + V_l} \quad (2.1)$$

Où V_g et V_l sont respectivement les volumes des phases gazeuse et liquide.

Titre volumique β

Le titre volumique d'un écoulement diphasique est le ratio du débit volumique de la phase gazeuse sur le débit volumique total :

$$\beta = \frac{Q_g}{Q_g + Q_l} \quad (2.2)$$

Où Q_g et Q_l sont respectivement les débits volumiques des phases gazeuse et liquide.

Vitesses de phase U_g et U_l

U_g et U_l sont les vitesses de phase, respectivement gazeuse et liquide, moyennées sur l'aire transverse de la conduite, A :

$$U_g = \frac{Q_g}{A \varepsilon} \quad (2.3)$$

et

$$U_l = \frac{Q_l}{A (1 - \varepsilon)} \quad (2.4)$$

Vitesses superficielles j_g et j_l

j_g et j_l sont les vitesses superficielles, respectivement gazeuse et liquide, définies par le rapport entre le débit volumique correspondant et l'aire transverse de la conduite A :

$$j_g = \frac{Q_g}{A} \quad (2.5)$$

et

$$j_l = \frac{Q_l}{A} \quad (2.6)$$

Taux de glissement S

Pour un écoulement diphasique réel, les vitesses de phase U_g et U_l ne sont pas égales, et il existe un glissement entre les deux phases. On définit ainsi le taux de glissement S par le ratio de la vitesse de phase gazeuse sur la vitesse de phase liquide :

$$S = \frac{U_g}{U_l} \quad (2.7)$$

Il y a de fait une différence entre le taux de vide ε et le titre volumique β , qui sont reliés par la relation :

$$\frac{1 - \beta}{\beta} = \left(\frac{1 - \varepsilon}{\varepsilon} \right) \frac{1}{S} \quad (2.8)$$

Modèles d'analyse des écoulements diphasiques

Les équations fondamentales qui permettent de décrire les écoulements diphasiques sont basées sur les équations de conservation de masse, de quantité de mouvement et d'énergie. A cause de la nature diphasique de l'écoulement, il faut y ajouter des équations qui prennent en compte les interactions entre les phases, à l'interface. Les modèles les plus courants permettant de décrire le comportement du mélange diphasique sont (Wallis, 1969):

- le modèle *homogène* : le mélange diphasique est simplement considéré comme un fluide monophasique, et les propriétés du mélange sont alors obtenues en moyennant les propriétés de chaque phase.
- le modèle à *écoulement séparé* : on considère séparément chaque phase du mélange mais une interface commune est définie, au niveau de laquelle on détaille les interactions des deux phases.
- le modèle à *écart de vitesse* : on tient compte du glissement entre les phases en introduisant une vitesse de glissement.

Lorsqu'il s'agit d'établir une carte d'écoulement, le modèle homogène est généralement utilisé, car il permet de relier directement le taux de vide aux débits volumiques imposés dans la section d'essai, tout en constituant une approximation acceptable dans ce contexte particulier.

On utilisera donc le modèle homogène dans la suite de ce mémoire.

2.1.2 Modèle homogène

Comme mentionné précédemment, en choisissant le modèle homogène pour décrire l'écoulement diphasique, on n'étudie pas séparément chacune des phases, mais on s'intéresse au mélange dans lequel les deux phases sont supposées être réparties de façon homogène et se déplacer à la même vitesse : la vitesse du mélange homogène, U_{hom} .

Dans les équations précédentes, cela se traduit par :

$$U_g = U_l = U_{hom} \quad \Leftrightarrow \quad S = 1 \quad \Leftrightarrow \quad \varepsilon = \beta \quad (2.9)$$

On peut donc définir une vitesse du mélange homogène en fonction des paramètres expérimentaux : les débits volumiques de chaque phase et l'aire transverse de la section :

$$U_{hom} = j = \frac{Q_g + Q_l}{A} \quad (2.10)$$

Et les vitesses superficielles s'expriment alors ainsi :

$$j_g = \beta j \quad (2.11)$$

Et

$$j_l = (1 - \beta) j \quad (2.12)$$

Cela présente le grand avantage de pouvoir imposer directement dans la section d'essai un taux de vide et des vitesses superficielles donnés, par le biais du titre volumique, qui est défini en fonction des débits volumiques Q_g et Q_l qu'on peut facilement régler sur le montage expérimental.

2.2 Configurations d'écoulement diphasique

Du fait de sa nature diphasique, l'écoulement peut s'organiser de diverses manières selon ses propriétés physiques, les vitesses des phases, le taux de vide... il s'agit alors de décrire la distribution spatiale des interfaces qui séparent les phases, et donc de définir différentes configurations de l'écoulement. Celles-ci sont ensuite positionnées sur une carte d'écoulement, qui permet de représenter les frontières entre les configurations dans un repère approprié, comme celui des vitesses superficielles - définies à la section 2.1.1 : (j_g, j_l) , ou leur équivalent adimensionnel, ou encore des grandeurs faisant intervenir d'autres paramètres physiques des deux phases.

2.2.1 Configurations d'un écoulement vertical ascendant dans un tube

Avant de s'intéresser aux configurations qui ont été observées dans les faisceaux de tubes, il paraît important de décrire celles que l'on observe à l'intérieur d'un tube. Étant mieux connues, elles pourront servir d'élément de comparaison. On s'est notamment appuyé sur le travail réalisé au sein de la Chaire par Giraudeau (2011).

Les quatre configurations communément décrites, notamment d'après Taitel et al. (1980), sont : la configuration à bulles (*bubbly*), à bouchons (*slug*), agitée (*churn*) et annulaire (*annular*), représentées Figure 2.1.

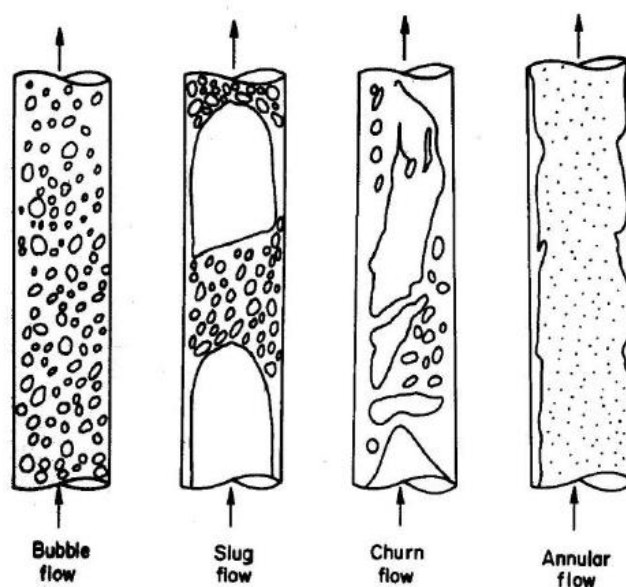


Figure 2.1 : Configurations d'un écoulement vertical ascendant dans un tube d'après Taitel et al. (1980)

Configuration à bulles

La phase gazeuse se trouve sous la forme de petites bulles uniformément réparties au sein d'une phase liquide continue. L'écoulement est strictement ascendant.

D'après Taitel et al., la configuration à bulles ne peut exister que pour des conduites de faible diamètre à des débits de liquide assez élevés donc des forces de dispersion par turbulence assez élevées elles-aussi.

Configuration à bouchons

La majorité de la phase gazeuse se trouve sous la forme de grandes bulles de la forme d'une balle (*bullet*) dont le diamètre est quasiment égal au diamètre de la conduite. Elles sont souvent appelées *bulles de Taylor* et se déplacent uniformément vers le haut. Elles sont séparées par des bouchons continus de liquide qui contiennent des petites bulles, et relient les parois de la conduite. Entre les bulles de Taylor et les parois se trouve un fin film de liquide descendant.

Configuration agitée

Cette configuration est une version plus chaotique, mousseuse et désordonnée de la configuration à bouchons. Les bulles de Taylor deviennent déformées et plus étroites. A cause de l'augmentation locale de la concentration de la phase gazeuse dans les bouchons de liquide, ceux-ci finissent par se désagréger et s'effondrer dans la bulle de Taylor qui les suivait, jusqu'à ce que le liquide s'accumule assez pour pouvoir être de nouveau être porté par le gaz et constituer un nouveau bouchon. La configuration agitée présente donc un mouvement oscillant typique, puisque le mélange diphasique se déplace alternativement vers le haut et vers le bas.

Configuration annulaire

La phase gazeuse occupe entièrement le cœur de la conduite, et la phase liquide est entraînée vers le haut à la fois sous forme d'un film liquide ondulé sur les parois et de gouttelettes portées par le gaz qui se trouve au centre de la conduite. (On trouve aussi la configuration *whispy-annular* lorsque la phase liquide entraînée se trouve sous la forme de filaments plus grands).

D'après les équations des modèles théoriques développés par Taitel et al., basés sur la compréhension des phénomènes physiques qui régissent les transitions, ces quatre configurations peuvent être positionnées sur la carte d'écoulement présentée à la Figure 2.2, dans le repère des vitesses superficielles :

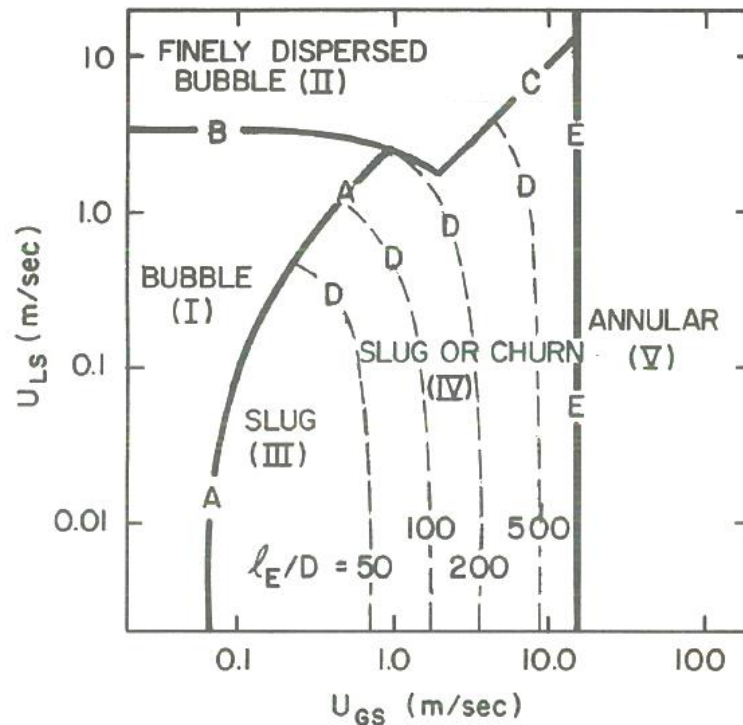


Figure 2.2 : Carte des configurations d'un écoulement vertical ascendant dans un tube d'après Taitel et al. (1980)

(tube de diamètre $D=5\text{cm}$, mélange air-eau à 25°C)

Taitel et al. distinguent deux configurations différentes à bulles selon la dimension de celles-ci (d'où l'appellation « à bulles finement dispersées ») – cf Figure 2.2, mais cette distinction ne fait pas l'unanimité. Costigan et Whalley (1997), confirmant les conclusions de Kaichiro et Ishii (1984) ont montré expérimentalement que bulles et bulles finement dispersées constituent en fait une seule et même configuration, car elles génèrent le même profil de PDF (*Probability Density Function*) du taux de vide.

Ces mêmes expériences ont en revanche montré l'existence d'une configuration différente : les *bulles à calotte sphérique*, pour lesquelles les profils de PDF du taux de vide se distinguent de ceux qu'on obtient pour la configuration à bulles. Taitel et al. font bien mention de bulles de Taylor trop petites pour remplir complètement la conduite et constituer un bouchon, et observent qu'elles ont une calotte sphérique. Ils n'en avaient toutefois pas fait une configuration différente.

Par ailleurs, la transition entre les configurations à bouchons et annulaire a suscité beaucoup de discussions, et plusieurs auteurs ont proposé des mécanismes de transition différents, donnant lieu à des distinctions entre les configurations.

Des observations de Jayanti et al. (1993) ont confirmé le mécanisme proposé par Jayanti et Hewitt (1992), et amélioré par Costigan et Whalley (1997) : dans l'écoulement à bouchons, avec l'augmentation de l'épaisseur du film de liquide entourant les bulles de Taylor, ces bulles finissent par devenir instables (ce mécanisme est appelé *flooding*). Cela conduit à un écoulement agité, aussi décrit comme un *écoulement à bouchons instables*. Dans ce cas, la configuration agitée est scindée en deux configurations moins ambiguës : à bouchons instables et semi-annulaire, en anglais *Churn-slug* et *Churn-Annular* d'après Jayanti et al. (1993).

La configuration semi-annulaire correspond alors à une forme dégénérée de la configuration annulaire, où le cœur du tube est bien occupé par la phase gazeuse, mais où on observe des vagues d'amplitude non négligeable à l'interface entre le gaz et le liquide, avec des changements de direction possibles.

Afin de décrire plus précisément les configurations et d'éviter l'ambiguïté liée au qualificatif « agité », certains auteurs ont donc choisi de se référer à un ensemble de 6 configurations, parmi lesquels Giraudeau et al. (2011) et Rosa et al. (2010) – cf Figure 2.3 et Figure 2.4.

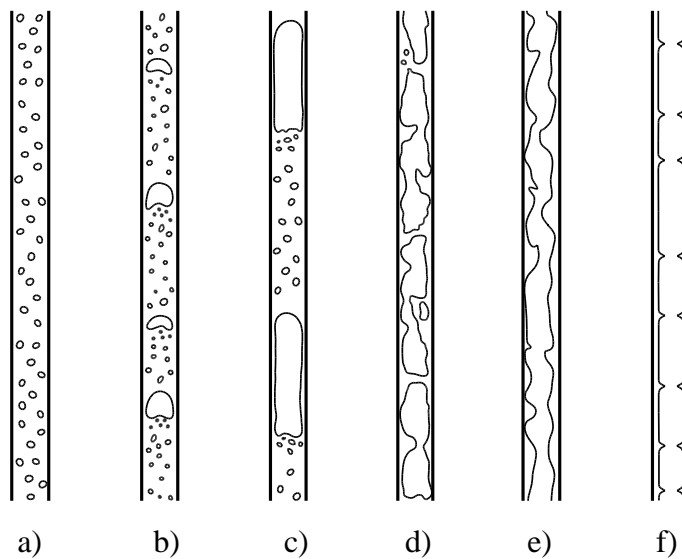


Figure 2.3 : Configurations d'écoulement d'après Giraudeau (2011) :
 écoulement a) à bulles, b) à bulles à calotte sphérique, c) à bouchons ou intermittent,
 d) à bouchons instables, e) agité/annulaire ou semi-annulaire et f) annulaire

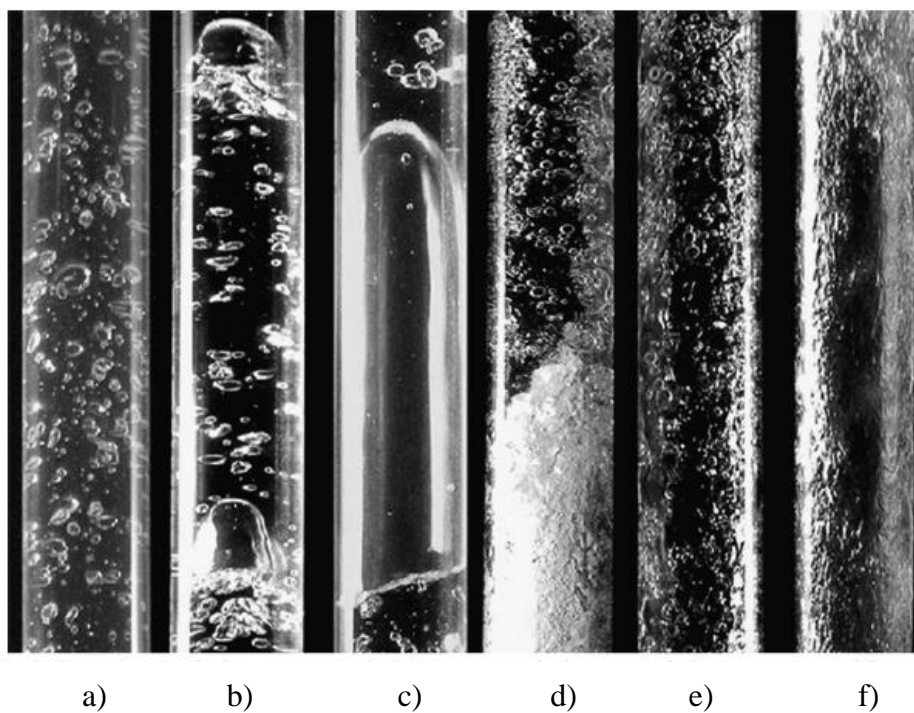


Figure 2.4 : Photographies de l'écoulement d'après Rosa et al. (2010) :
 a) à bulles, b) à bulles à calotte sphérique, c) à bouchons, d) à bouchons instables,
 e) semi-annulaire et f) annulaire

2.2.2 Configurations d'écoulement en faisceau de tubes

La revue de littérature suivante se concentre sur les travaux réalisés en faisceaux de tubes horizontaux, pour un écoulement vertical transverse air-eau, ascendant vis-à-vis de la pesanteur, qui présentent une étude détaillée des configurations d'écoulement observées et si possible une carte d'écoulement associée. Elle s'appuie notamment sur les différentes revues des articles cités, ainsi que sur la revue de Ribatski & Thome (2007) sur les écoulements diphasiques en faisceaux de tubes horizontaux. La description des configurations d'écoulement citées sera reprise plus en détail dans la section 4.2.

Avant tout, un faisceau de tubes peut se trouver dans différentes configurations géométriques, dont deux exemples sont représentés à la Figure 2.5. On définit le pas du faisceau P , le diamètre des tubes D , et le rapport pas sur diamètre P/D (g représentant la pesanteur, et les vitesses ayant été définies en 2.1.1). La configuration du faisceau est appelée normale ou tournée selon l'angle entre la direction de l'écoulement et la base du triangle dont les sommets sont les centres de tubes adjacents. Dans les échangeurs de chaleur et les générateurs de vapeur, le rapport P/D est typiquement compris entre 1.2 et 3.0.

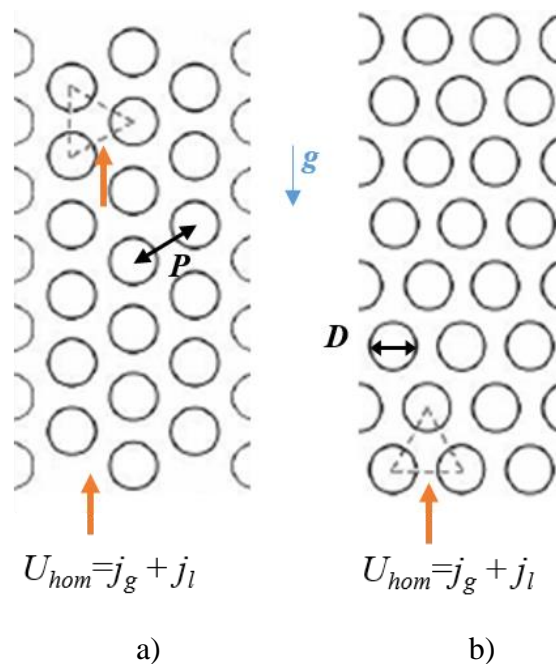


Figure 2.5 : Configurations d'un faisceau triangulaire :

a) triangulaire tourné, b) triangulaire normal

Le premier travail de recherche qui traite des configurations d'écoulement en faisceau de tubes tout en proposant une carte d'écoulement est celui de Grant et Murray en 1972, au début d'une série d'études détaillées menées au *National Engineering Laboratory* sur les faisceaux de tubes des échangeurs de chaleur, qui ont été synthétisées par Grant et Chisholm (1979). Grant et Murray ont étudié la chute de pression dans un échangeur de chaleur tubulaire, pour 39 tubes de diamètre 19 mm en configuration triangulaire normale avec un ratio pas sur diamètre (P/D) de 1.25. Ils ont observé visuellement trois configurations dans la zone où l'écoulement était ascendant : à bulles, intermittente et dispersée (*spray*).

La carte d'écoulement définie par Grant et Murray à la Figure 2.6 a été ramenée dans le repère des vitesses superficielles adimensionnées $u_{sG}^* = f(u_{sL}^*)$ par Ulbrich et Mewes (1994), les vitesses superficielles adimensionnées (coordonnées de Bell et al. (1970)) étant :

$$u_{sG}^* = u_{sG} \left(\frac{\rho_G}{\rho_L} \right)^{0.5} \quad (2.13)$$

$$u_{sL}^* = u_{sL} \left(\frac{\rho_L \eta_L}{\sigma} \right)^{0.33} \quad (2.14)$$

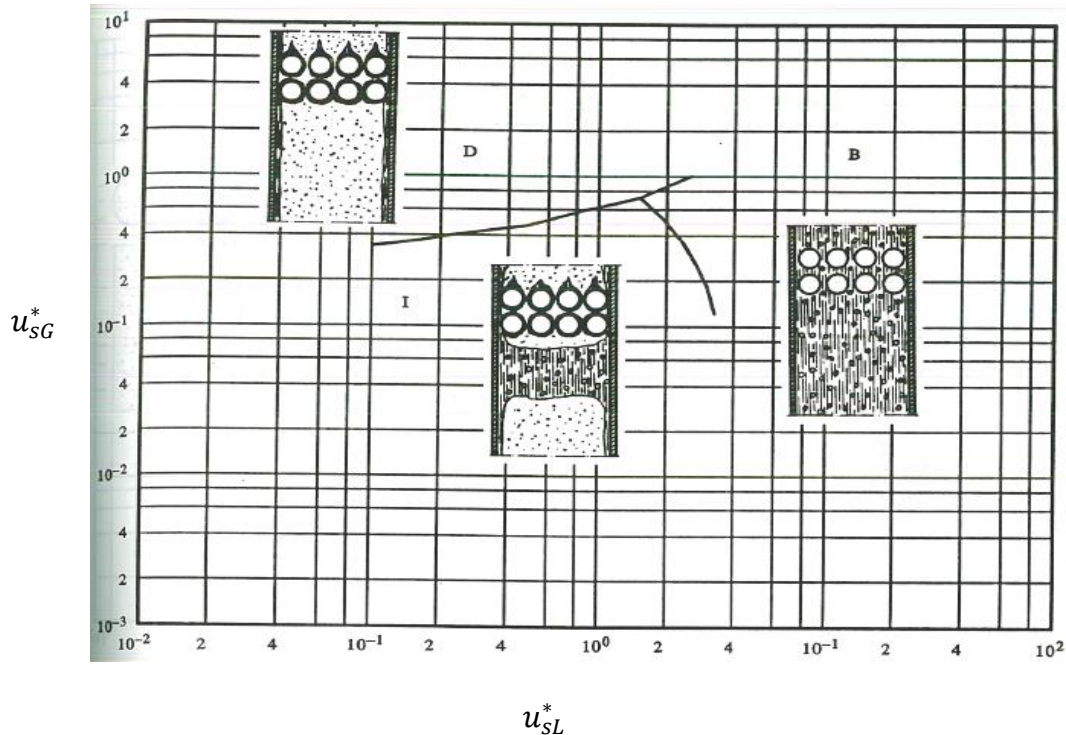


Figure 2.6 : Carte de Grant et Murray (1972) citée par Ulbrich et Mewes (1994)

Les configurations sont : B=à bulles, I=intermittente, et D=dispersée

Kondo et Nakajima (1980) ont également observé visuellement les configurations de l'écoulement, dans des faisceaux de configuration triangulaire de différents ratios P/D : 1.08, 1.28, et 1.4, pour des tubes de diamètre extérieur 25 mm. Ces configurations sont : à bulles, à bouchons, à mousse (*froth*) et dispersée (*spray*). D'après leurs observations, seule l'augmentation de la vitesse superficielle du gaz génère le changement de configuration, dans l'ordre indiqué ci-dessus. Une corrélation entre le taux de vide moyen et la hauteur des bulles a été obtenue, ainsi qu'une carte d'écoulement. Celle-ci est déduite à la fois des observations visuelles d'après des photos prises par caméra haute vitesse, et de distributions de fréquence relative calculées pour les ratios hauteur/longueur de bulles observés, qui présentent des allures distinctes sur lesquelles baser la classification.

Ces résultats ne sont pas directement utilisables car les vitesses utilisées sont relativement faibles : $j_{g\max} = 0.5$ m/s et $j_{l\max} = 0.0032$ m/s. La description des configurations observées est toutefois intéressante, tout en gardant à l'esprit les limites liées aux observations visuelles. Les configurations sont représentées par les schémas de la Figure 2.7.

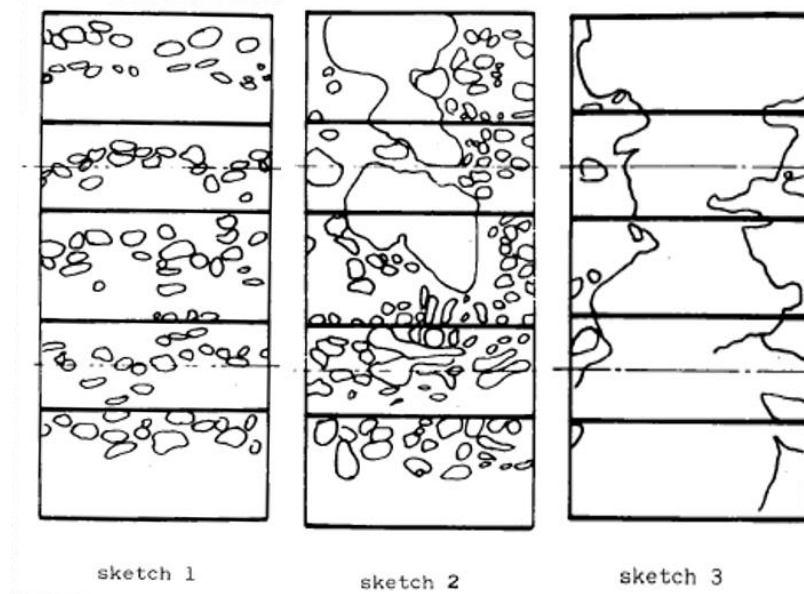


Figure 2.7 : Schémas de certaines configurations observées par Kondo et Nakajima (1980) :
1 : A bulles (Bubbly), 2 : A bouchons (Slug), 3 : A mousse (Froth)

Ulbrich et Mewes (1994) ont été les premiers à coupler les observations visuelles avec des mesures quantitatives, pour une plage de vitesses plus standard que celle de Kondo et Nakajima (1980). Les fluctuations de la chute de pression ont été mesurées entre la 2^e et la 9^e rangée d'un faisceau en ligne carré, comportant dix rangées de 5 tubes de diamètre extérieur 20 mm, avec un ratio P/D de 1.5. Différents points de mesure ont été choisis parmi la plage de vitesses suivante : $j_g \in [0.001; 0.65]$ m/s et $j_l \in [0.047; 9.3]$ m/s. Des photographies et vidéos de l'écoulement par caméra haute vitesse, couplées à l'analyse statistique de l'évolution temporelle de la chute de pression normalisée, ont permis d'identifier les configurations A bulles, Intermittente, et Dispersée, avec les sous-catégories Intermittent-Dispersé et Annulaire-Dispersé – cf Figure 2.8.

La comparaison des frontières obtenues avec celles de Grant et Murray (1972) n'est pas concluante, mais cela peut être dû à la nécessité d'extrapoler les frontières de Grant et Murray pour faire correspondre les plages de mesures, ce qui peut se révéler incorrect. La carte d'écoulement ainsi obtenue présente cependant un degré d'accord de 85% avec des données variées de la littérature, ce qui est considéré comme satisfaisant dans ce domaine. Elle est tracée dans le repère des vitesses superficielles : ($u_{sG} = j_g$; $u_{sL} = j_l$)

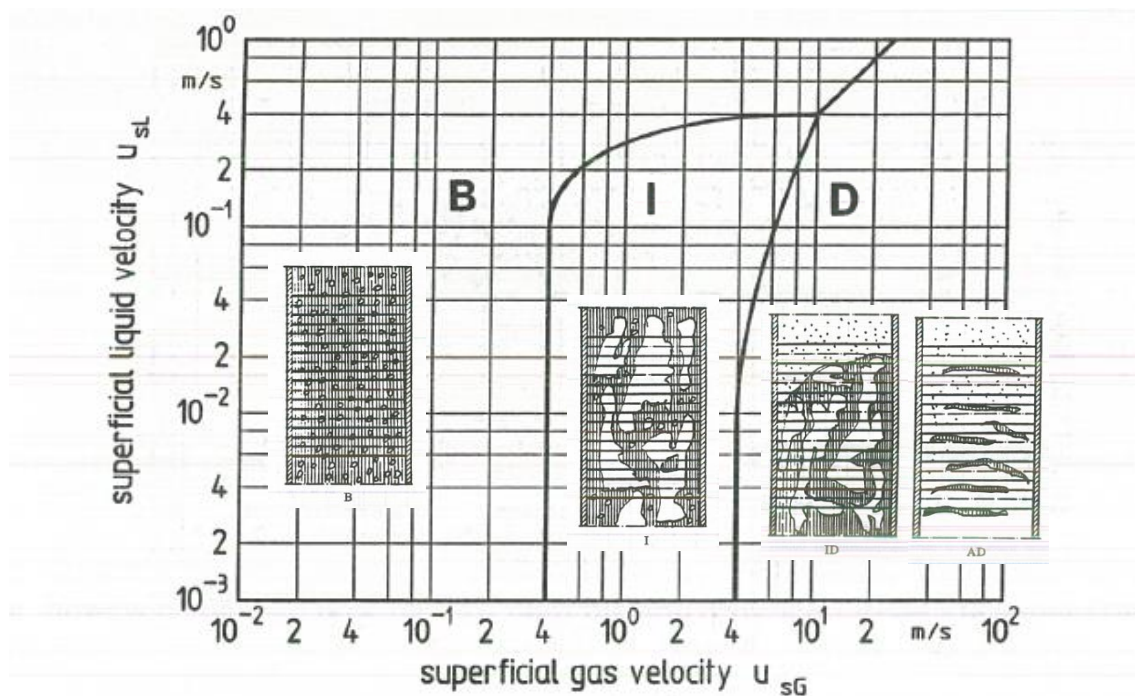


Figure 2.8 : Carte obtenue par Ulbrich et Mewes (1994) pour les configurations : B= A bulles, I=Intermittente, D=Dispersée; et schémas des écoulements correspondants

Ueno et al. (1995) ont étudié plusieurs grandeurs locales dans une section d'essai plus particulière : une conduite présentant des demi-tubes en regard sur les parois. Cela représente un des chemins d'écoulement d'un faisceau en ligne carré, comportant 20 rangées de deux demi-tubes de diamètre 19.05 mm, avec un ratio P/D de 1.42. Les vitesses superficielles varient ainsi : $j_g \in [1.0; 3.5]$ m/s et $j_l \in [0.1; 0.6]$ m/s. Bien que les configurations ne soient pas étudiées en détail, on trouve des informations intéressantes comme des distributions spatiales du taux de vide, ou des comparaisons entre les observations réalisées et les configurations connues en tube vertical : en faisceau, les bouchons de gaz n'ont pas la forme arrondie qu'ils ont en conduite, et leur taille est relativement plus petite qu'en conduite. Par ailleurs, on observe peu d'oscillation de l'écoulement dans cette section d'essai. La raison avancée pour justifier cette divergence est celle de la faible dimension des espaces inter-tubes qui empêcheraient une inversion locale du flux. Cet espace vaut en effet 8 mm, ce qui est relativement étroit, et expliquerait pourquoi peu d'oscillations ont été observées, contrairement à la plupart des autres études pour les mêmes gammes de vitesses.

Xu et al. (1998) se sont intéressés à l'hydrodynamique d'un écoulement diphasique dans un échangeur de chaleur, comportant 20 rangées de 3 tubes (plus deux demi-tubes sur les parois), de diamètre 9.79 mm, qui constituent un faisceau en ligne carré de ratio P/D 1.28. L'étude porte sur la chute de pression, le taux de vide, et les configurations d'écoulement ascendant et descendant dans les zones transverses. Celles-ci ont été observées visuellement dans des zones transparentes de la section d'essai, pour des vitesses superficielles telles que $j_g \in [0.047; 9.3]$ m/s et $j_l \in [0.001; 0.65]$ m/s. Les configurations observées pour l'écoulement ascendant sont : à bulles, agitée, intermittente et annulaire. La carte d'écoulement et les schémas de ces configurations sont donnés en Figure 2.9 et Figure 2.10. La comparaison avec la carte de Grant et Chisholm (1979) montre que les configurations d'écoulement sont similaires quelle que soit la disposition du faisceau (en ligne carré ou triangulaire). Les différences majeures observées concernent la transition vers la configuration annulaire, et la configuration agitée, qui n'a pas été observée par Grant et Chisholm – elle se situe en dehors de leur zone d'étude. La comparaison avec la carte de Ulbrich et Mewes (1994) révèle peu d'accord, excepté pour la transition vers la configuration annulaire.

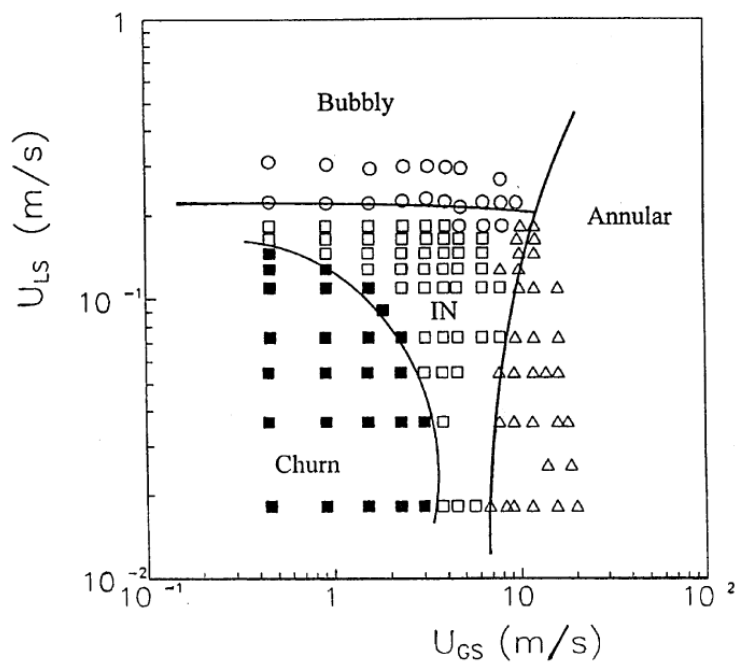


Figure 2.9 : Carte d'écoulement obtenue par Xu et al. (1998) sur laquelle sont représentées les configurations à bulles (*Bubbly*), intermittente (IN), agitée (*Churn*), et annulaire (*Annular*)

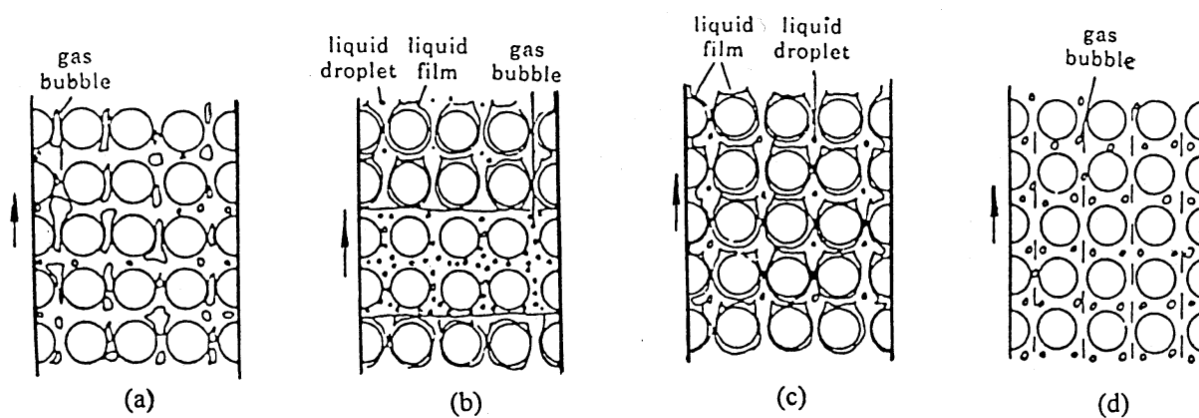


Figure 2.10 : Schémas des configurations décrites par Xu et al. (1998)

a) agitée, b) intermittente, c) annulaire et d) à bulles

Noghrehkar et al. (1999) ont étudié les configurations d'écoulement dans deux types de faisceaux : en ligne et triangulaire normal, comportant respectivement 24 rangées (5 tubes et 2 demi-tubes) et 26 rangées (5 tubes ou 4 tubes et 2 demi-tubes), de diamètre extérieur 12.7 mm, avec un ratio P/D de 1.47. Le taux de vide local a été mesuré à l'aide d'une sonde à résistance, et les PDF (fonctions de densité de probabilité) du taux de vide ont été obtenues. Elles ont permis d'obtenir les cartes d'écoulement de la Figure 2.11, jugées beaucoup plus objectives que celles qui se basent seulement sur des observations visuelles. En effet, Noghrehkar et al. ont également montré que la configuration observée visuellement depuis l'extérieur du faisceau peut différer de celle qu'indique la courbe de PDF du signal mesuré à l'intérieur du faisceau. Les configurations déduites des courbes de PDF sont : à bulles, intermittente et annulaire. La frontière à bulles/intermittente est légèrement différente selon la disposition du faisceau : cela semble provenir du fait que la géométrie du faisceau triangulaire normal est plus apte à scinder de grandes bulles en bulles plus petites, donc la configuration à bulles occupe une plus grande plage de vitesses dans ce cas.

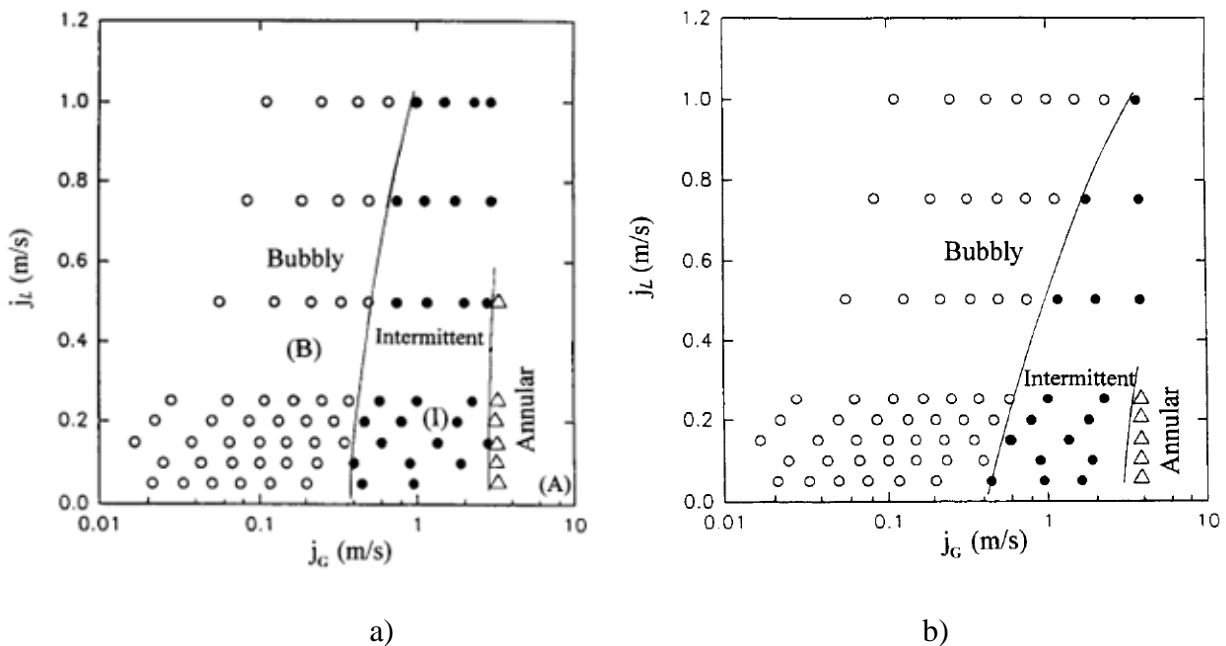


Figure 2.11 : Cartes des configurations de Noghrehkar et al. (1999) pour a) un faisceau en ligne et b) un faisceau triangulaire normal

Plus récemment, d'autres études ont utilisé des caméras haute vitesse plus perfectionnées pour assister l'identification visuelle par des images ou des vidéos prises à haute fréquence.

Parmi celles-ci, Chung et Chu (2006) ont étudié l'instabilité fluidélastique de faisceaux de tubes en configuration en ligne et carrée tournée, comportant respectivement 28 et 20 rangées de tubes de diamètre 12.7 mm, avec un ratio $P/D=1.633$. Les parois et les demi-tubes latéraux transparents ont permis d'observer les configurations suivantes : à bulles pour $\beta < 80\%$ environ, avec une transition vers un écoulement intermittent pour $\beta > 90\%$, quelle que soit la disposition du faisceau. Ces observations sont en accord avec une prédiction de configurations basée sur la carte de Ulbrich et Mewes (1994), comme le montre la Figure 2.12.

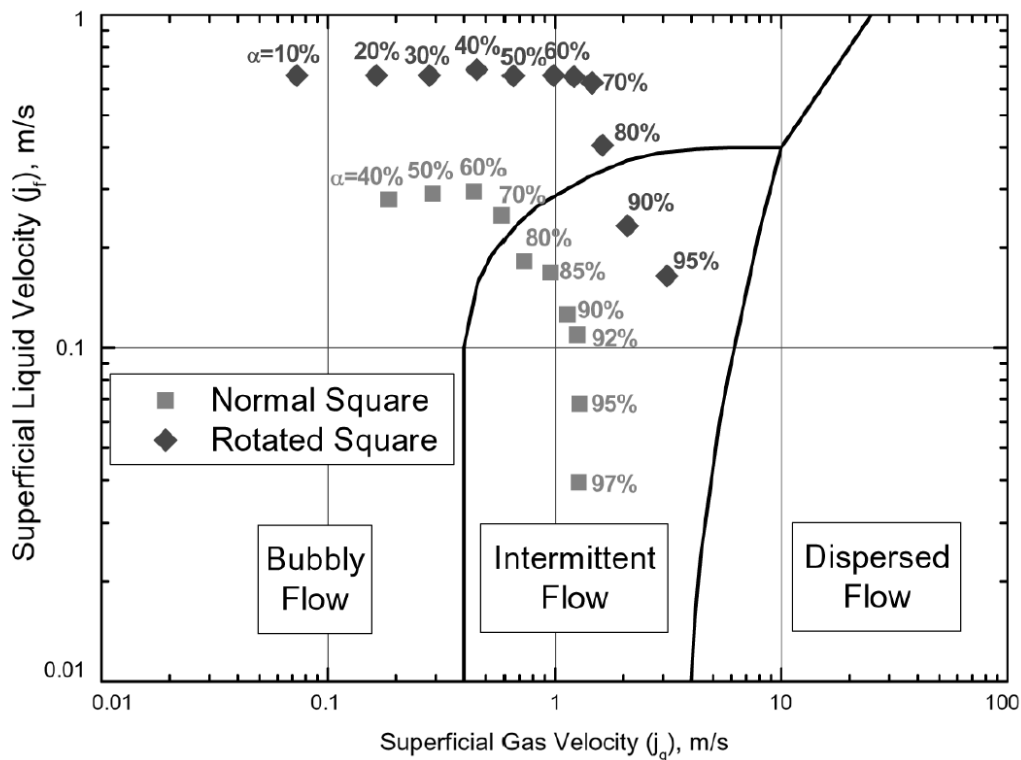


Figure 2.12 : Points de mesure de Chung et Chu (2006) sur la carte de Ulbrich et Mewes (1994)

Disposition du faisceau : ■ = en ligne, ◆ = carrée tournée

Hong et Liu (2010) ont également utilisé une caméra haute vitesse dans le but d'identifier les configurations d'écoulement, dans des faisceaux en ligne, avec des tubes de diamètre extérieur 20 mm, et des rapports P/D de 1.3 et 1.8. Les vitesses superficielles étudiées sont $j_g \in [0.03; 1.1]$ m/s et $j_l \in [0.02; 0.4]$ m/s. L'analyse des images obtenues a permis d'identifier les configurations à bulles, intermittente et dispersée (*spray*) – cf Figure 2.13. Pour les faisceaux présentant un petit pas P, une sous-configuration de l'écoulement à bulles a été observée pour les hautes vitesses et nommée « S-bubbly » à cause de la forme particulière des chemins d'écoulement autour des tubes. La comparaison avec les cartes de Ulbrich et Mewes (1994), Xu et al. (1998) et Noghrehkar et al. (1999) - Figure 2.14 et Figure 2.15 - révèle peu de similitudes, excepté pour les basses vitesses et la configuration à bulles. La configuration agitée de Xu et al. en particulier, n'a pas été détectée. Ces divergences sont attribuées à l'emploi de sections d'essai de dimensions différentes (rapport P/D par exemple).

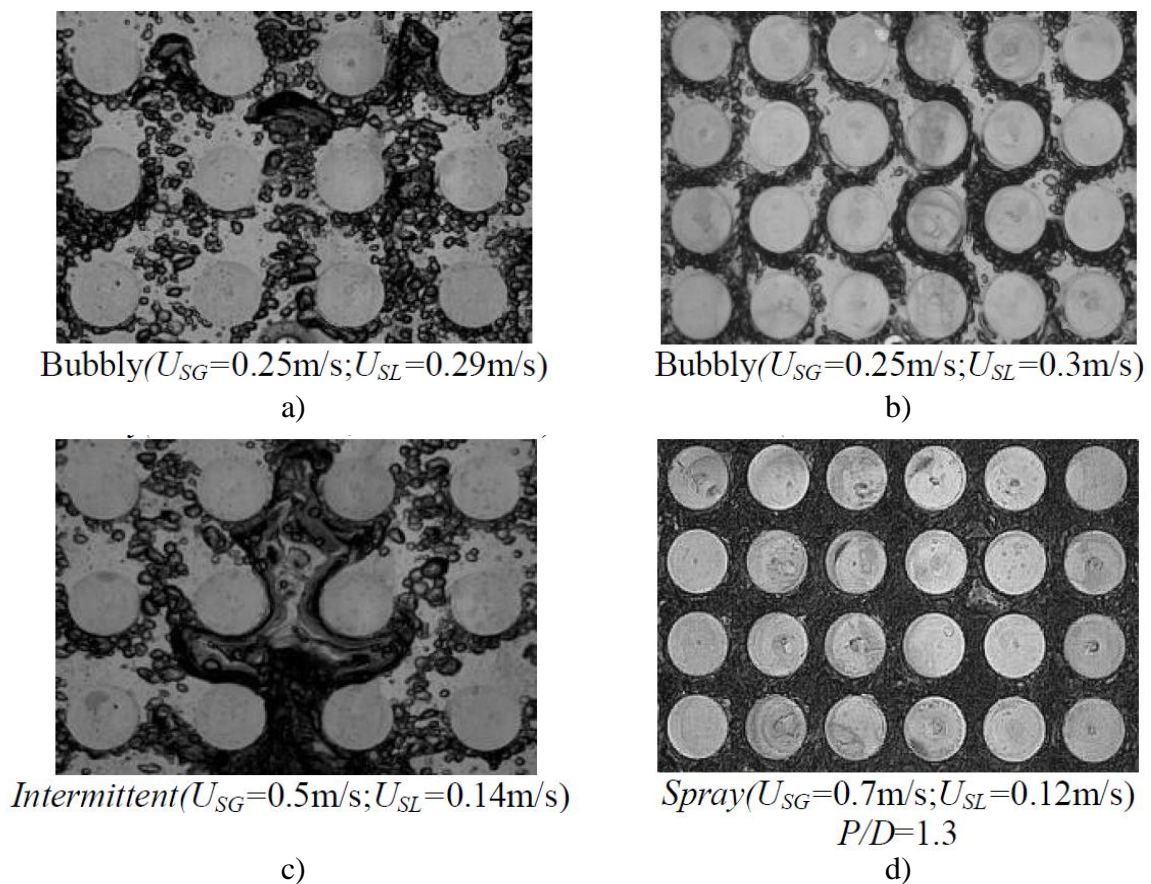


Figure 2.13 : Photographies caractéristiques obtenues par Hong et Liu (2010)

a) A bulles, b) A bulles « en forme de S », c) Intermittente, d) Dispersée

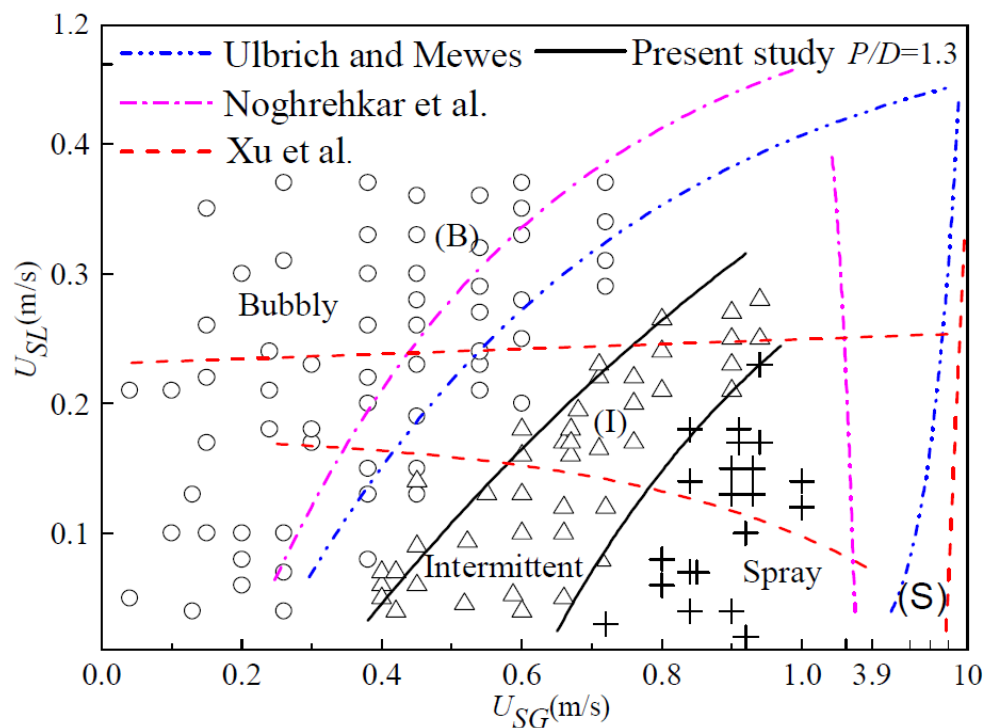


Figure 2.14 : Carte de Hong et Liu (2010) et comparaison avec la littérature pour $P/D=1.3$

Les symboles représentent les points de Hong et Liu : les cercles la configuration à bulles, les triangles la configuration intermittente, et les croix la configuration dispersée.

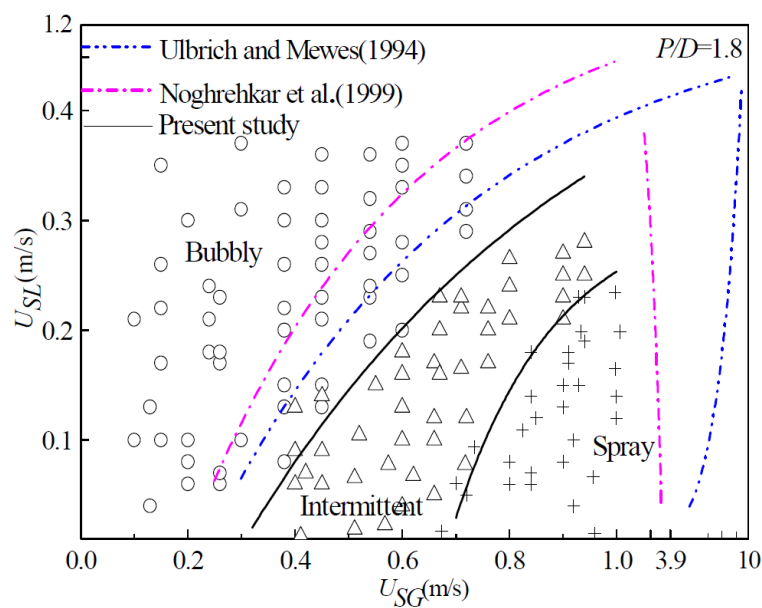


Figure 2.15 : Carte de Hong et Liu (2010) et comparaison avec la littérature pour $P/D=1.8$

On trouve une synthèse des comparaisons des différentes cartes obtenues pour un écoulement air-eau vertical ascendant dans des faisceaux en ligne dans la revue de Ribatski et Thome (2007) sur les écoulements diphasiques en faisceaux de tubes, citées dans la Figure 2.16.

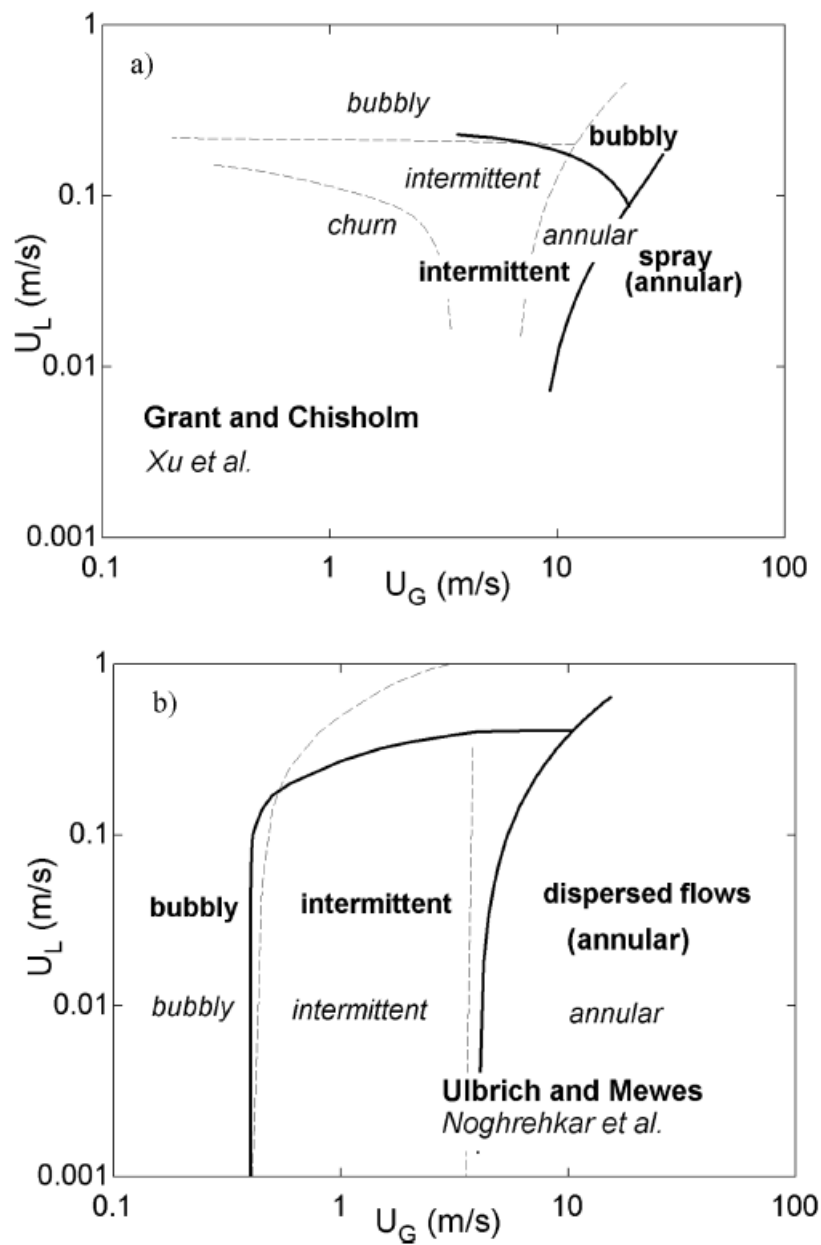


Figure 2.16 : Comparaison par Ribatski et Thome (2007) de cartes obtenues :
 a) par des observations visuelles seulement
 b) par l'analyse de mesures objectives (couplées aux observations visuelles)

Les principales caractéristiques des sections d'essai utilisées pour établir ces cartes sont rappelées dans le Tableau 2.1.

Tableau 2.1 : Caractéristiques des sections d'essai correspondant aux cartes de la Figure 2.16

| <i>Auteurs</i> | <i>Diamètre des tubes D (mm)</i> | <i>Rapport Pas sur diamètre P/D</i> | <i>Espace entre les tubes P – D (mm)</i> | <i>Rangées de tubes (sans compter les demi-tubes des parois)</i> |
|--------------------------|--------------------------------------|-------------------------------------|--|--|
| Grant et Chisholm (1979) | 19 | 1.25 | 4.8 | 11 rangées de 3 tubes |
| Xu et al. (1998) | 9.79 | 1.28 | 2.7 | 20 rangées de 3 tubes |
| Ulbrich et Mewes (1994) | 20 | 1.5 | 10.0 | 10 rangées de 5 tubes |
| Noghrehkar et al. (1999) | 12.7 | 1.47 | 6.0 | 24 rangées de 5 tubes |

Ribatski et Thome apportent les commentaires suivants :

- Les cartes basées seulement sur des observations visuelles - Figure 2.16 a) - ne donnent pas des frontières très similaires. La seule différence notable dans les conditions expérimentales concerne le diamètre des tubes, 1.94 fois plus petit pour le faisceau de Xu et al. par rapport à celui de Grant et Chisholm, et l'espace entre les tubes.
- Les frontières des cartes basées sur des méthodes plus objectives sont beaucoup plus proches, malgré une différence entre les diamètres des tubes là aussi et surtout l'utilisation de méthodes différentes : mesure instantanée de chute de pression ou analyse des PDF du taux de vide.

Étant donné le caractère subjectif des observations visuelles, et le risque d'observer des configurations d'écoulement différentes à la paroi et à l'intérieur du faisceau (Noghrehkar et al., 1999), qui met en doute le fait d'utiliser uniquement des observations depuis l'extérieur du faisceau, Ribatski et Thome préconisent d'employer une méthode qui combine observations visuelles et mesures objectives permettant de caractériser les configurations.

CHAPITRE 3 MÉTHODE EXPÉRIMENTALE

Ce chapitre décrit le montage expérimental, le choix des points de mesure, et les conditions dans lesquelles ces mesures ont été faites. Les appareils de mesure utilisés y sont également présentés, en particulier la sonde optique ayant permis d'obtenir les données principales.

3.1 Montage expérimental

3.1.1 Boucle d'essai

La boucle d'essai utilisée dans ce projet de maîtrise est une des boucles secondaires du laboratoire de la Chaire industrielle de recherche BWC/EACL/CRSNG en Interactions Fluide-Structure de l'École Polytechnique de Montréal. Elle a été conçue et mise en place en 2005 (Coquand, 2005). La section d'essai en elle-même a été adaptée en 2009 (Faure, 2009) dans le but d'effectuer des mesures du champ de vitesses de l'écoulement diphasique par PIV (*Particle Image Velocimetry*), pour étudier en détail la structure de l'écoulement (Mureithi & Masabarakiza, 2011).

Cette boucle d'essai est constituée d'une part d'un circuit d'air comprimé, et d'autre part d'un circuit d'eau provenant d'un réservoir, qui permettent d'injecter le mélange diphasique air-eau dans la section d'essai. L'ensemble des connexions - et notamment les tuyaux flexibles - permet de fonctionner en boucle fermée pour l'eau. La séparation des deux phases en sortie se fait naturellement par gravité au niveau du coude supérieur (élément 12 sur la Figure 3.1), et l'air est évacué au niveau d'une ouverture du réservoir.

La Figure 3.1 représente la boucle d'essai, dans sa configuration initiale (Coquand, 2005). Des modifications mineures au niveau de l'injection ont été apportées. Chaque circuit est muni d'un ou plusieurs débitmètres, présentés plus en détail en Annexe A. Suivant les points de mesure, deux pompes différentes ont été utilisées pour le circuit d'eau.

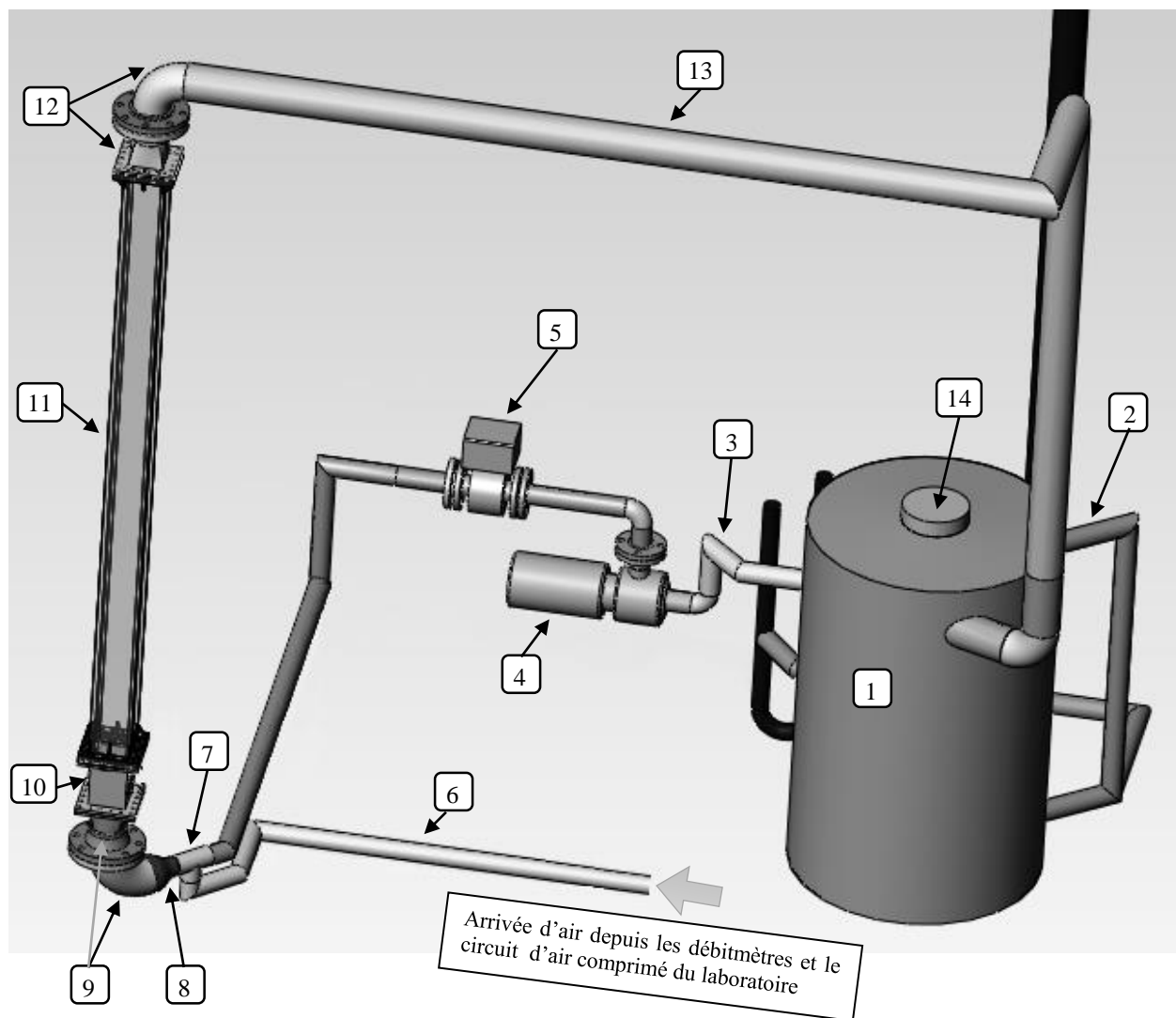


Figure 3.1 : Configuration initiale de la boucle d'essai

Tableau 3.1 : Éléments numérotés de la boucle d'essai (Figure 3.1)

| # de l'élément | Élément | # de l'élément | Élément |
|----------------|--|----------------|---|
| 1 | Réservoir | 8 | Divergent |
| 2 | Conduites de vidange et trop-plein du réservoir | 9 | Coude inférieur |
| 3 | Sortie du réservoir vers la boucle | 10 | Ancien injecteur d'air (<i>non utilisé</i>) |
| 4 | Pompe (Goulds Pumps) | 11 | Section d'essai |
| 5 | Débitmètre massique Krohne | 12 | Coude supérieur |
| 6 | Arrivée d'air depuis les débitmètres et le circuit d'air comprimé du laboratoire | 13 | Conduite flexible de retour au réservoir |
| 7 | Connecteur | 14 | Ouverture (évacuation de l'air) |

3.1.2 Section d'essai

La section d'essai et l'équipement de mesure associé sont présentés Figure 3.2.

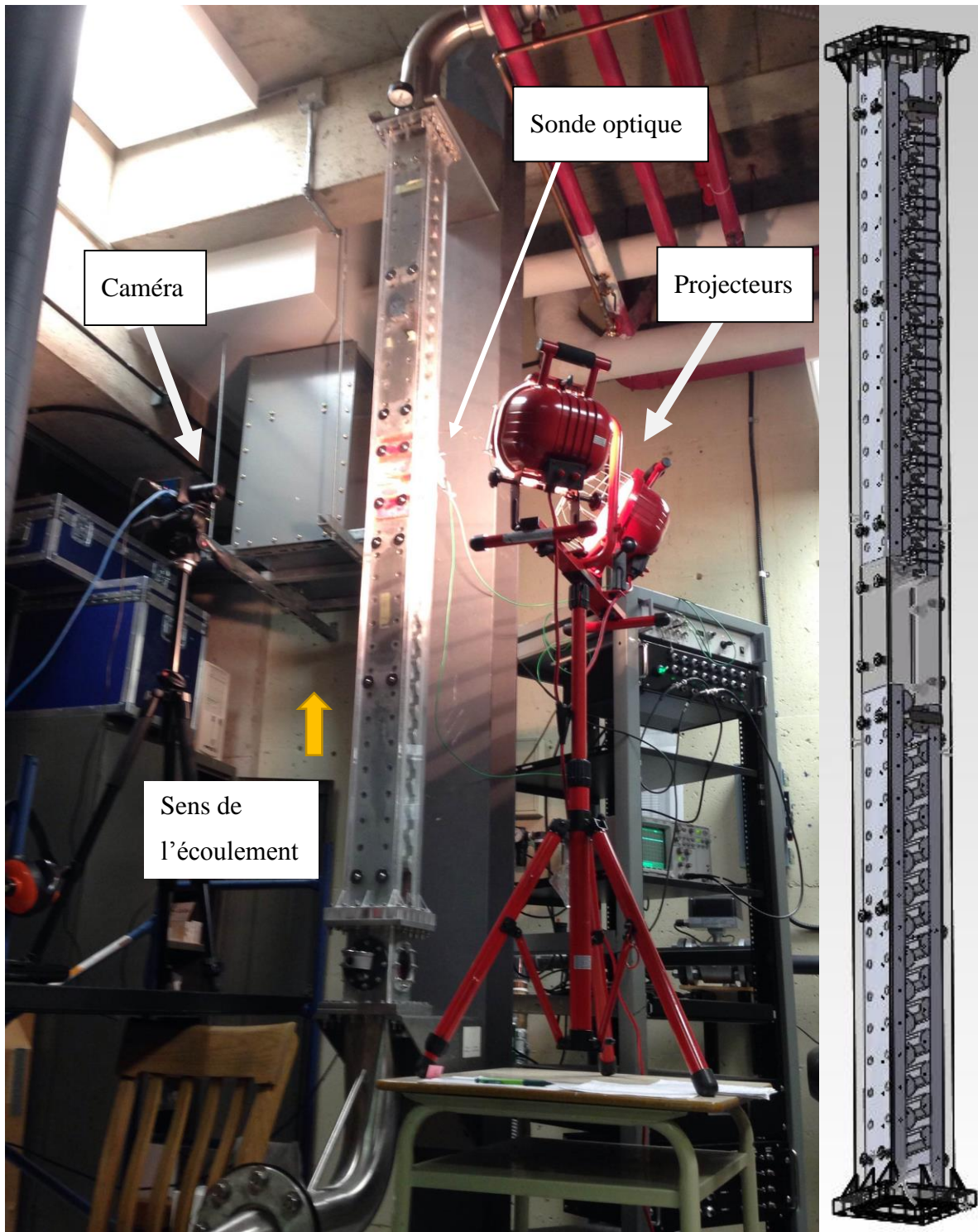


Figure 3.2 : Caméra et projecteurs

Caractéristiques géométriques et techniques

La section d'essai est constituée de plusieurs ensembles :

- Une conduite de section carrée.
- Des plaques vissées sur les parois de la conduite, sur lesquelles sont fixés les demi-tubes.

On obtient ainsi 33 rangées de demi-tubes sur chaque paroi, décalées pour obtenir la configuration triangulaire tournée (cf Figure 2.5). Le Tableau 3.2 présente les différents éléments de cette section d'essai, dont les détails des dimensions sont donnés en Annexe A.

Tableau 3.2 : Éléments de la section d'essai

| <i>Élément</i> | <i>Matériau</i> | <i>Quantité</i> | <i>Dimensions principales</i> |
|-------------------------------------|-------------------|---|---|
| Conduite de section carrée | Acrylique | 1 | Dimensions internes : 10 cm x 10 cm Hauteur : 2 m |
| Plaques métalliques | Aluminium anodisé | 4 | Hauteur \approx 88 ou 90 cm suivant la position, Épaisseur : 2.54 cm (1") |
| Demi-tubes | Aluminium anodisé | 58 | Rayon : $R=19$ mm |
| Vis | Acier inoxydable | 16 (plaques sur conduite) + 58 ($\frac{1}{2}$ tubes métalliques sur plaques) | |
| Plaques transparentes (pour la PIV) | Acrylique | 2 | Hauteur : 23 cm Épaisseur : 2.54 cm (1") |
| Demi-tubes (collés sur les plaques) | Acrylique | 8 | Rayon : $R=19$ mm |

La section d'essai est fixée par des vis sur deux supports métalliques, eux-mêmes vissés sur une poutre. Dans le cas de la partie inférieure, une pièce permettant d'autres types d'injection d'air se trouve entre le support métallique et la section d'essai (élément 10 sur la Figure 3.1), comme le montre la Figure 3.3.

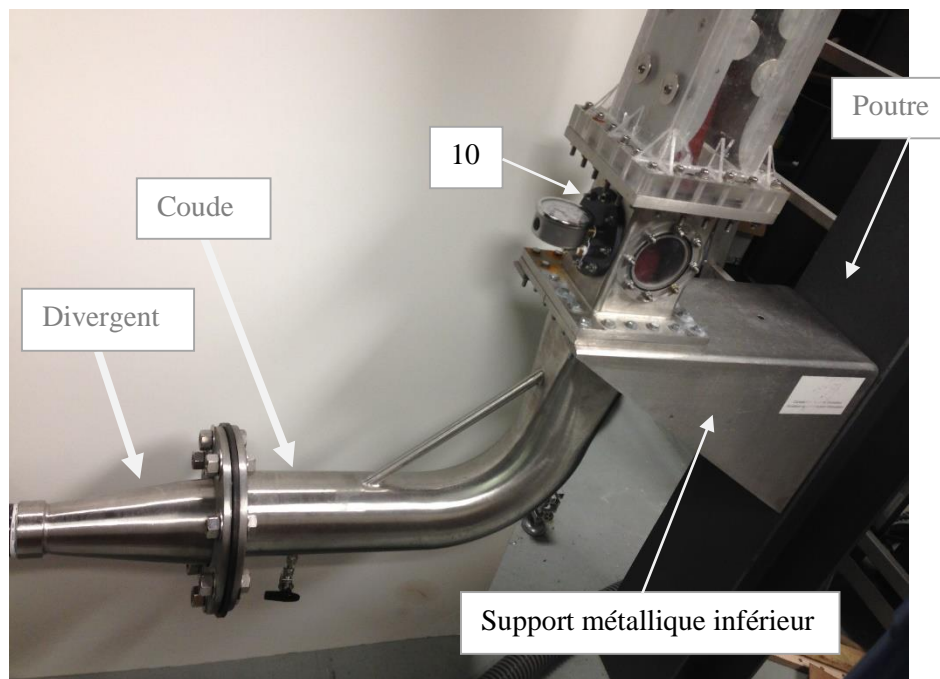


Figure 3.3 : Fixation de la section d'essai (support inférieur)

L'étanchéité entre la section d'essai et les bases du support est assurée par deux joints de forme carrée (*O-rings*) insérés dans des nervures.

Cette section d'essai représente le chemin d'écoulement d'un faisceau triangulaire tourné, comme l'illustre la Figure 3.4 (où on voit le haut de la section).

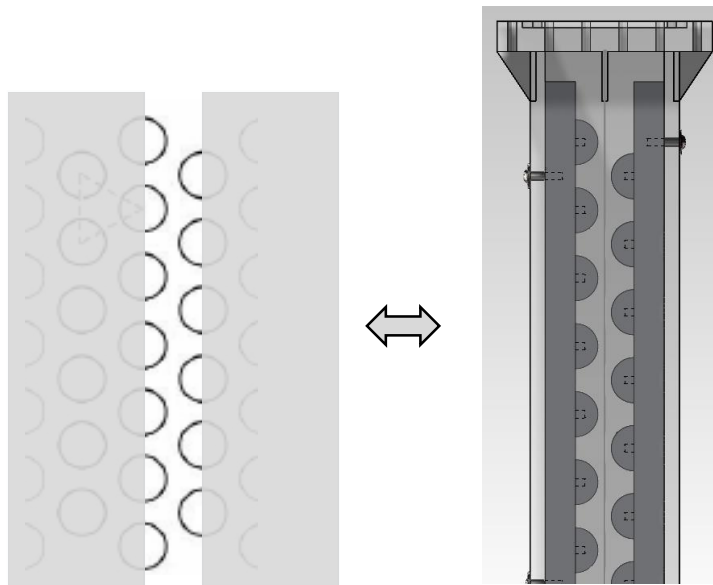


Figure 3.4 : Chemin d'écoulement d'un faisceau de tubes triangulaire tourné et portion de la section d'essai

L'intérêt de cette section d'essai est d'être en quelque sorte à mi-chemin entre une conduite et un faisceau de tubes, ce dernier étant difficile à étudier à cause de l'influence de la géométrie complexe sur l'écoulement. Tout en bénéficiant d'une géométrie plus facile à appréhender, on s'attend néanmoins à obtenir des configurations qui se rapprochent de celles qu'on observe en faisceau : à bulles, intermittente, et éventuellement annulaire.

Le rapport pas/diamètre, qui indique la « densité » du positionnement des tubes dans le faisceau (défini à la Figure 2.5) est :

$$\frac{P}{D} = 1.5 \quad (3.1)$$

Chemin d'écoulement

La Figure 3.5 représente le champ des vitesses moyennes de la phase gazeuse dans la section d'essai, obtenu par PIV (*Particle Image Velocimetry*) lors d'un projet mené au sein de la Chaire en 2011, sur la même section d'essai (Mureithi & Masabarakiza, 2011) :

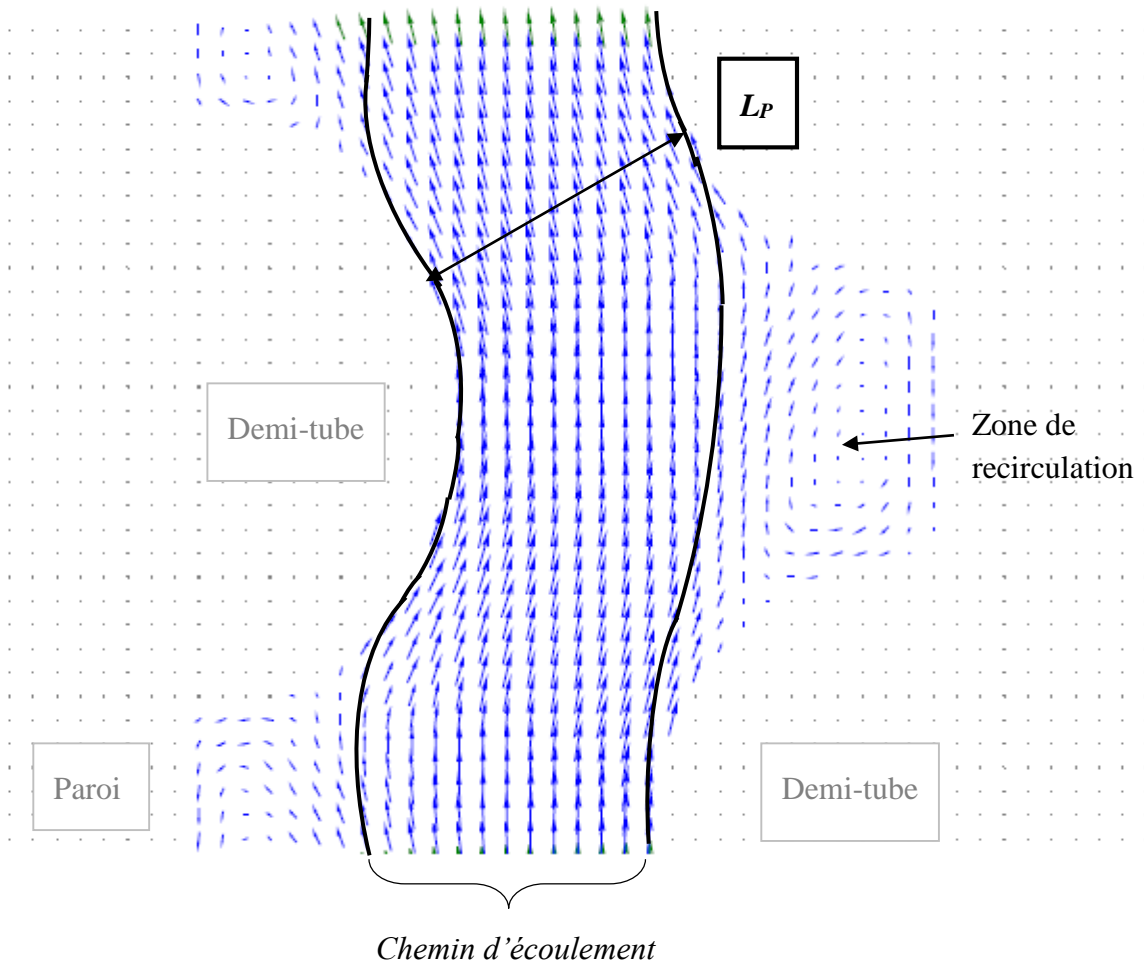


Figure 3.5 : Champ des vitesses moyennes de la phase gazeuse pour une vitesse inter-tubes $j_P = 0.3$ m/s, $\beta = 10\%$, d'après Mureithi & Masabarakiza (2011)

La vitesse moyenne en chaque point est représentée par une flèche de longueur proportionnelle à sa valeur, et le chemin d'écoulement déduit de ce champ des vitesses y est superposé.

La distance L_P indiquée sur la Figure 3.5 a été choisie comme largeur de référence du chemin d'écoulement, compte-tenu de l'allure du champ des vitesses.

Comme elle est positionnée sur la ligne reliant les centres des deux demi-tubes adjacents, de par la géométrie d'un faisceau de tubes de configuration triangulaire tourné, définie à la Figure 2.5 a), on a la relation :

$$L_P = P - D \quad (3.2)$$

On gardera donc la valeur : $L_P = 19 \text{ mm}$ comme référence pour la largeur du chemin d'écoulement. En notant L_{pr} la profondeur de la section d'essai, l'aire de passage de l'écoulement est donc : $A = L_{pr} \times L_P = 0.0019 \text{ m}^2$. C'est cette valeur qui sera utilisée pour le calcul des débits d'air et d'eau.

3.1.3 Points de mesure et paramètres mesurés

Afin de réaliser une carte d'écoulement, on veut pouvoir obtenir grâce à la sonde optique – présentée à la section 3.2 – des informations quantitatives liées au taux de vide, pour un certain nombre de points de mesure. Il faut tenir compte des limites imposées par l'équipement expérimental tout en faisant en sorte que la plage choisie couvre différentes configurations supposées.

Comme un grand nombre d'auteurs –cf Chapitre 2, on choisit de représenter la carte d'écoulement dans le repère des vitesses superficielles des phases gazeuse et liquide : (j_g, j_l) . Ne sachant pas à l'avance quelles configurations on obtiendra exactement et dans quelle plage de vitesses, on choisit comme point de départ les frontières de la carte de Noghrehkar et al. (1999), établie de façon plus objective que la plupart des autres cartes pour les faisceaux de tubes. Les frontières affichées sont celles obtenues pour le faisceau triangulaire normal (appelé *staggered* par les auteurs) représenté à la Figure 3.6, dont la géométrie est assez proche de la section d'essai utilisée pour ce mémoire.

Ce faisceau de tubes a un rapport pas sur diamètre $P/D=1.47$, pour un diamètre extérieur de tubes de 12.7 mm, la profondeur de la section étant 200 mm et la largeur du faisceau environ 100mm. Il comporte 26 rangées de tubes ainsi que des demi-tubes aux parois.

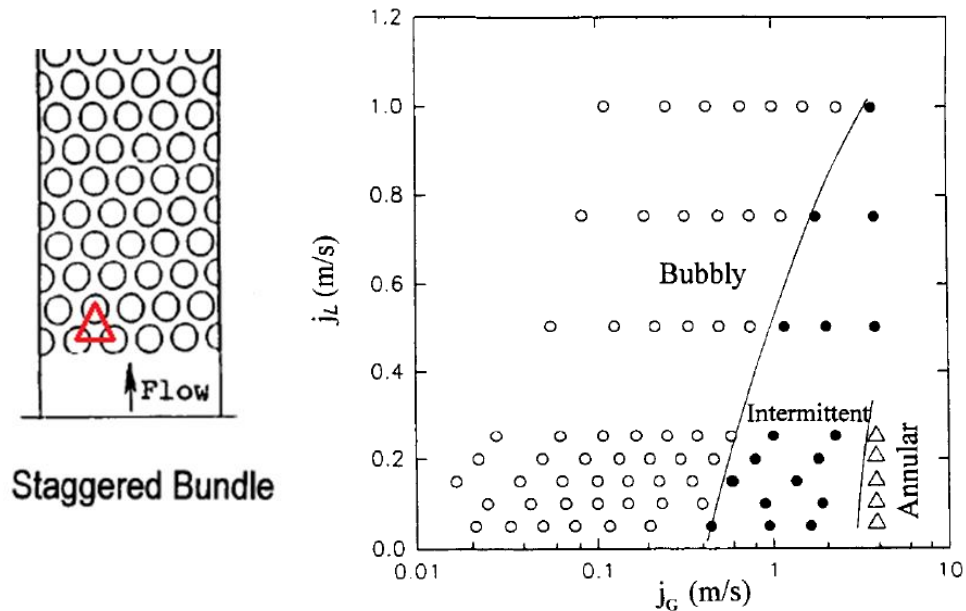


Figure 3.6 : Carte de Noghrehkar et al. pour le faisceau triangulaire normal

Comme Noghrehkar et al., on choisit comme paramètres d'entrée le titre volumique β et la vitesse superficielle de liquide j_l . Sachant que la pompe impose une vitesse superficielle de liquide maximale de 0.79 m/s pour les taux de vide envisagés, on a choisi de faire varier ces paramètres dans les plages suivantes :

$$\beta \in [0.12 - 0.97] \%$$

$$j_l \in [0.1 - 0.75] \text{ m/s}$$

Les points de mesure effectués sont répertoriés dans le Tableau 3.3 et représentés à la Figure 3.7.

Tableau 3.3 : Couples (β, j_l) choisis et numéros des points correspondants

| # des points | Titre volumique - modèle homogène | | | | | | | | | | | |
|----------------|-----------------------------------|------|------|------|------|------|------|------|------|------|------|------|
| | β | 0.12 | 0.23 | 0.34 | 0.45 | 0.55 | 0.64 | 0.74 | 0.83 | 0.92 | 0.96 | 0.97 |
| j_l (m/s) | 0.75 | 11 | 22 | 33 | 44 | 55 | 66 | 77 | 88 | 99 | - | - |
| | 0.68 | 10 | 21 | 32 | 43 | 54 | 65 | 76 | 87 | 98 | - | - |
| | 0.60 | 9 | 20 | 31 | 42 | 53 | 64 | 75 | 86 | 97 | - | - |
| | 0.50 | 8 | 19 | 30 | 41 | 52 | 63 | 74 | 85 | 96 | - | - |
| | 0.40 | 7 | 18 | 29 | 40 | 51 | 62 | 73 | 84 | 95 | 106 | - |
| | 0.35 | 6 | 17 | 28 | 39 | 50 | 61 | 72 | 83 | 94 | 105 | - |
| | 0.30 | 5 | 16 | 27 | 38 | 49 | 60 | 71 | 82 | 93 | 104 | - |
| | 0.25 | 4 | 15 | 26 | 37 | 48 | 59 | 70 | 81 | 92 | 103 | 110 |
| | 0.20 | 3 | 14 | 25 | 36 | 47 | 58 | 69 | 80 | 91 | 102 | 109 |
| | 0.15 | 2 | 13 | 24 | 35 | 46 | 57 | 68 | 79 | 90 | 101 | 108 |
| | 0.10 | 1 | 12 | 23 | 34 | 45 | 56 | 67 | 78 | 89 | 100 | 107 |

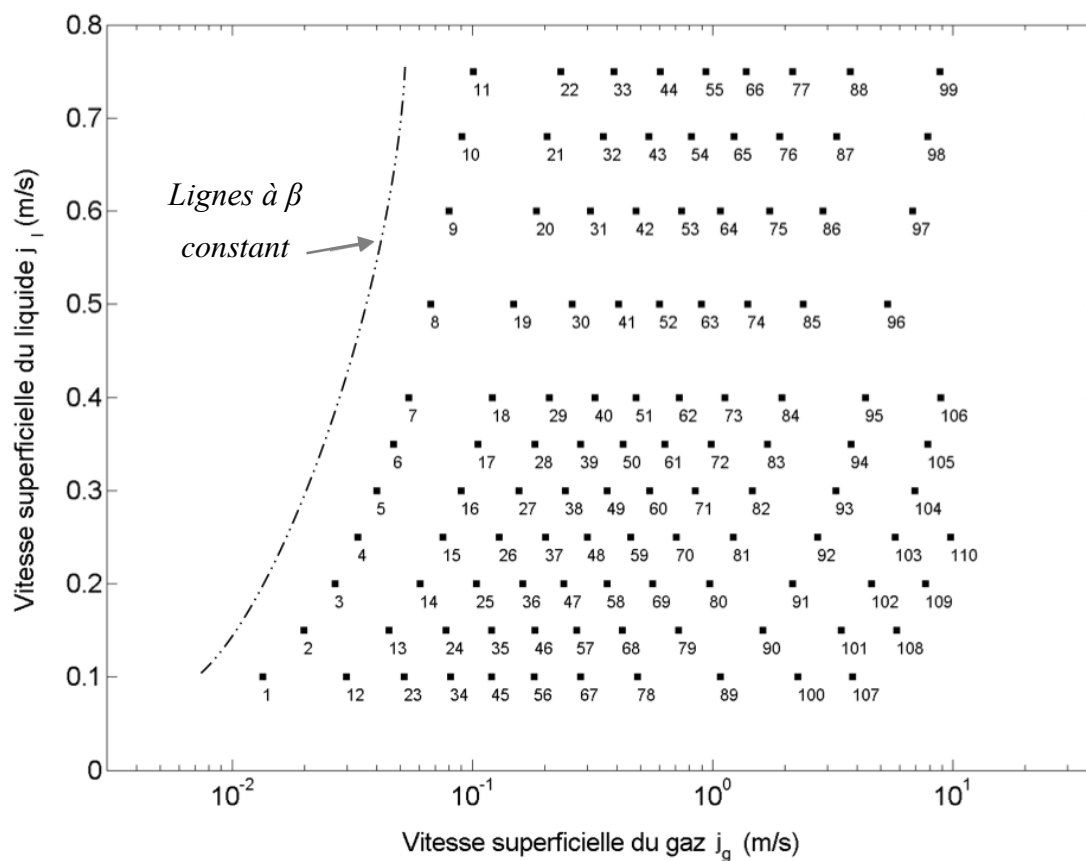


Figure 3.7 : Position des points de mesure effectués

Cela donne la plage suivante sur la carte de Noghrehkar et al. :

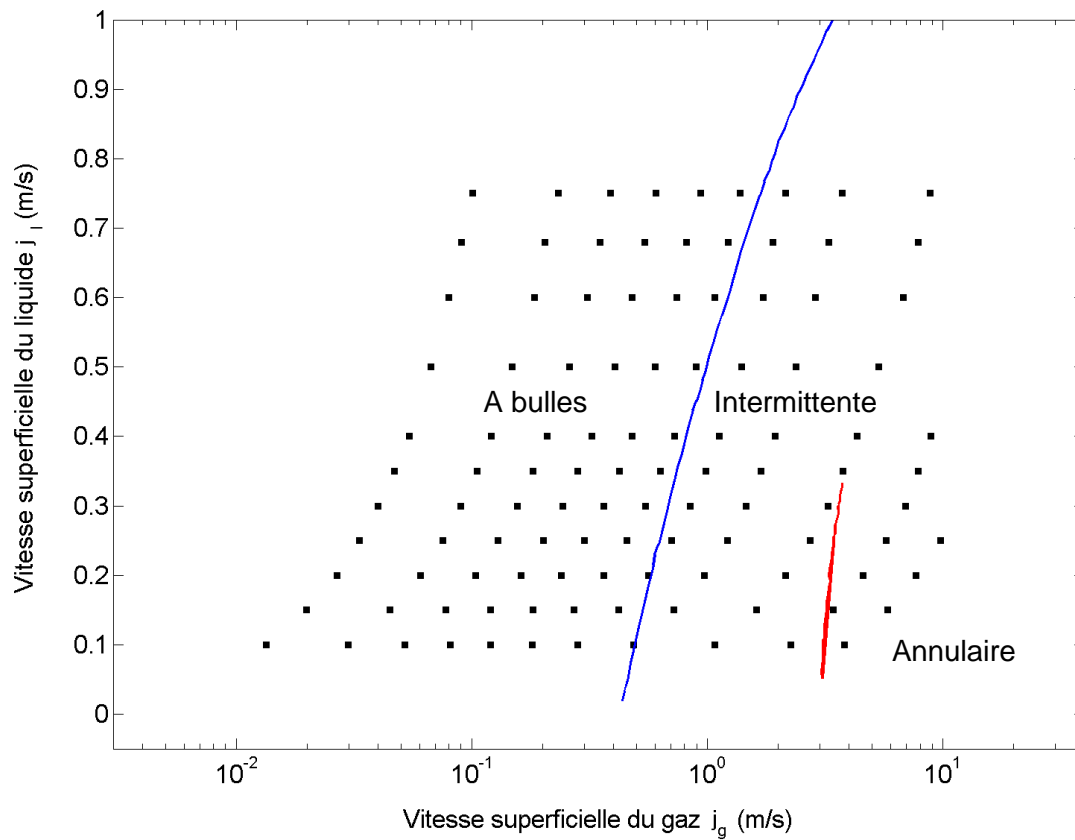


Figure 3.8 : Position des points de mesure par rapport aux frontières de Noghrehkar et al.

On devrait donc bien pouvoir observer différentes configurations dans la section d'essai.

3.1.4 Prise d'images par caméra haute vitesse

En plus de mesurer l'évolution temporelle du signal donné par la sonde optique, des images de l'écoulement ont été prises par caméra haute vitesse. Il est en effet nécessaire de fournir à la SVM des points de référence dont on connaît la configuration, ou plus exactement sur lesquels on fait l'hypothèse d'une configuration donnée. Pour cela, il est nécessaire de s'appuyer sur des observations visuelles, malgré leur caractère subjectif et le problème des différences entre la configuration observée à la paroi et celle qui existe réellement dans le faisceau, relevé par Noghrehkar et al. (1999). Ce point sera discuté plus en détail en section 6.2.2.

Étant donné les vitesses importantes pour une grande partie des points, il est difficile de distinguer réellement la structure de l'écoulement à l'œil nu. On a donc utilisé la caméra MotionBLITZ Cube de Mikrotron (modèle : CFH24HA-1), qui permet de prendre des images à haute fréquence, dans ce cas 200 images par secondes, pendant 30s. Ces images permettent d'étudier la distribution instantanée des deux phases, ainsi que d'observer le comportement dynamique au ralenti par le biais de vidéos.

La Figure 3.2 montre le positionnement de la caméra et des projecteurs, qui permettent d'obtenir des images de la zone située entre le capteur de pression et la sonde optique.

3.1.5 Ajustement des débits

Comme expliqué dans la section 2.1.2, ce sont les débits d'air, Q_g , et d'eau, Q_l , qui permettent de régler les conditions expérimentales pour qu'elles correspondent à un titre volumique β et une vitesse j_l donnés.

Le principe de réglage de chaque débit est détaillé en Annexe A, mais il faut savoir qu'une modification du débit d'air affecte le débit d'eau et vice-versa. En pratique, on a d'abord fixé le débit d'eau voulu, puis après stabilisation de celui-ci, effectué le processus itératif de réglage du débit d'air pour tenir compte de la pression relative à l'intérieur de la section d'essai. Il faut alors vérifier que le débit d'eau n'a pas trop varié, et si nécessaire, recommencer le réglage du débit d'eau puis vérifier la valeur du débit d'air, et ainsi de suite jusqu'à obtenir un réglage satisfaisant.

3.2 Instrument de mesure principal : sonde optique

Principe de fonctionnement

Une sonde optique est composée d'une ou plusieurs fibres optiques, qui permettent de détecter les changements de phase. Un signal laser est émis à l'entrée de la fibre par une diode, avec un angle assez grand pour permettre la propagation du signal dans la fibre.

Celle-ci est composée de couches concentriques successives : une enveloppe protectrice, une gaine, et un cœur en silice. La Chaire de Recherche utilise des fibres multimodes à saut d'indice : la gaine possède un indice de réfraction inférieur à celui du cœur, ce qui permet au signal laser de se propager à l'intérieur du cœur par réflexion totale interne. Il existe également une différence entre les indices de réfraction du cœur de la fibre d'une part : $n_{cœur}=1.46$, et de l'eau et de l'air dans lequel elle plongée, d'autre part : $n_{eau}=1.33$ et $n_{air}=1$. Ainsi, lorsque la fibre est plongée dans de l'air, dont l'indice de réfraction est bien inférieur à celui du cœur de la fibre, si l'angle d'incidence est suffisamment grand, le signal va être totalement réfléchi et l'intensité lumineuse captée en sortie de la fibre sera élevée - cf Figure 3.9, a).

Inversement, lorsque la fibre est plongée dans de l'eau, dont l'indice de réfraction est légèrement inférieur à celui du cœur, le signal va être réfracté dans l'eau et l'intensité lumineuse captée à la sortie sera quasiment nulle - cf Figure 3.9, b).

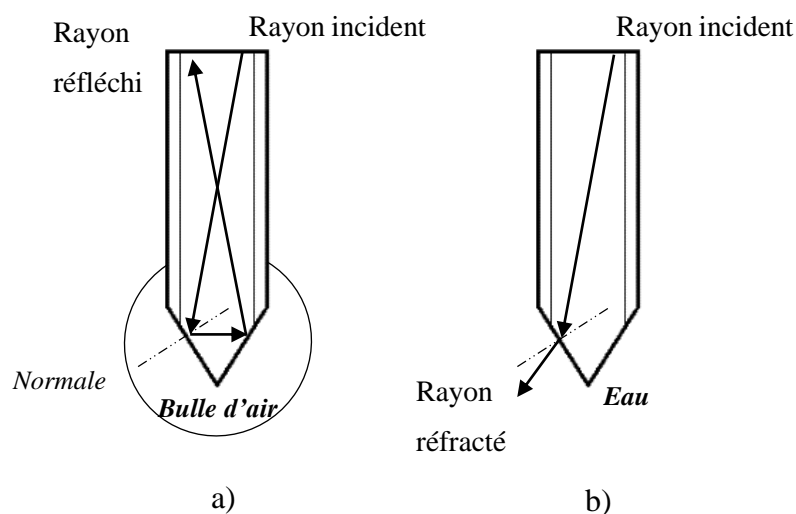


Figure 3.9 : Principe de fonctionnement de la fibre optique :
a) Réflexion totale dans l'air et b) Réfraction dans l'eau

Après traduction du signal lumineux obtenu en sortie en signal électrique, on obtient donc un signal de type échelon qui permet de différencier la phase liquide (amplitude la plus faible) de la phase gazeuse (amplitude la plus élevée), comme l'illustre la Figure 3.10, où l'on peut voir le signal affiché par un oscilloscope, obtenu en trempant plusieurs fois l'extrémité de la sonde dans un béccher rempli d'eau :

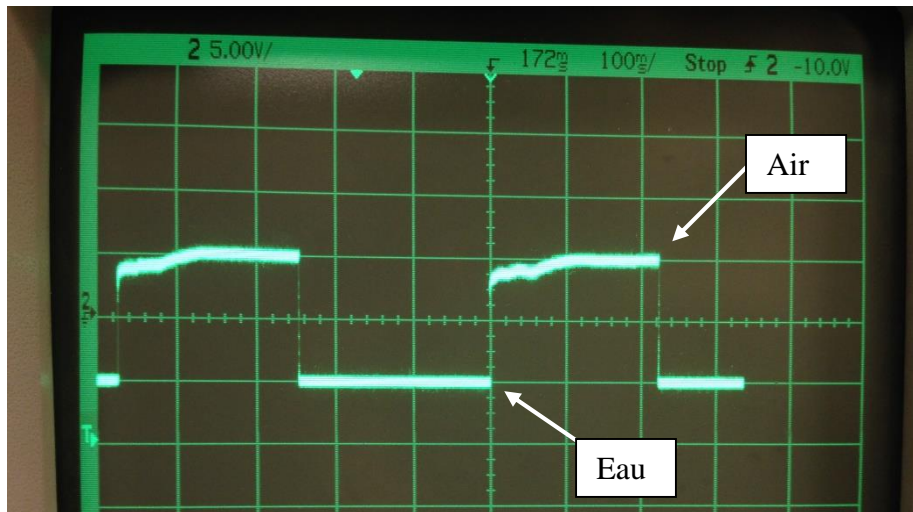


Figure 3.10 : Exemple de signal échelon donné par la sonde optique

Ce signal est ensuite traduit par une application développée au sein de la Chaire sous LabVIEW, qui permet d'obtenir un pur signal échelon ramené sur l'intervalle $[0;1]$ en amplitude (se rapporter à la section 5.1.1 pour plus de détails). La Figure 3.11 illustre la correspondance entre une succession de bulles de l'écoulement et le signal finalement obtenu, reconstruit d'après les instants auxquels une bulle touche la sonde : t_i et les durées de chaque bulle : τ_i .

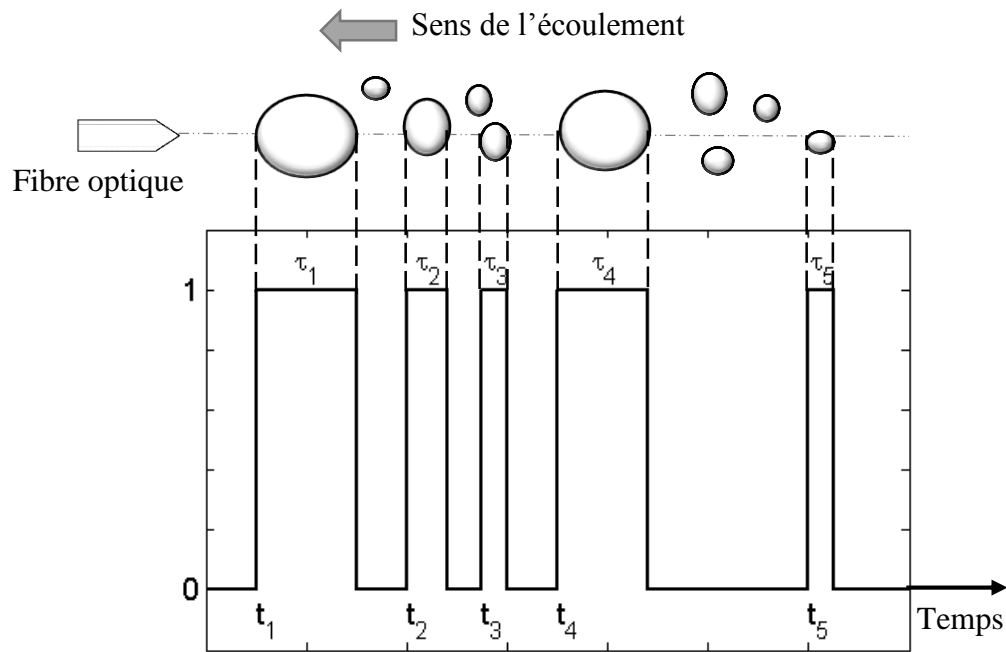


Figure 3.11 : Correspondance entre l'écoulement diphasique et le signal donné par la sonde optique

Pour mesurer le taux de vide local ε , on peut donc rapporter le ratio des volumes de chaque phase à un ratio de temps passé dans chaque phase :

$$\varepsilon_{local} = \frac{V_g}{V_g + V_l} = \frac{T_g}{T_g + T_l} \quad (3.3)$$

Ce qui donnerait par exemple sur l'intervalle de temps T affiché à la Figure 3.11, dans lequel la sonde capte 5 bulles :

$$\varepsilon_{local} = \frac{1}{T} \sum_{i=1}^5 \tau_i \quad (3.4)$$

La sonde optique permet ainsi de mesurer directement le taux de vide local, noté ε par la suite.

Pour l'étude des configurations de l'écoulement, on fait l'hypothèse, en première approximation, que le taux de vide *local* restitue correctement l'information sur la configuration *globale* de l'écoulement. Cette approximation est notamment exacte dans le cas d'une configuration à bulles, où l'on peut raisonnablement supposer que les bulles sont réparties de façon homogène sur la section de passage de l'écoulement A_P .

Caractéristiques de la sonde utilisée

Une seule fibre optique est nécessaire pour obtenir un signal dont on peut déduire le taux de vide, directement lié à la configuration de l'écoulement. La sonde est placée à 1.15m au-dessus du bas de la section, et au milieu de la section dans la direction de la profondeur. Dans le plan des observations visuelles, l'extrémité de la sonde optique se trouve juste en-dessous de la distance L_P (Figure 3.5) et a été positionnée soigneusement de façon à être parallèle à l'écoulement, en s'aidant du champ des vitesses de la Figure 3.5, pour optimiser la détection des phases.

Les détails de la fabrication de la sonde et du dispositif utilisé pour générer et mesurer le signal de sortie, ainsi que du positionnement de la sonde dans la section, sont donnés en Annexe A.

Pour tous les points de mesure, on a fait l'acquisition de 60s de signal donné par la sonde, échantillonné à 50 kHz (valeur standard utilisée dans la Chaire lorsqu'on s'intéresse seulement au taux de vide avec une seule fibre optique).

CHAPITRE 4 RÉSULTATS BRUTS : FORME DE L'ÉCOULEMENT ET DONNÉES DE SONDE OPTIQUE

Ce chapitre présente les résultats bruts obtenus à l'issue de la phase de mesure : les images de l'écoulement, d'après lesquelles on peut déjà identifier et décrire des configurations, et les signaux donnés par la sonde optique.

4.1 Images de l'écoulement et formes observées

Pour chaque point de mesure, la caméra haute vitesse a permis d'obtenir 6000 images réparties sur 30s d'acquisition. La zone montrée sur les photos se situe au-dessus du capteur de pression et sous la sonde optique comme le montre la Figure 4.1. Les Figure 4.2 et Figure 4.3 montrent les différentes formes qu'on peut observer selon les valeurs de β pour une valeur donnée de j_i .

On peut observer un ensemble plutôt homogène de petites bulles - Figure 4.2 a) à basse vitesse de gaz, mais aussi de plus grosses bulles comme celles qu'on distingue à la Figure 4.2 b) et c) lorsque j_g augmente. A haut débit de gaz, on observe une alternance entre des structures beaucoup plus grandes dans un écoulement plus agité comme à la Figure 4.3. a), et de nouveau une succession de petites bulles - Figure 4.3 b).

Quelles que soient les conditions, on observe une zone de recirculation à l'arrière de chaque tube où se trouvent toujours des petites bulles dans du liquide.

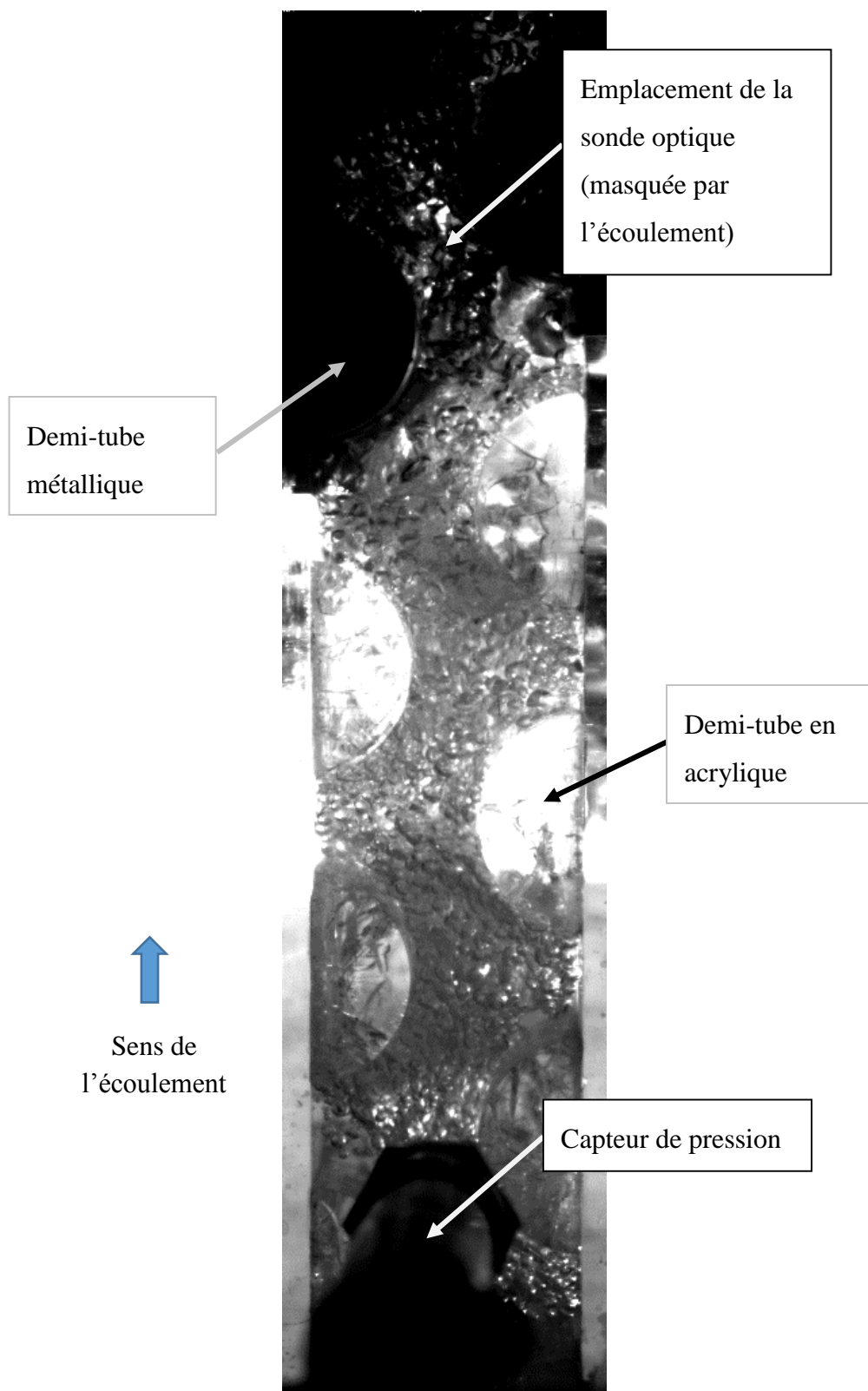


Figure 4.1 : Zone observée avec la caméra

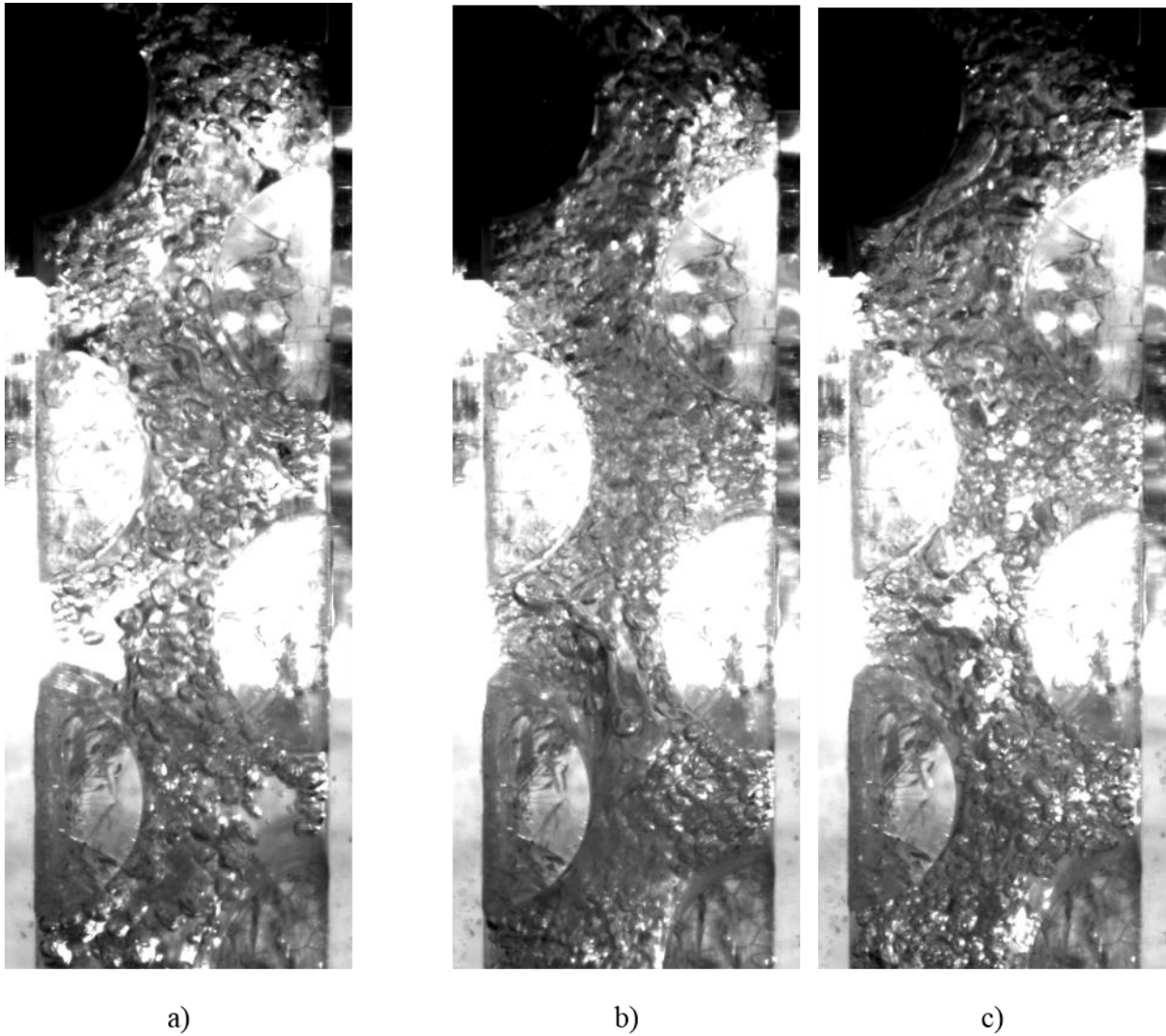


Figure 4.2 : Exemples d'images obtenues à différents instants et β , pour $j_i=0.3$ m/s
a) $\beta=12\%$, b) et c) $\beta=55\%$

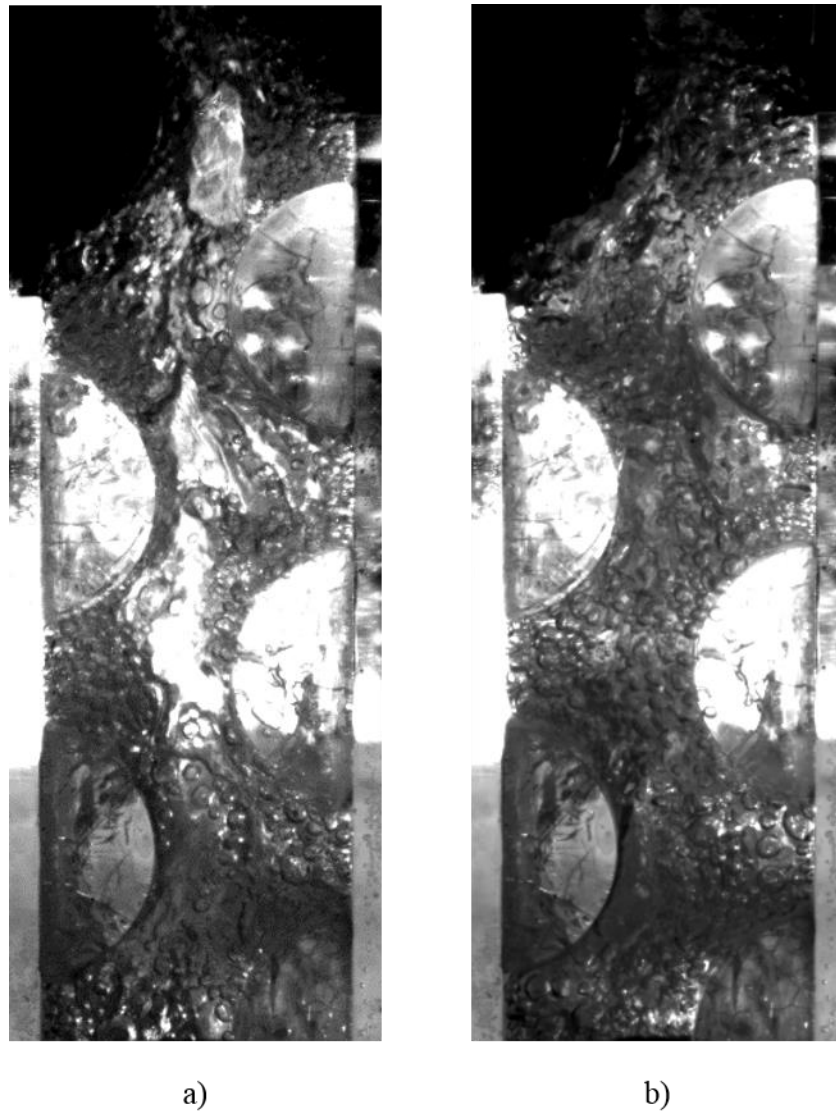


Figure 4.3 : Exemples d'images obtenues à différents instants pour $j_l=0.3$ m/s et $\beta=96\%$

4.2 Définition des configurations observées

En suivant pour le moment la méthode la plus ancienne d'identification des configurations, c'est-à-dire de façon purement visuelle depuis l'extérieur de la section d'essai, on peut déjà définir des configurations d'écoulement.

La revue de littérature a montré que dans les faisceaux de tubes, on observe toujours les mêmes configurations : à bulles, intermittente, et annulaire/dispersée, même si selon les auteurs, un même terme ne recouvre pas toujours exactement le même type d'écoulement. Globalement, c'est principalement l'augmentation du débit de gaz qui entraîne le passage d'une configuration à une autre, dans l'ordre cité.

Une comparaison des photos de l'écoulement prises dans la section d'essai avec une synthèse de caractéristiques visuelles données par Ulbrich et Mewes (1994) et Kondo et Nakajima (1980) (ici traduites en français) est donnée pour chaque configuration dans les Tableau 4.1, Tableau 4.2 et Tableau 4.3. La description de chaque configuration donnée dans le corps du texte correspond à ce qui a été observé dans la section d'essai utilisée pour cette étude, en utilisant conjointement les images et les vidéos (créées à partir de celles-ci).

Bien que les plages de vitesses de Kondo et Nakajima soient bien inférieures à celles de cette étude, le lien entre le nom donné aux configurations et les observations rattachées a paru constituer une référence intéressante. Dans tous les cas, la section d'essai ici employée est différente d'un faisceau de tubes complet, au-delà des divergences de conditions expérimentales, mais on a constaté expérimentalement que les configurations observées sont très proches de celles des faisceaux.

La **configuration à bulles** est caractérisée par un écoulement uniforme à bulles de tailles variées mais très inférieures au diamètre des tubes, strictement ascendant. C'est ce qui est observé sans conteste pour certains points pour lesquels la structure à un instant quelconque est semblable à celle de la photo du Tableau 4.1.

La **configuration intermittente** est caractérisée par un écoulement irrégulier, on observe à la fois un écoulement assez uniforme à bulles par moments, mais qui présente une oscillation locale des bulles, avec des bulles de taille plus variée, et le passage de structures gazeuses de taille au moins équivalente au diamètre des tubes, et de largeur équivalente à l'espace entre les tubes, comme on peut le voir sur la photo du Tableau 4.2 (mais qui ne se distinguent vraiment bien que sur les vidéos, ou en naviguant entre deux images successives ce qui permet de « voir » la structure se déplacer légèrement donc de définir ses frontières). Ces structures semblent contenir quelques gouttes liquides.

Tableau 4.1 : Caractéristiques visuelles de la configuration à bulles

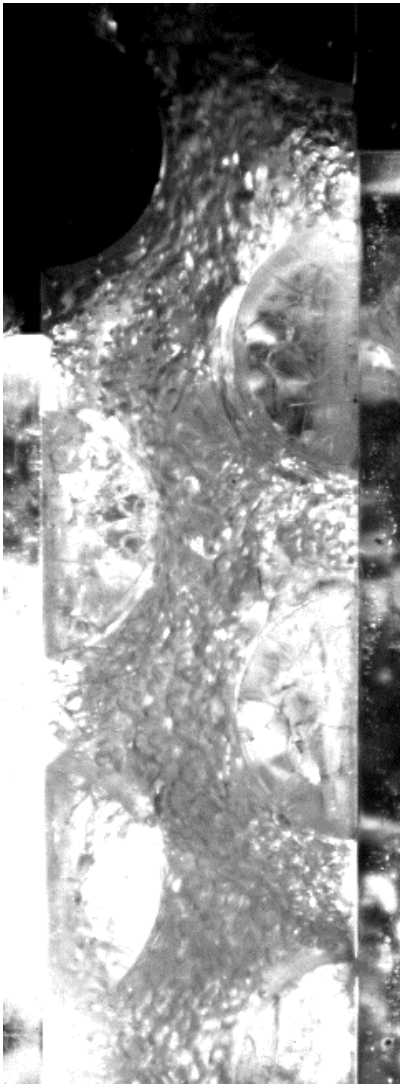
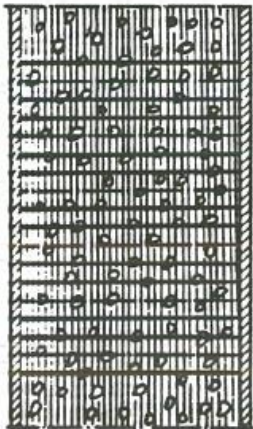
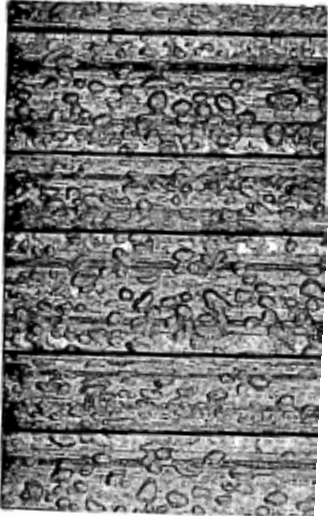
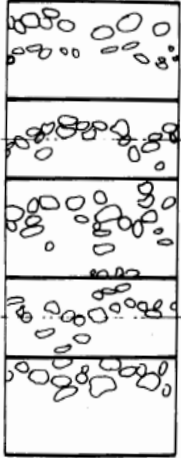
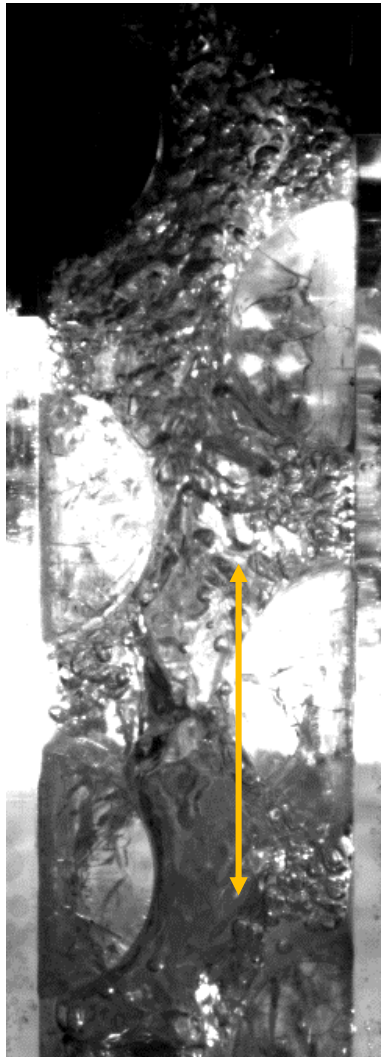

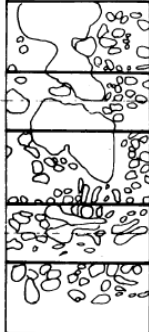
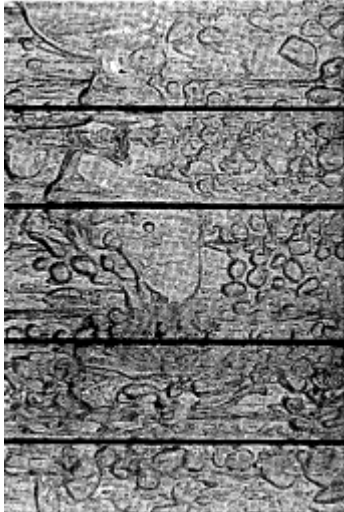
| Image prise par caméra HV | Ulbrich et Mewes (1994) | Kondo et Nakajima (1980) |
|---|--|--|
|  <p data-bbox="342 1591 456 1623">Point #9</p> <p data-bbox="277 1661 521 1696">$\beta=12\%$, $j_l=0.6$ m/s</p> | <p data-bbox="678 365 911 401">« <i>Bubbly flow</i> » :</p> <p data-bbox="621 436 967 852">« Phase liquide continue majoritaire (sans oscillations locales) et phase gazeuse dispersée en petites bulles dont le diamètre est plus petit que l'espace caractéristique entre les tubes. »</p>  <p data-bbox="802 1423 821 1451">B</p> | <p data-bbox="992 365 1422 730">« <i>Bubbly flow</i> » : « Pour de faibles débits d'air, les diamètres des bulles sont beaucoup plus petits que le diamètre extérieur des tubes, et les bulles sont distribuées uniformément au sein de la phase liquide. »</p>   |

Tableau 4.2 : Caractéristiques visuelles de la configuration intermittente

| Image prise par caméra HV | Ulbrich et Mewes (1994) | Kondo et Nakajima (1980) |
|---|--|---|
| <p data-bbox="207 415 591 562"><i>La flèche repère la longueur approximative de la structure gazeuse adjacente.</i></p>  <p data-bbox="337 1734 461 1766">Point #48</p> <p data-bbox="266 1803 529 1839">$\beta=55\%$, $j_l=0.25$ m/s</p> | <p data-bbox="618 386 967 1146">« <i>Intermittent flow</i> » : « mouvement oscillant irrégulier du liquide et du gaz. Des poches de liquide peuvent se déplacer vers le bas, et les bulles de gaz peuvent prendre des formes irrégulières dont la hauteur est plusieurs fois plus grande que le diamètre des tubes, et de largeur équivalente à celle de l'espace inter-tubes.»</p>  <p data-bbox="808 1707 824 1728">I</p> | <p data-bbox="984 386 1425 974">« <i>Slug flow</i> » : « Lorsque le débit d'air augmente, une partie des bulles grandit et remplit l'espace entre les tubes. La hauteur et la longueur de ces bulles est comparable au pas du faisceau. Dans cet intervalle de débits d'air, l'écoulement contient des bulles de tailles différentes. Certaines bulles de petite taille se déplacent en sens inverse à cause du déplacement des grosses bulles dans le sens de l'écoulement.»</p>   |

La configuration observée à plus hauts débits a été nommée **semi-annulaire** car dans les conditions étudiées, il est apparu qu'on n'a pas atteint un débit de gaz suffisant pour observer une phase gazeuse continue au centre de la section d'essai, caractéristique de la configuration annulaire.

La configuration annulaire, ou dispersée, est en effet décrite par Ulbrich et Mewes (1994) comme comportant des « gouttes régulièrement dispersées transportées par le gaz. Cette configuration existe d'abord au-dessus du faisceau puis dans le faisceau. » Ils distinguent deux sous-configurations : la configuration intermittente-dispersée (ID) : « En plus des gouttes dispersées, on a des poches de liquide de forme et mouvement irréguliers » et la configuration annulaire-dispersée (AD) caractérisée par un « film mince de liquide avec des « vagues de surface » qui occupent la paroi des tubes ou de la coque». Les schémas correspondants sont donnés Figure 4.4.

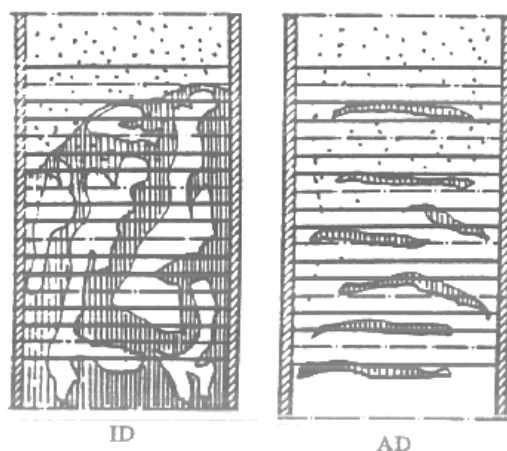


Figure 4.4 : Configurations intermittente-dispersée (ID) et annulaire dispersée (AD)
selon Ulbrich et Mewes (1994)

En revanche, la description de la configuration nommée « *froth flow* » par Kondo et Nakajima, qui est la configuration de transition entre les configurations intermittente et annulaire, semble correspondre assez bien à ce qu'on observe expérimentalement. En effet dans la section d'essai, pour les débits d'air les plus élevés, l'écoulement devient très agité (avec une oscillation globale de l'écoulement dans la plupart des cas, c'est-à-dire à l'échelle de toute la section) et on observe une alternance entre un écoulement plutôt à bulles (qui peuvent être assez petites), et de grandes structures gazeuses, de longueur au moins égale à 2-3 fois le diamètre des tubes.

Ces structures se distinguent particulièrement des zones de recirculation qui se trouvent à l'arrière des tubes, et comportent toujours des bulles d'air au sein de liquide.

Le terme semi-annulaire a été choisi pour cette configuration car il désigne souvent la configuration de transition entre les configurations intermittente (à bouchons instables dans le cas d'un tube) et annulaire, et on semble observer un début de configuration annulaire sur certaines portions de l'écoulement comme le montre la Figure 4.5 (qui a été prise avec un appareil photo standard contrairement aux autres images). On évite ainsi l'ambiguïté liée au terme « agité » (*churn*), qui est utilisé pour décrire des écoulements parfois assez différents (Rosa et al. 2010).

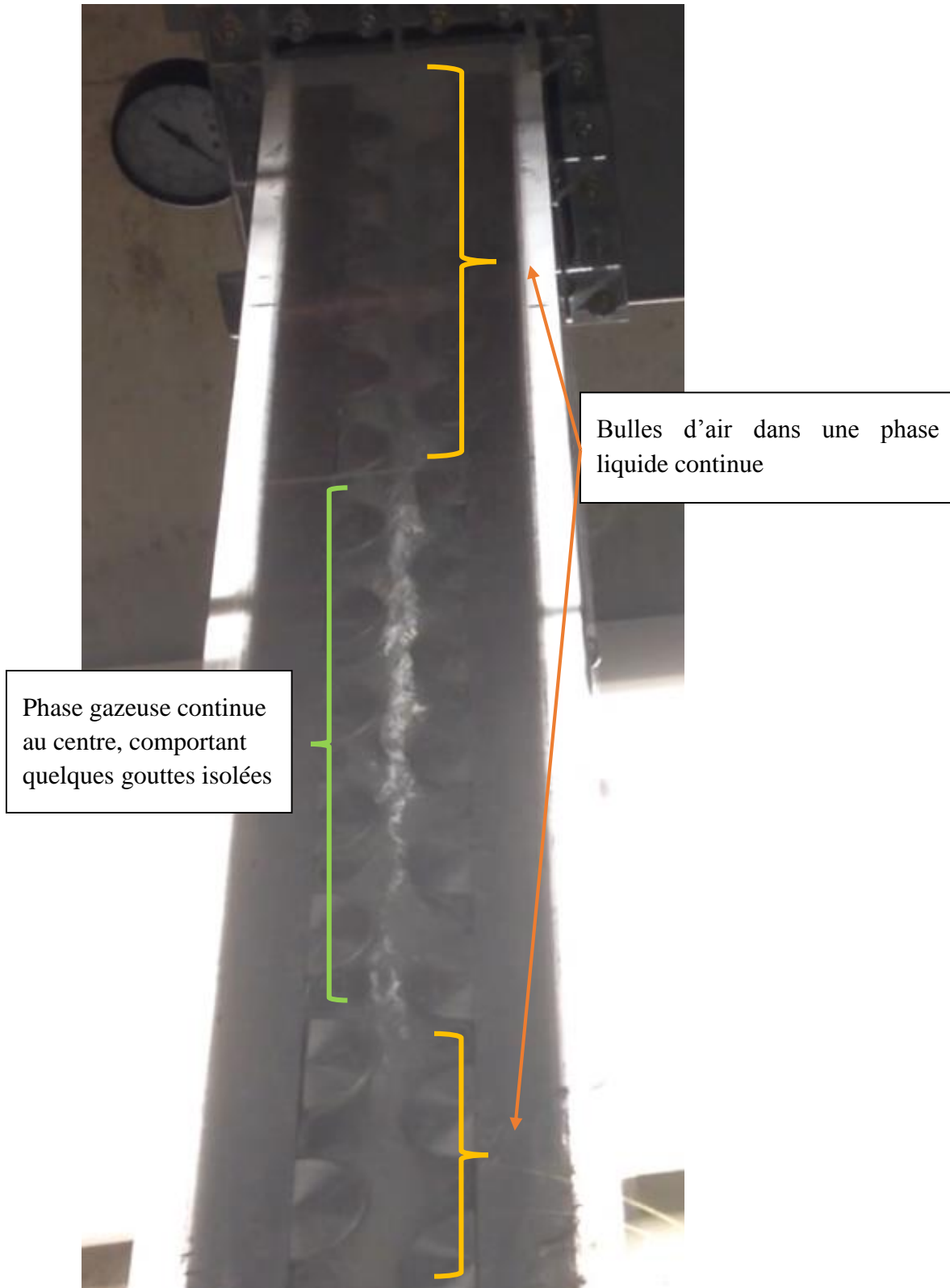
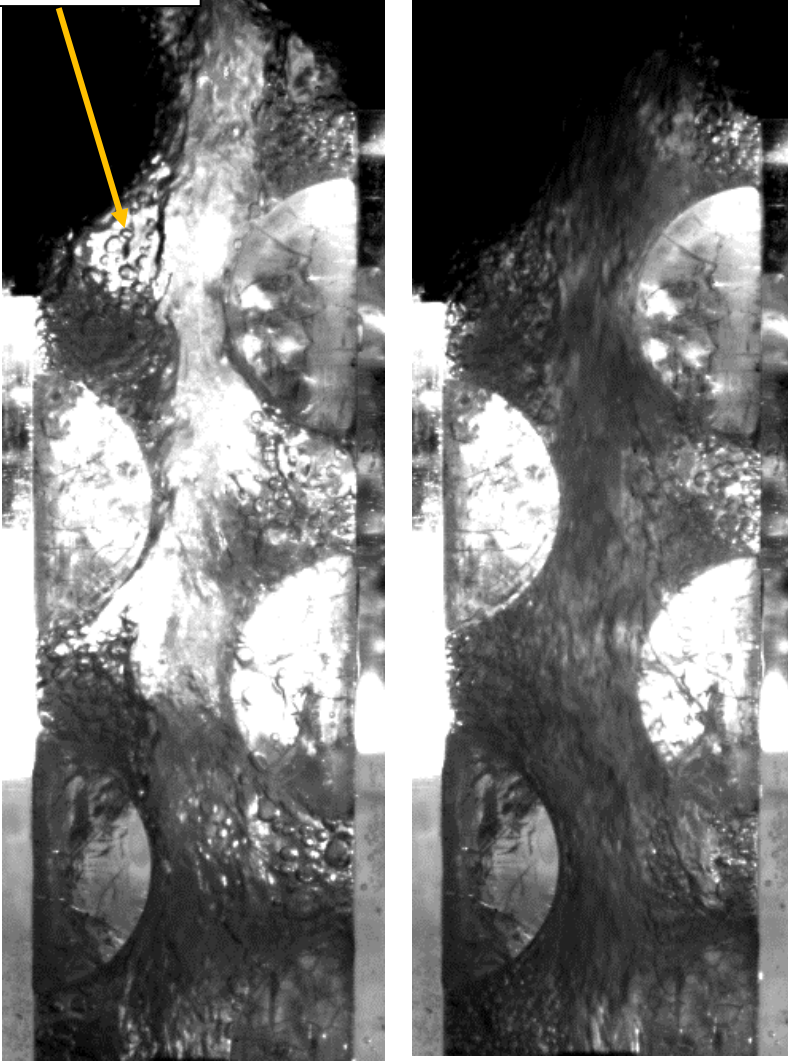


Figure 4.5 : Exemple typique de la configuration semi-annulaire

Tableau 4.3 : Caractéristiques visuelles de la configuration semi-annulaire

| Images prises par caméra HV (à deux instants différents) | | | Kondo et Nakajima (1980) |
|---|--|--|---|
| Zone de recirculation : liquide + bulles | Phase principalement gazeuse au centre | Mélange plus difficilement identifiable (vitesse élevée), a priori bulles au sein du liquide | <p>« <i>Froth flow</i> » : « Lorsque le débit d'air est élevé, la hauteur des grosses bulles s'étend sur plusieurs fois la taille du pas du faisceau et leur largeur est équivalente à l'espace entre les tubes. Le mouvement à contre-courant de certaines petites bulles devient plus fort et le mélange est agité par des pulsations violentes. La phase gazeuse commence à occuper presque tout l'espace entre les tubes, sans jamais les recouvrir tout à fait. Cette partie de l'écoulement ressemble à celui qu'on observe en configuration annulaire dans un tube : phase gazeuse au centre et fin film de liquide sur les parois des tubes, alors que le reste est constitué d'un écoulement à bulles contenant des bulles relativement petites. »</p> |
|  | | | |
| <p>Point #106</p> <p>$\beta=96\%$, $j_l=0.4$ m/s</p> | | | |

4.3 Signaux bruts de sonde optique

La Figure 4.6 montre un exemple d'évolution de l'amplitude du signal de la sonde en fonction du temps sur 1s, pour $\beta=12\%$ et $j_l=0.3$ m/s. Les valeurs maximales correspondent aux instants où le signal laser a été réfléchi car la sonde était dans l'air, et inversement les valeurs minimales correspondent à l'eau. On peut donc voir dans ce signal une succession de bulles d'air plus ou moins grosses séparées par du liquide.

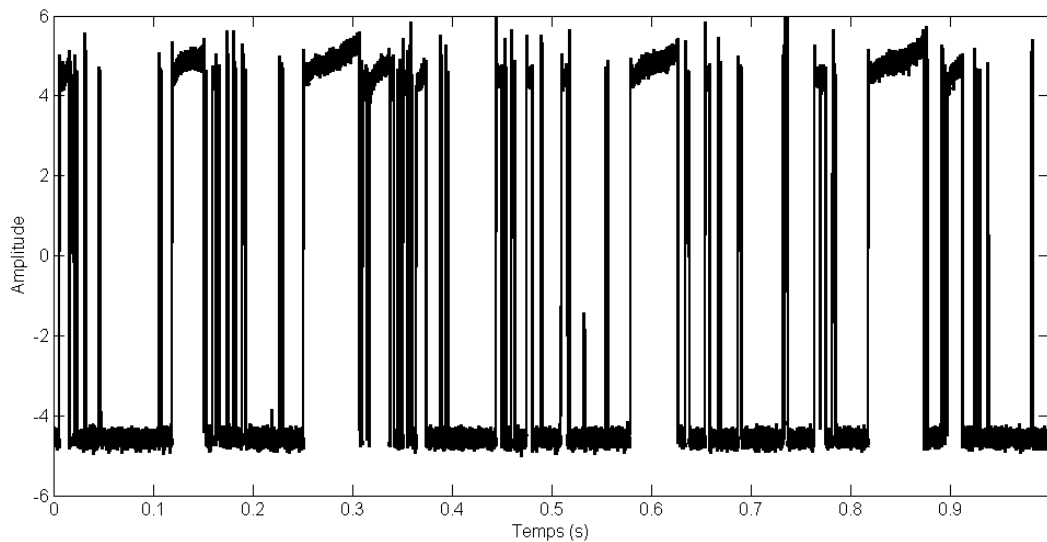


Figure 4.6 : Exemple d'évolution temporelle du signal brut

On peut déjà observer des différences dans le tracé de ces signaux lorsque le titre volumique β et la vitesse superficielle de liquide j_l varient, comme l'illustre la Figure 4.7. On y retrouve les signaux correspondant aux images de la Figure 4.2 et de la Figure 4.3 pour $\beta=12\%$, 55% et 96%.

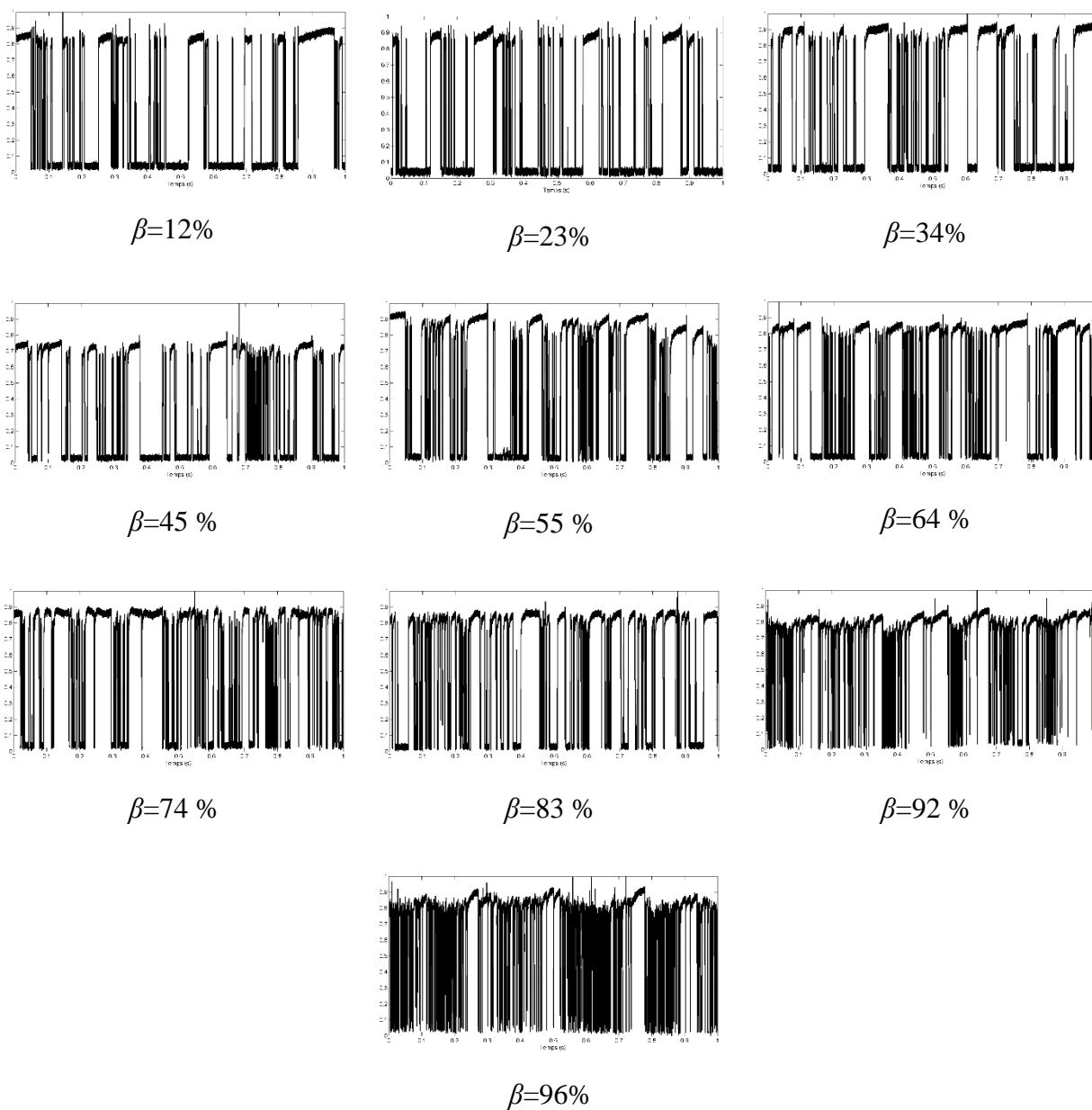


Figure 4.7 : Évolution du tracé du signal brut sur 1s, pour $j_l=0.3$ m/s et différentes valeurs de β

La Figure 4.7 permet de voir l'évolution de l'allure du signal par rapport à une même durée, mais si l'on veut comparer ces signaux avec la structure observée sur les images, il faut absolument tenir compte du fait que la vitesse du mélange $j = \frac{j_l}{(1-\beta)}$ varie. On choisit donc une longueur de référence L_0 , représentée à la Figure 4.8, environ 2 fois plus grande que les structures de la configuration semi-annulaire.



Figure 4.8 : Longueur L_0

En prenant cette fois-ci non pas une durée fixe, mais la durée Δt correspondant à cette longueur fixe L_0 rapportée à chaque vitesse du mélange j : $\Delta t = \frac{L_0}{j}$, on obtient les signaux de la Figure 4.9, pour lesquels la durée des bulles affichée est représentative de leur taille réelle.

On peut voir que ces signaux reflètent bien les types de structures observées sur les images correspondantes (qui ne sont pas prises aux mêmes instants d'où les légères différences observées), selon la valeur de β : des bulles plutôt petites dans du liquide pour $\beta=12\%$, la même chose avec des bulles plus grosses pour $\beta=55\%$, et une majorité de phase gazeuse dans laquelle on trouve des gouttes d'eau pour $\beta=96\%$, qui correspondrait à une partie d'une grande structure gazeuse comme sur la Figure 4.3 a).

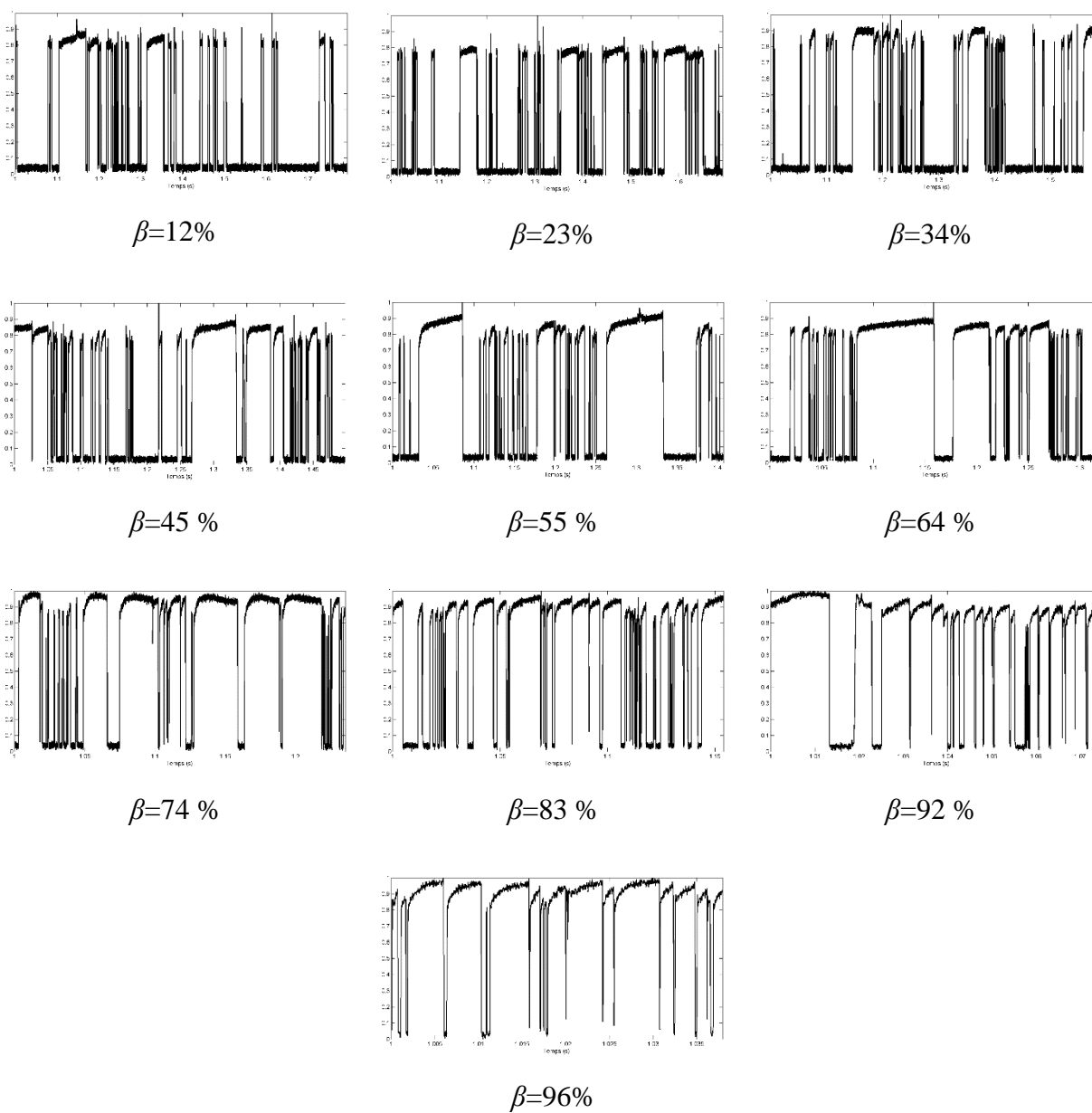


Figure 4.9 : Évolution du tracé du signal brut sur une durée variable= L_0/j
pour $j=0.3$ m/s et différentes valeurs de β

Mais telles quelles, ces données ne permettent pas vraiment une identification des frontières des différentes configurations : une classification « manuelle » est rendue difficile par la continuité que l'on observe d'un point à un autre, particulièrement pour des signaux qui semblent se situer dans une zone de transition entre deux types plus identifiables. Il serait de plus assez fastidieux de s'assurer que la zone observée est représentative de l'ensemble d'un signal donné, en plus d'être un choix subjectif.

Et une classification objective directement effectuée sur ces données ne serait pas optimale, à cause du trop grand nombre de valeurs (50 000 points pour 1 seconde de signal), même en envisageant un ré-échantillonnage.

Conclusion

Pour ces raisons, avant de pouvoir construire une SVM pour rendre plus objective la classification des points, il a été nécessaire d'effectuer un traitement du signal donné par la sonde optique pour en obtenir un nombre limité de grandeurs caractéristiques, qui soient en même temps reliées physiquement aux configurations de l'écoulement.

CHAPITRE 5 ANALYSE DES DONNÉES EXPÉRIMENTALES

Ce chapitre présente le post-traitement des signaux de sonde optique, c'est-à-dire la façon dont ils ont été analysés pour en faire ressortir des indicateurs quantitatifs liés à la configuration de l'écoulement, pouvant ensuite être utilisées par la machine à vecteurs de support (SVM) pour la classification.

5.1 Obtention des PDF du taux de vide

Dans un premier temps, pour représenter de façon plus synthétique l'information contenue dans le signal donné par la sonde, les PDF (*Probability Density Function* : fonction de densité de probabilité) du taux de vide ont été calculées et tracées.

En effet, Jones & Zuber (1975) ont montré que cette transformation du signal est une image plus explicite de l'allure de la distribution du taux de vide, qui est elle-même porteuse de l'information reliée à la configuration de l'écoulement. Les PDF du taux de vide ont par la suite été utilisées par de nombreux auteurs pour l'identification des configurations d'écoulement, parmi lesquels on peut citer Noghrehkar et al. (1999) en faisceau de tubes ou plus récemment, Rosa et al. (2010) pour un tube vertical.

5.1.1 Filtrage du signal de la sonde

Dans le but d'obtenir les PDF du taux de vide, il faut pouvoir calculer celui-ci sur une période donnée. Pour cela, on procède à un filtrage du signal brut donné par la sonde pour le traduire sous forme d'échelon, en utilisant LabVIEW.

On obtient ainsi directement les instants auxquels une bulle touche la sonde et la durée de la bulle correspondante. Cela permet de ne conserver que les bulles ayant été entièrement « mesurées », généralement celles dont la trajectoire n'a pas trop été modifiée par la présence de la sonde, comme le montre la Figure 5.1. Pour améliorer la visualisation, les aires se trouvant sous l'échelon ont été remplies, ainsi les zones foncées représentent le temps passé dans l'air (l'amplitude vaut 1), et les zones blanches le temps passé dans l'eau (l'amplitude vaut 0).

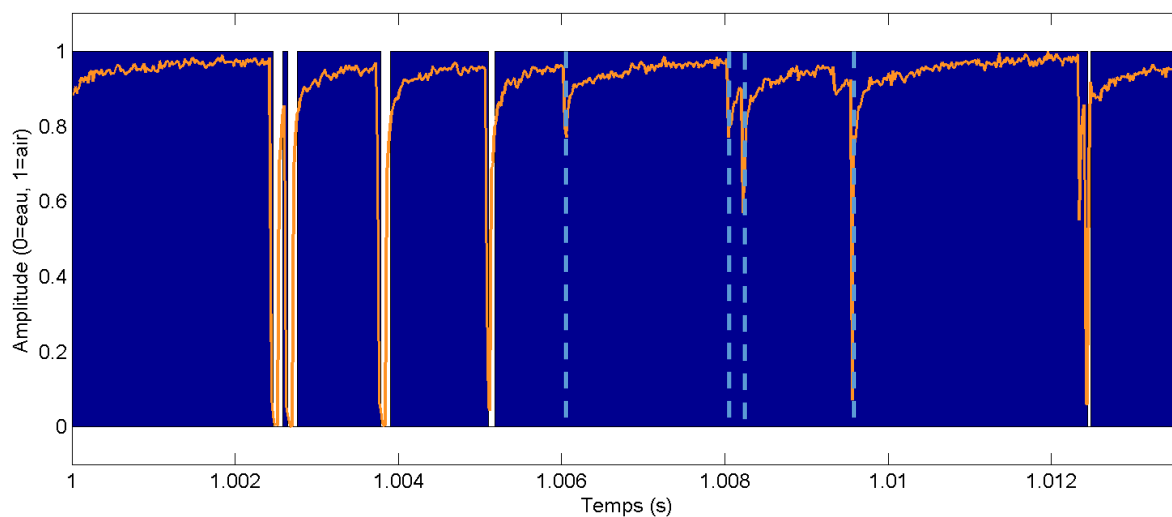


Figure 5.1 : Différences entre le signal brut (ligne claire) et le signal filtré (échelon)

Les pointillés - - correspondent aux frontières des bulles ignorées par le filtrage

Pour le signal brut de la Figure 5.2, on obtient donc le signal filtré de la Figure 5.3. L'amplitude des signaux est ramenée à l'intervalle 0-1, pour faire en sorte que la moyenne corresponde à peu près au taux de vide moyen mesuré.

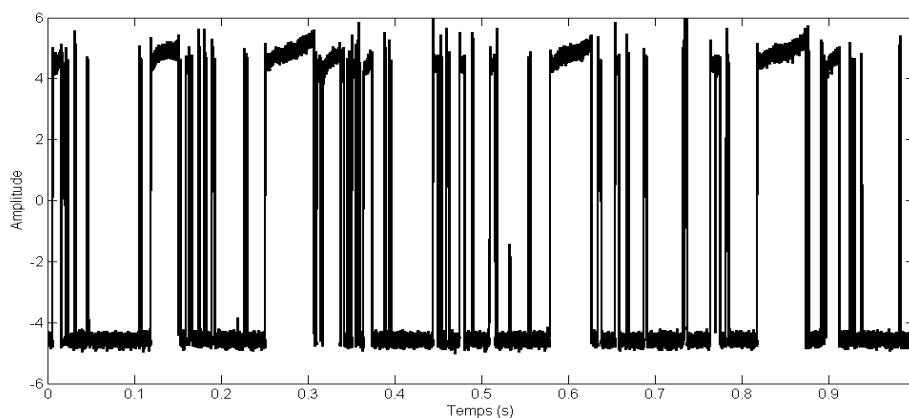


Figure 5.2 : Évolution temporelle du signal brut de la sonde ($\beta=12\%$, $j_l=0.3$ m/s)

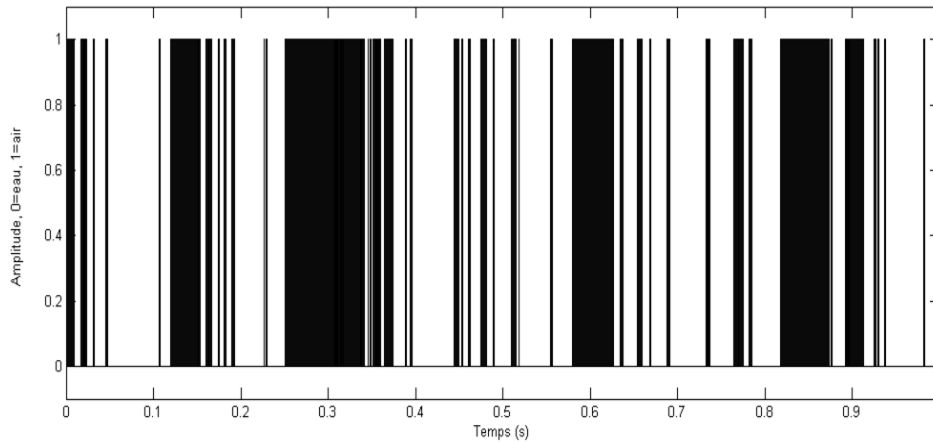


Figure 5.3 : Évolution temporelle du signal filtré correspondant ($\beta=12\%$, $j_I=0.3$ m/s)

5.1.2 Principe de calcul des PDF

L'idée des PDF est de représenter pour chaque valeur du taux de vide la probabilité de retrouver cette valeur dans le signal. Par exemple, pour un écoulement à bulles, on s'attend à avoir un pic de probabilité autour des faibles valeurs du taux de vide, alors qu'inversement, pour un écoulement annulaire, le pic de probabilité devrait être proche des hauts taux de vide.

Pour cela, on choisit tout d'abord un court intervalle de temps Δt , et on divise la durée totale d'acquisition, ici $T=60$ s, en intervalles de durée Δt , ici $\Delta t=0.015$ s, comme illustré par la Figure 5.4 :

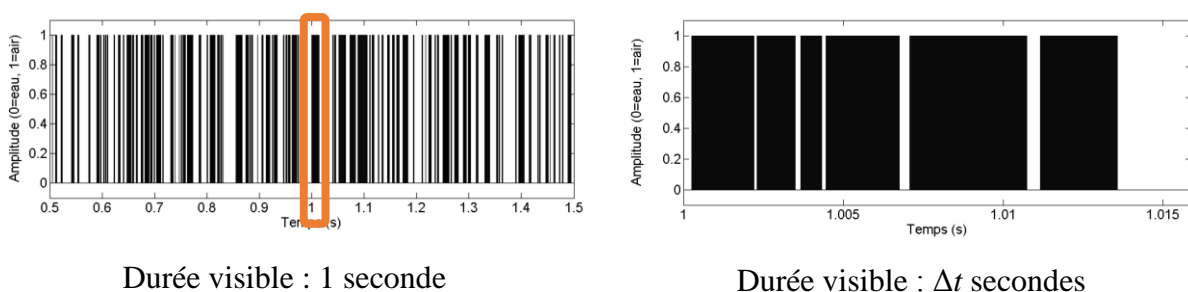


Figure 5.4 : Division du signal temporel en intervalles Δt

On calcule alors le taux de vide moyen (qui est alors quasiment « instantané » si Δt est assez petit) sur chaque portion du signal.

Comme le montre la Figure 5.5, l'utilisation du signal filtré permet de calculer directement le taux de vide, sur un intervalle de temps donné, par le biais du ratio du temps passé dans la phase gazeuse T_g par rapport au temps total (gaz+liquide) :

$$\varepsilon = \frac{T_g}{T_g + T_l} \quad (5.1)$$

où :

$$T_g = \sum_1^6 T_{gi} \quad (5.2)$$

Les zones foncées représentent le temps passé dans l'air (l'amplitude vaut 1), et les zones blanches le temps passé dans l'eau (l'amplitude vaut 0). Sur l'intervalle de temps Δt affiché à la Figure 5.5, le taux de vide moyen vaut 0.755.

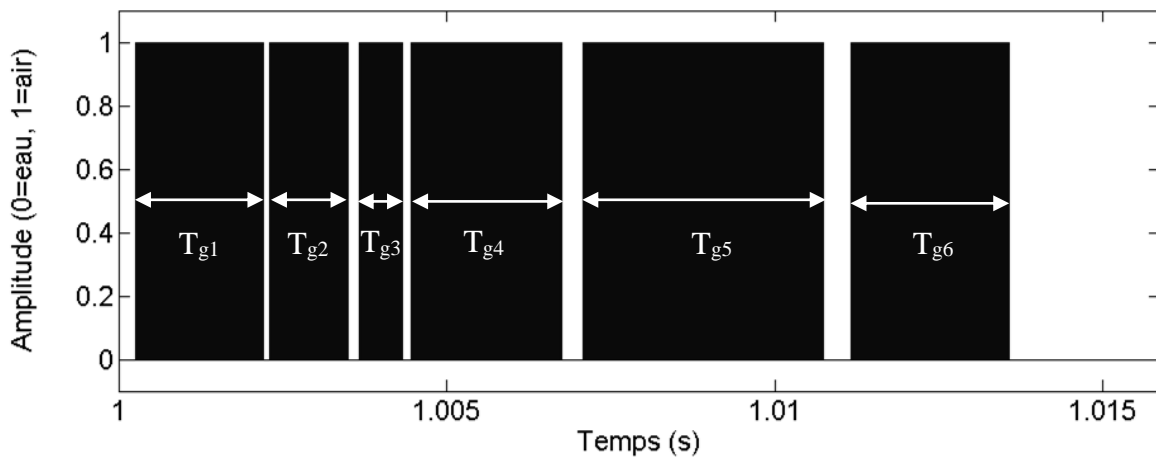


Figure 5.5 : Calcul du taux de vide moyen sur un intervalle Δt

On calcule ainsi la valeur du taux de vide moyen sur chaque intervalle de la durée d'acquisition, comme le traduit la Figure 5.6, où sont tracées pour l'ensemble du signal les valeurs successives du taux de vide moyen, sous forme de barres de largeur Δt :

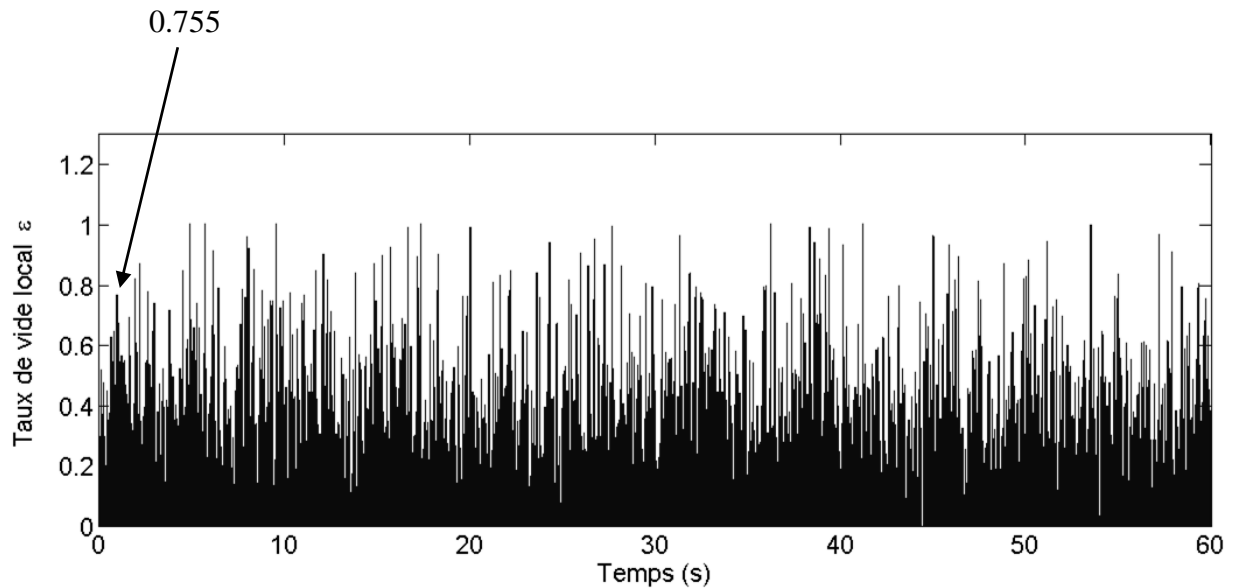


Figure 5.6 : Exemple d'évolution du taux de vide calculé sur chaque intervalle Δt , sur toute la durée d'acquisition

La valeur calculée précédemment, 0.755, est repérée par la flèche.

L'étape suivante consiste à diviser l'intervalle des valeurs du taux de vide, à savoir $[0;1]$, en intervalles de longueur $\Delta \varepsilon$. $\Delta \varepsilon$ correspondant à la précision voulue pour le tracé des PDF. Pour chaque intervalle $\Delta \varepsilon$, on calcule ensuite la densité de probabilité de retrouver dans l'ensemble des N valeurs calculées précédemment, des valeurs appartenant à ce $\Delta \varepsilon$:

$$PDF(i) = \frac{n_i/N}{\Delta \varepsilon_i} \quad (5.3)$$

où l'indice i repère un intervalle $\Delta \varepsilon_i$ donné, n_i est le nombre de valeurs de ε qui appartiennent à $\Delta \varepsilon_i$, et N le nombre total d'intervalles sur lesquels on calcule le taux de vide moyen ε : $N = \frac{T}{\Delta t}$

Après normalisation pour que la somme des probabilités vaille 1 (selon la théorie relative aux PDF), on obtient les valeurs de PDF et le graphe associé, dont un exemple est donné Figure 5.7.

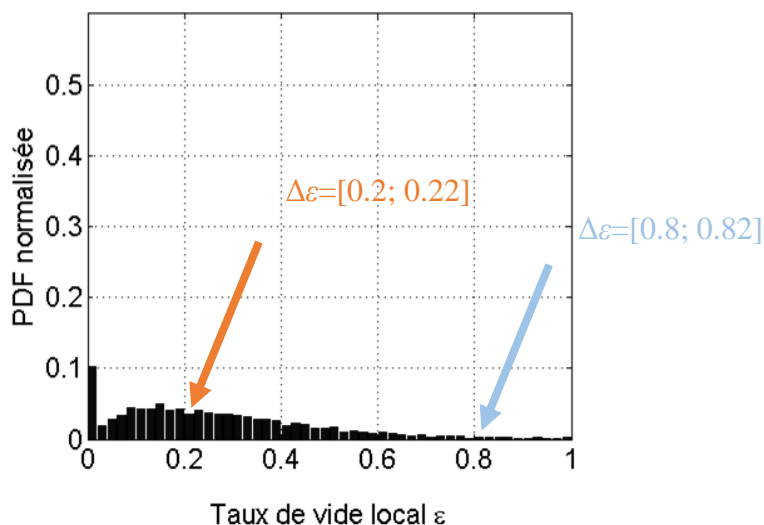


Figure 5.7 : Exemple de graphe de PDF

Pour mieux illustrer cette dernière étape : si on compare par exemple l'intervalle $\Delta\varepsilon=[0.2; 0.22]$ et l'intervalle $\Delta\varepsilon=[0.8; 0.82]$ repérés sur la Figure 5.8 par les lignes claires, on peut voir qu'il est assez probable de retrouver des valeurs de ε appartenant à $[0.2; 0.22]$ (car la ligne de ce $\Delta\varepsilon$ est située dans la partie foncée de la figure), contrairement à l'intervalle $[0.8; 0.82]$.

La valeur de PDF correspondante est donc plutôt élevée dans le premier cas mais très faible dans le deuxième, comme on le voit sur la Figure 5.7.

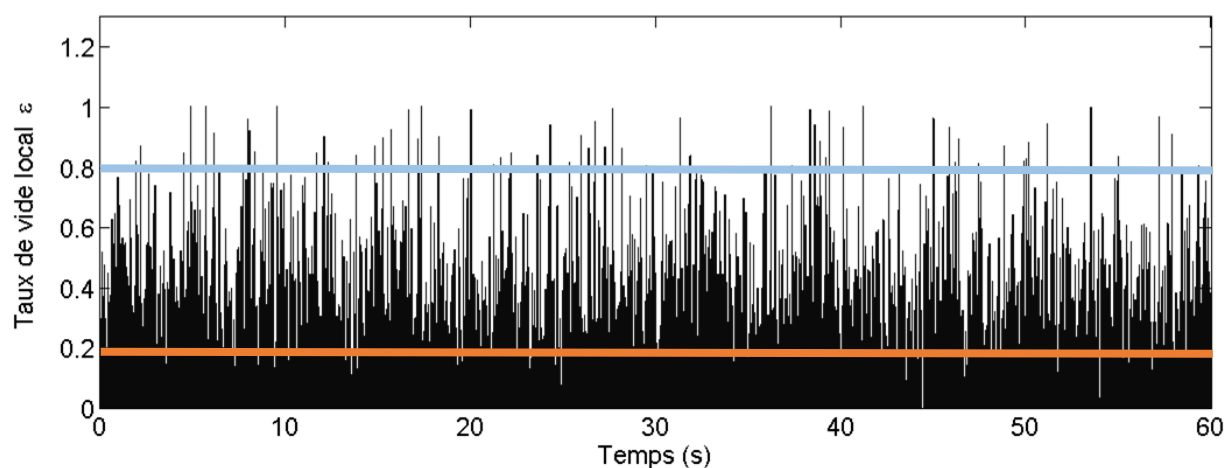
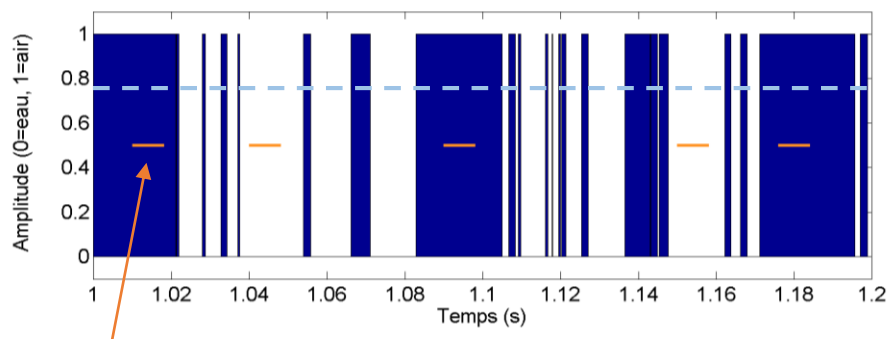


Figure 5.8 : Illustration du calcul de la PDF associée à un $\Delta\varepsilon$ donné

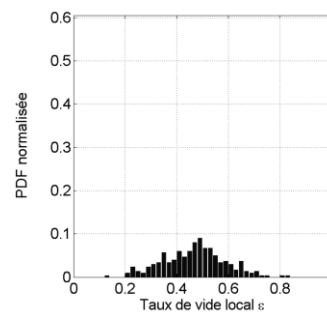
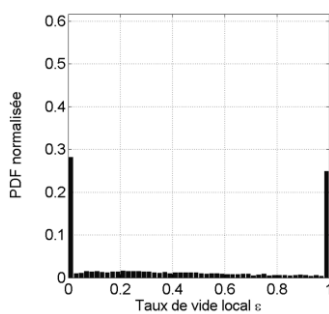
5.1.3 Choix de l'intervalle de temps Δt

On peut voir que dans le principe de calcul des PDF, l'intervalle de temps Δt est un paramètre clé, qui peut donner des résultats non représentatifs de la réalité s'il est mal choisi.

En effet, comme le montre la Figure 5.9, pour un point de mesure donné, un Δt trop petit par rapport aux structures de l'écoulement génèrera davantage de valeurs de ε purement à 0 ou 1, qui donneront l'illusion d'avoir des structures purement liquides ou gazeuses dont les tailles ne correspondront pas à l'écoulement réel. On observe des pics en 0 et 1 dans la PDF associée. Tandis qu'un Δt trop grand ne reflètera que le taux de vide moyen global (PDF centrée sur ε_{moy}) et ne permettra pas de différencier les configurations.



La ligne à l'intérieur des structures gazeuses et liquides représente une durée de 0.008s (exemple de Δt trop petit), et la durée totale affichée correspond à 0.2s (exemple de Δt trop grand – en pointillés)



- a) Δt trop petit : 0.008 s \Rightarrow pics en 0 et 1 b) Δt trop grand : 0.2 s \Rightarrow reflète seulement ε_{moy}

Figure 5.9 : Illustrations de mauvais choix de Δt pour un même point de mesure :

a) Δt trop petit et b) Δt trop grand

A ces différences d'allure de PDF, qu'on observe pour un même point de mesure (pour β et j_i donnés), s'ajoute un autre effet : une dépendance de Δt vis-à-vis de la vitesse du mélange homogène, j .

(On rappelle qu'en adoptant le modèle homogène pour décrire l'écoulement, on suppose que les deux phases se déplacent à la même vitesse, et on peut donc définir une vitesse du mélange homogène : j).

Suivant la valeur de la vitesse j , si on prend une seule et même valeur pour Δt , on va retrouver les mêmes problèmes que précédemment : les bulles de certains points à basse vitesse vont être « vues » comme des bouchons par un Δt qui sera trop petit par rapport à la vitesse de l'écoulement - Figure 5.9, cas a), et inversement, pour les points à haute vitesse de mélange, on n'observera que le taux de vide moyen - Figure 5.9, cas b).

Il faut donc tenir compte de la vitesse du mélange homogène j , et ainsi définir un Δt variable.

Choix final de Δt :

Finalement, la meilleure façon de choisir cet intervalle de temps semble être de tenir compte de la physique de l'écoulement diphasique. Le Δt doit permettre entre autres de distinguer une configuration à bulles d'une configuration intermittente, c'est-à-dire avec des structures plus grosses (qui correspondraient aux bouchons et bulles de Taylor observés en conduite interne), pour lesquelles il sera légitime d'observer des pics en 0 et en 1, qui correspondront bien à une réalité physique. On peut se reporter à la définition de ces configurations et aux images de l'écoulement correspondantes dans les Tableau 4.1 et Tableau 4.2.

Pour cela, au lieu de fixer directement une durée Δt , on choisit une longueur L dont dépendra l'intervalle de temps des PDF :

$$\Delta t = \frac{L}{j} \quad (5.4)$$

Après étude des longueurs caractéristiques de bulles observées sur les images de l'écoulement, pour des points de mesure répartis sur toutes les valeurs de β et j_l , il apparaît que la longueur caractéristique des bulles observées varie peu, entre 1.7 mm et 3 mm, avec une moyenne de 2.7 mm pour les points étudiés, qui sera donc prise comme référence pour la longueur d'une bulle : $L_c = 2.7$ mm.

Pour que la notion de taux de vide moyen ait un sens physique, il faut que la longueur L permette d'observer plusieurs bulles. Elle doit néanmoins être plus petite que la longueur supposée des structures plus grandes, pour pouvoir distinguer un écoulement intermittent d'un écoulement à bulles. On cherche donc un multiple de L_c qui permette cette distinction.

Si on visualise par exemple la longueur $5 L_c$ sur les images de l'écoulement pour différents couples (β, j_l) – $L = 5 L_c$ étant la hauteur du cadre affiché, il semble bien qu'elle permette une distinction entre une succession de bulles – Figure 5.11 a) et des structures plus grosses – Figure 5.11 b).

Cette valeur est confirmée par la comparaison de l'allure des PDF obtenues pour des Δt correspondant à différents multiples de L_c allant de 1 à 7, où il apparaît qu'une longueur valant 5 fois L_c est un bon compromis entre des pics trop élevés et des valeurs trop moyennées.

De plus, si on visualise des intervalles de la durée appropriée ainsi calculée pour plusieurs couples (β, j_l) , on constate que le problème de surestimation de $\varepsilon=0$ et $\varepsilon=1$ est évité et qu'on a toujours plusieurs bulles pour les configurations supposées à bulles, comme par exemple à la Figure 5.10.

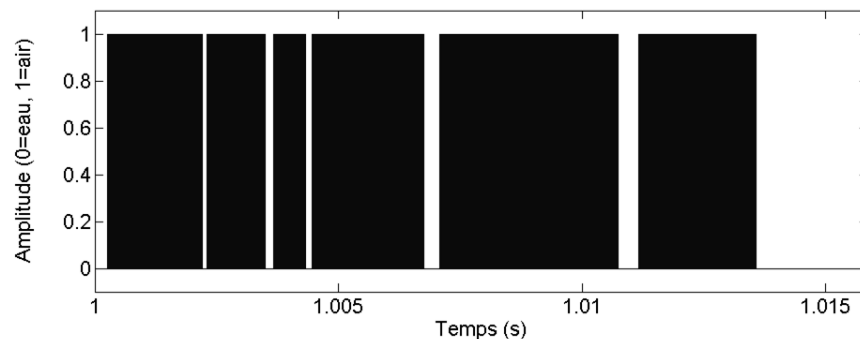


Figure 5.10 : Exemple de visualisation d'un Δt approprié

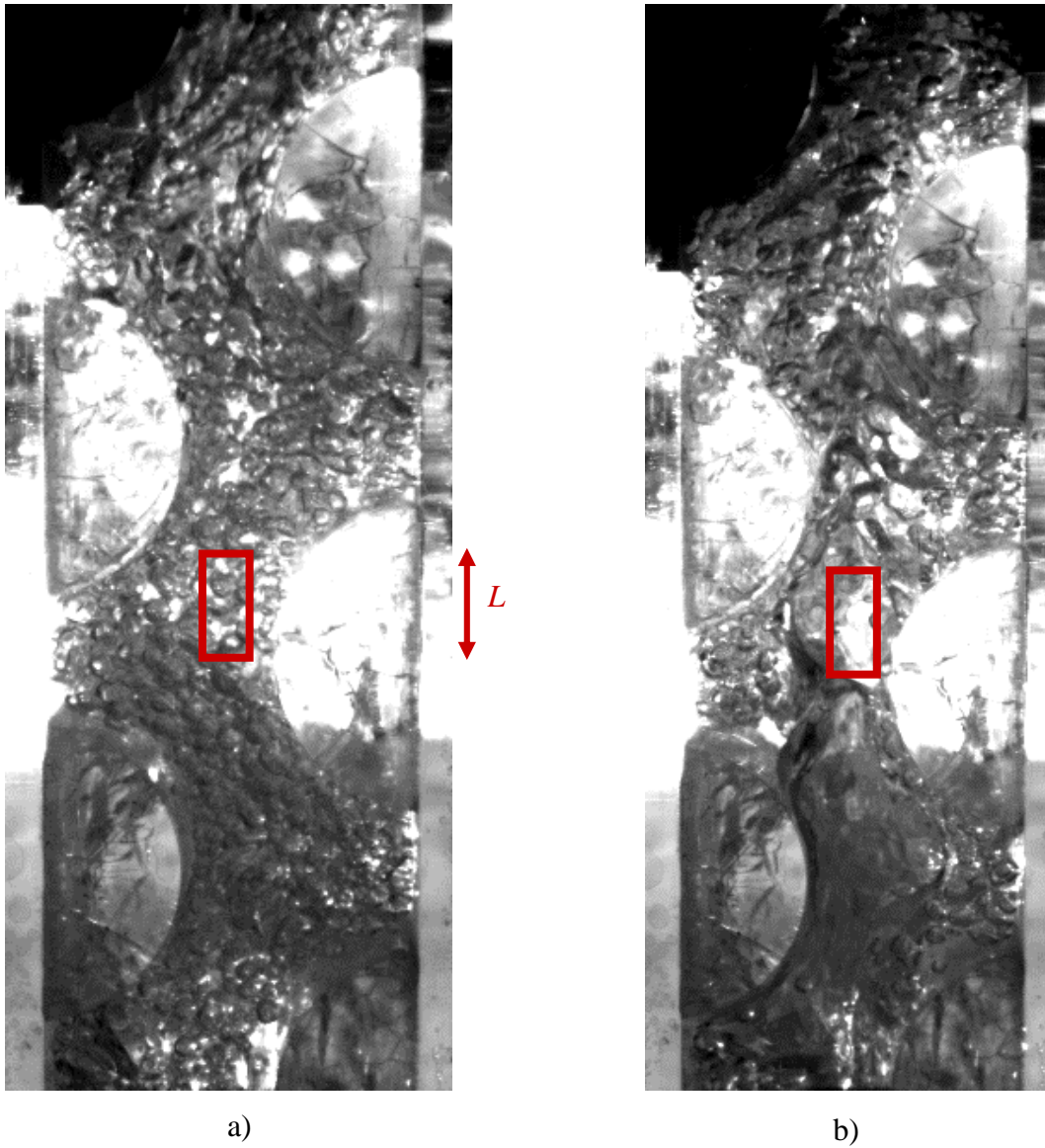


Figure 5.11 : Visualisation de $L=5 L_c$ sur les images de l'écoulement
 ($\beta=55\%$, $j_I=0.25$ m/s)

On choisit donc finalement :

$$\Delta t = \frac{5 L_c}{j} \quad (5.5)$$

5.1.4 Choix de l'intervalle $\Delta\varepsilon$

Choisir la valeur de $\Delta\varepsilon$ revient en fait à fixer la précision du tracé des PDF puisque le nombre total d'intervalles sur lesquels on calcule les PDF est :

$$N_\varepsilon = \frac{1}{\Delta\varepsilon} \quad (5.6)$$

Noghrehkar et al. (1999) ont par exemple choisi $\Delta\varepsilon=0.02$. La Figure 5.12 donne un exemple de calcul de PDF pour différentes valeurs de $\Delta\varepsilon$, et pour différents couples (β, j_i) .

Quel que soit le point de mesure, $\Delta\varepsilon=0.02$ est un bon compromis entre la précision et l'allure générale de la courbe, donc on retient cette valeur.

Tracés finaux des PDF pour tous les points de mesure

On retient donc finalement les paramètres suivants pour le tracé final des PDF du taux de vide :

$$\boxed{\Delta t = \frac{5 L_c}{j}} \text{ avec } L_c=2.7 \text{ mm et } j \text{ la vitesse homogène pour le point considéré, et } \boxed{\Delta\varepsilon = 0.02}.$$

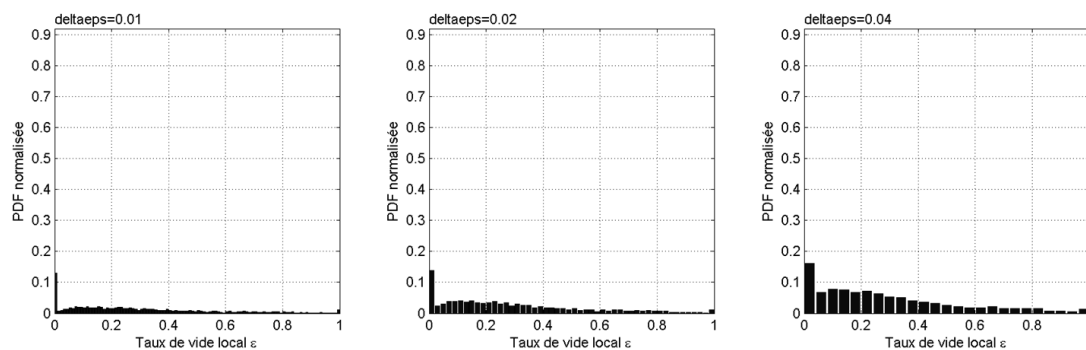
Cela donne par exemple les graphes de la Figure 5.13 lorsque β varie pour $j_i=0.3$ m/s. Les PDF du taux de vide pour tous les points de mesure se trouvent en Annexe B.

On observe que l'allure des PDF varie selon les valeurs de β et j_i , et on peut déjà identifier trois types d'allure illustrés par les trois lignes de la Figure 5.12, selon la présence et la position des pics.

Les PDF semblent donc déjà être de bons indicateurs des configurations d'écoulement, avec un nombre de grandeurs réduit (50 valeurs par graphe) par rapport aux évolutions temporelles du signal de la Figure 4.7.

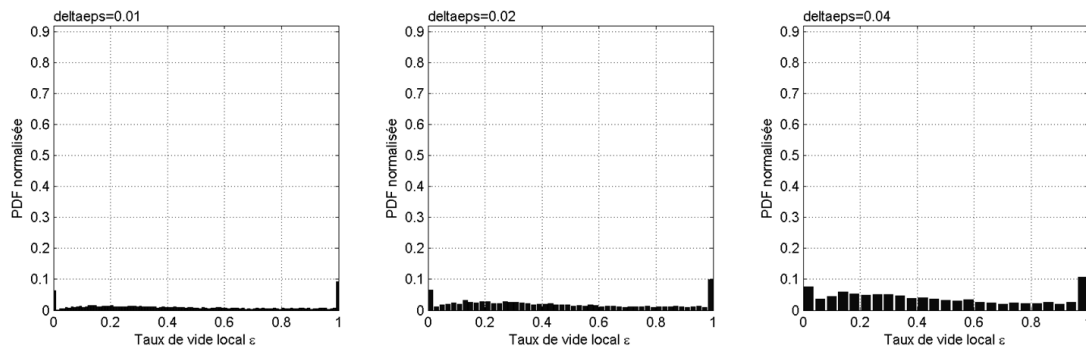
On constate là aussi sur la Figure 5.13 la même difficulté de classification manuelle due à une évolution continue de l'allure des PDF, qui fait qu'il serait assez difficile d'identifier d'emblée quelle serait la position exacte des frontières, d'où la nécessité d'utiliser la SVM.

PDF du taux de vide local sur la période [0,60]s pour $\beta=12\%$, $J_l=0.60$ m/s, $\Delta t=(L_c/J_l)*5.0$ ie $\Delta t=0.01984$ s



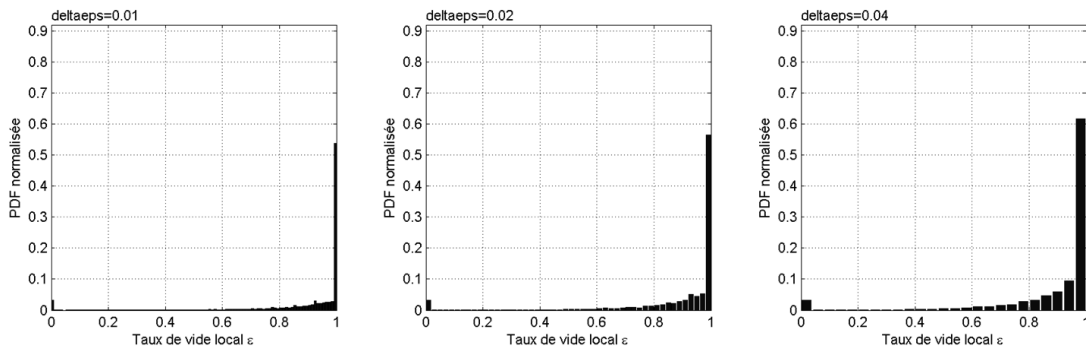
$$\beta=12\%, j_l=0.6 \text{ m/s}$$

PDF du taux de vide local sur la période [0,60]s pour $\beta=34\%$, $J_l=0.50$ m/s, $\Delta t=(L_c/J_l)*5.0$ ie $\Delta t=0.01779$ s



$$\beta=34\%, j_l=0.5 \text{ m/s}$$

PDF du taux de vide local sur la période [0,60]s pour $\beta=97\%$, $J_l=0.15$ m/s, $\Delta t=(L_c/J_l)*5.0$ ie $\Delta t=0.00227$ s



$$\beta=97\%, j_l=0.15 \text{ m/s}$$

a)

b)

c)

Figure 5.12 : Comparaison de différents choix pour $\Delta \epsilon$: a) 0.01, b) 0.02, c) 0.04

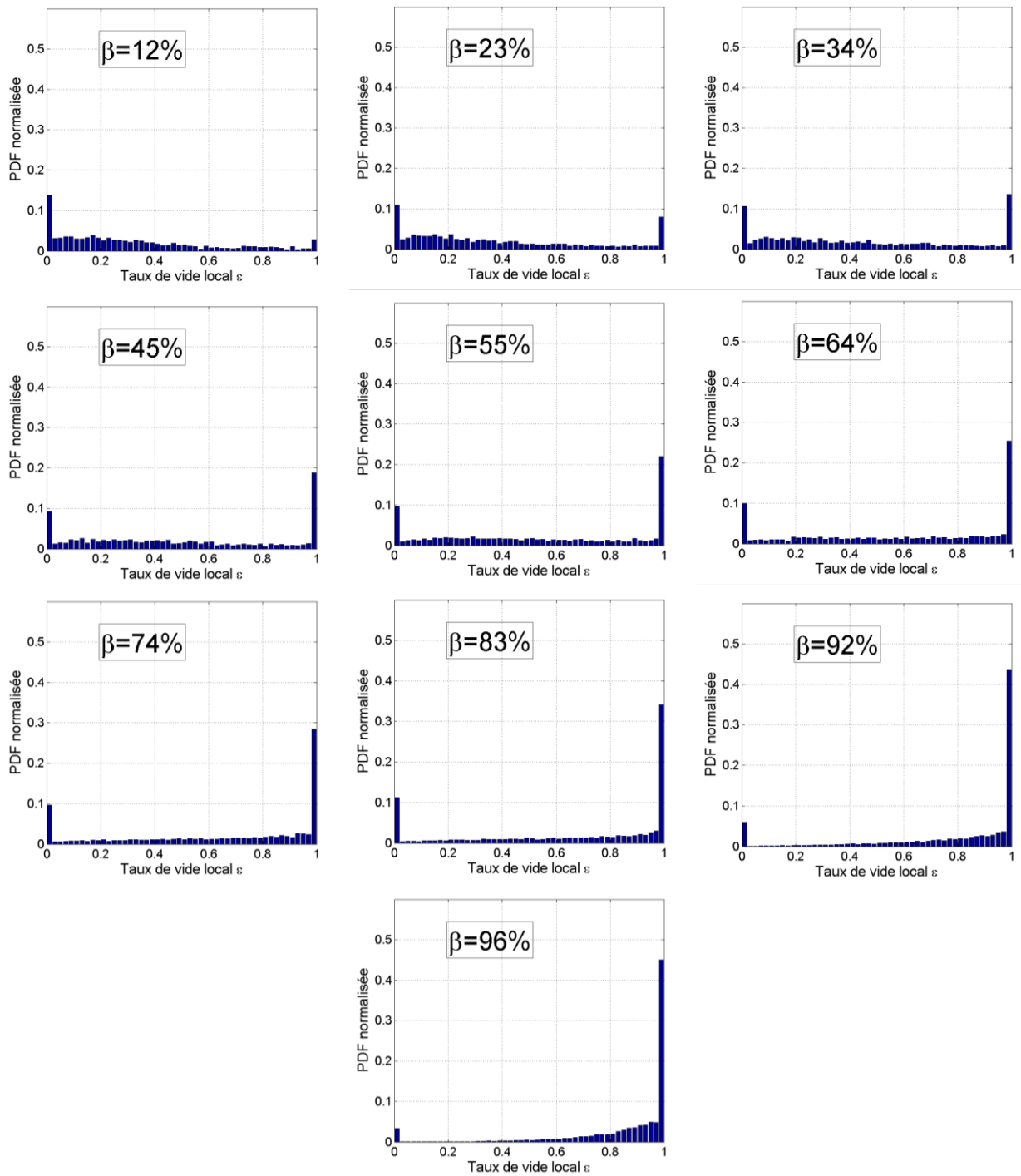


Figure 5.13 : Exemples de tracés finaux des PDF du taux de vide, pour $j_i=0.3$ m/s et différentes valeurs de β

5.2 Calcul des moments statistiques des signaux de la sonde

Un autre moyen d'exprimer l'information relative à la configuration d'écoulement contenue dans le signal de la sonde est de calculer les quatre premiers moments statistiques de ce signal.

Ces moments sont :

- La *moyenne* M :

$$M = \frac{1}{N} \sum_{i=1}^N x(i) \quad (5.7)$$

où x est un vecteur contenant N données scalaires

- L'*écart-type* Std (de l'anglais *Standard deviation*) :

$$Std = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |x(i) - M|^2} \quad (5.8)$$

où x est un vecteur contenant N données scalaires, M étant la moyenne de ces données.

- L'*asymétrie* Skw (de l'anglais *Skewness*) :

$$Skw = E \left[\left(\frac{x - M}{\sigma} \right)^3 \right] \quad (5.9)$$

où $E[t]$ représente l'espérance mathématique de la variable t , et M et σ sont respectivement la moyenne et l'écart-type des données contenues dans x .

- Le *kurtosis* K :

$$K = E \left[\left(\frac{x - M}{\sigma} \right)^4 \right] \quad (5.10)$$

où $E[t]$ représente l'espérance mathématique de la variable t , et M et σ sont respectivement la moyenne et l'écart-type des données contenues dans x .

Les 4 premiers moments statistiques sont des moments particuliers qui contiennent de l'information sur l'allure du tracé des données sur lesquelles ils sont calculés.

Étant donné leur définition et ce qu'ils expriment, la moyenne et l'écart-type sont dénommés « paramètres d'échelle », et l'asymétrie et le kurtosis sont dénommés « paramètres de forme » :

L'asymétrie, comme son nom l'indique, quantifie l'absence de symétrie d'une distribution par rapport à la moyenne. Lorsqu'elle est négative, les données sont réparties majoritairement à gauche de la moyenne – cf Figure 5.14. Inversement, lorsqu'elle est positive, les données sont réparties majoritairement à droite de la moyenne. Une distribution symétrique a ainsi une asymétrie nulle.

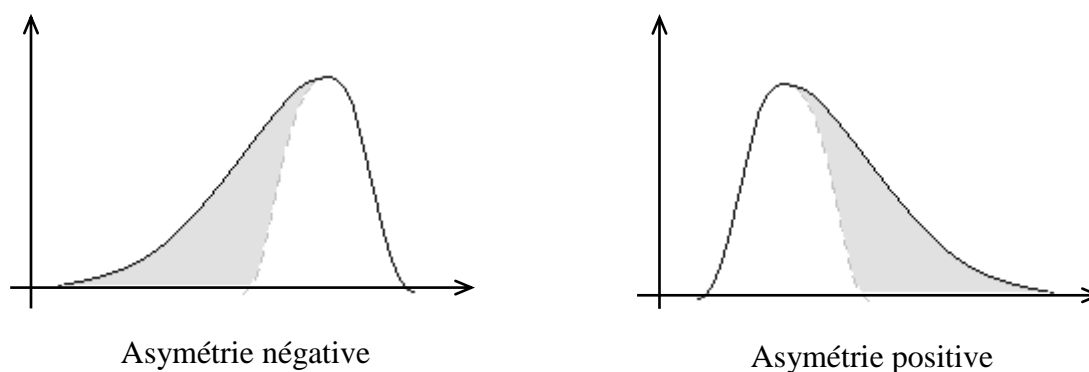


Figure 5.14 : Asymétrie d'une courbe

Le kurtosis exprime la présence de pics dans le signal. Le kurtosis de référence, de la distribution normale, est ici pris égal à 3. Une distribution présentant des pics plus resserrés que la distribution normale aura donc un kurtosis plus élevé, et inversement pour des pics plus larges.

Rosa et al. (2010), dans leur travail d'identification des configurations d'écoulement dans un tube vertical, ont étudié la possibilité d'utiliser ces 4 valeurs pour représenter l'information relative aux configurations.

Ils ont confirmé le fait que, de par leur définition, ces grandeurs statistiques reflètent l'allure des PDF :

- La moyenne reflète le taux de vide moyenné sur la durée d'acquisition (aux potentielles erreurs de centrage près, lors de l'acquisition du signal)
- L'écart-type représente la dispersion des valeurs autour de la moyenne
- L'asymétrie permet de quantifier le degré d'absence de symétrie de la courbe de PDF
- Le kurtosis est plus ou moins élevé selon la présence ou l'absence de pics dans la PDF

Afin d'illustrer ceci pour les données de cette étude, les Figure 5.15 à Figure 5.17 présentent pour trois allures typiques la PDF du taux de vide et les 4 moments statistiques associés (calculés ici seulement à partir du signal filtré, ré-échantillonné à 2000 Hz, pour permettre la comparaison avec les PDF) :

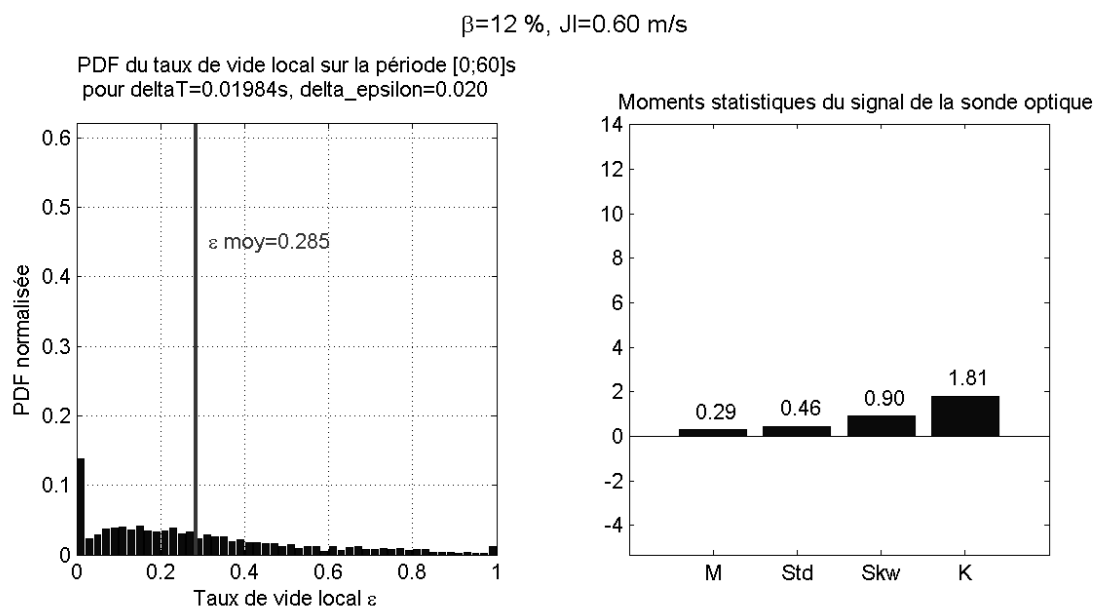


Figure 5.15 : Comparaison PDF/moments statistiques pour $\beta=12\%$, $j_l=0.6$ m/s

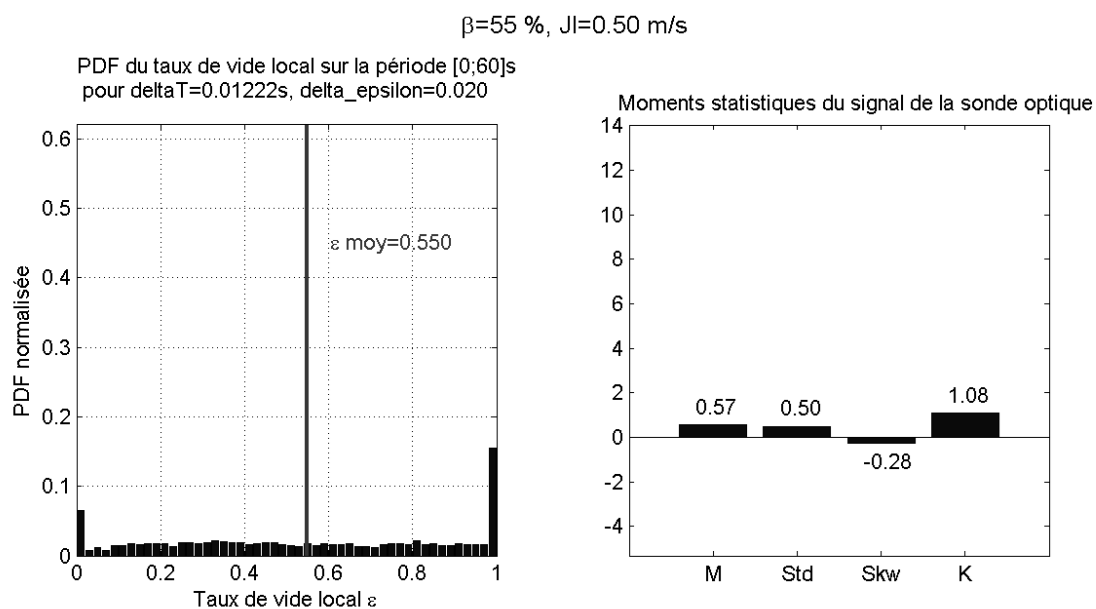


Figure 5.16 : Comparaison PDF/moments statistiques pour $\beta=55\%$, $j_I=0.5$ m/s

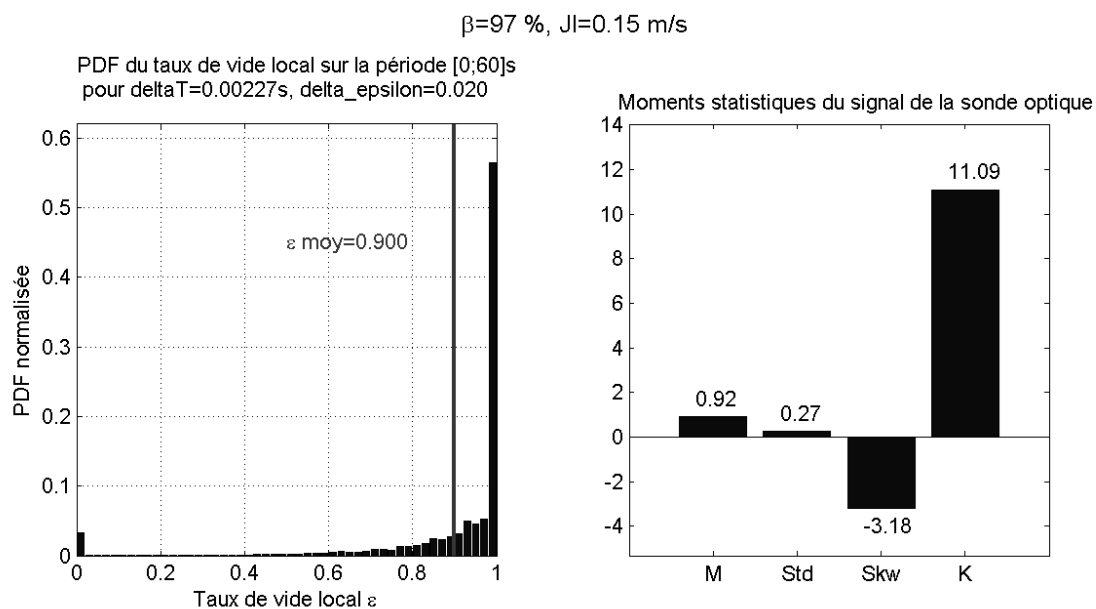


Figure 5.17 : Comparaison PDF/moments statistique pour $\beta=97\%$, $j_I=0.15$ m/s

Valeurs finales des moments statistiques

Le lien entre l'allure des PDF et les valeurs des moments statistiques étant connu, on calcule maintenant les valeurs de ces 4 moments directement à partir du signal brut (sur les 60s de signal échantillonné à 50 kHz). Les graphiques associés, pour tous les points de mesure, sont présentés en Annexe C, et illustrés Figure 5.18 pour les points tels que $j_i=0.3$ m/s, comme précédemment.

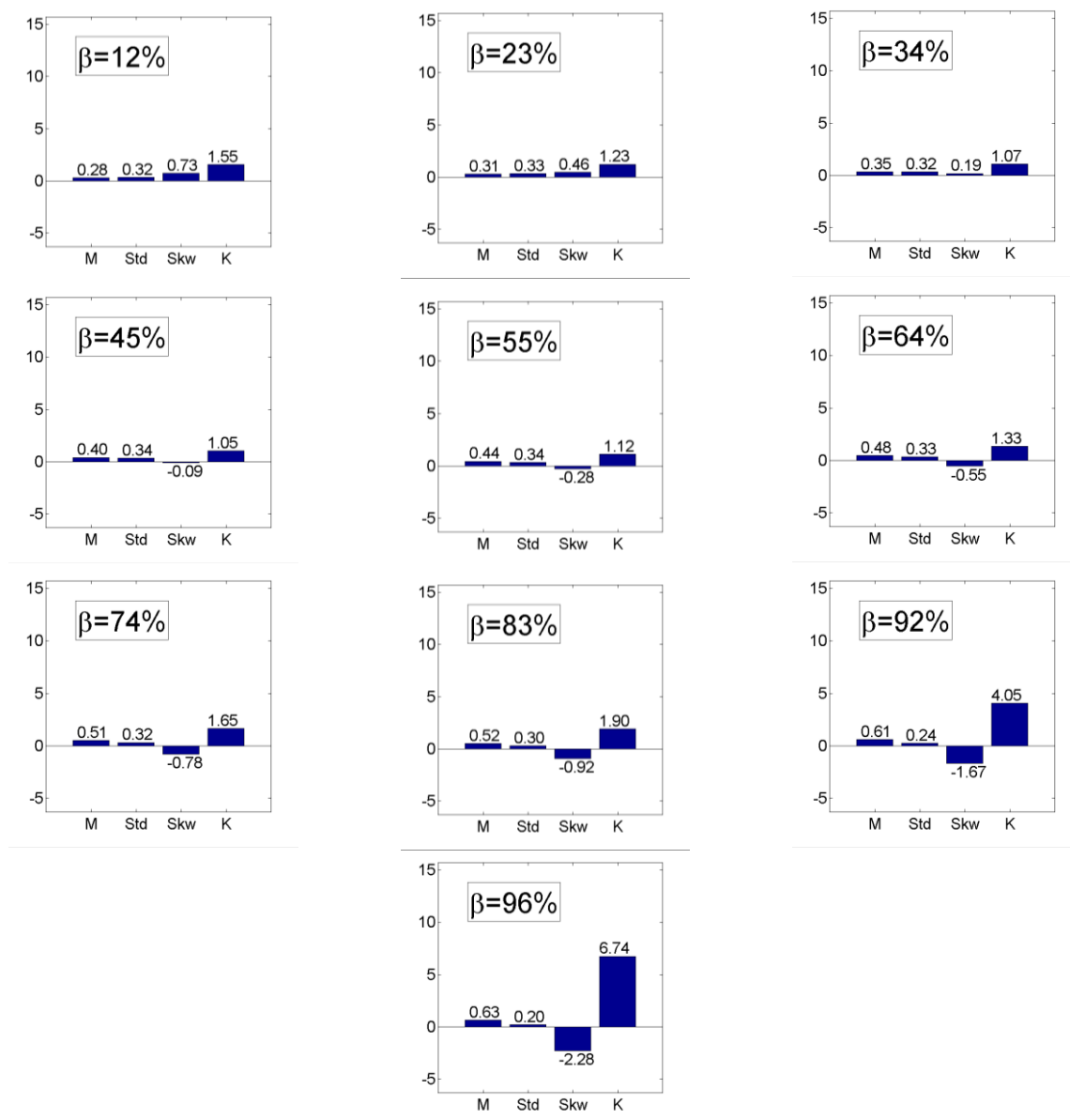


Figure 5.18 : Exemples de tracés finaux des 4 premiers moments statistiques du signal brut, pour $j_i=0.3$ m/s et différentes valeurs de β

5.3 Comparaison entre PDF et moments statistiques

Le calcul des moments statistiques présente plusieurs avantages non négligeables par rapport au tracé des PDF du taux de vide.

Le premier est qu'on n'a pas besoin de définir un intervalle de temps comme pour le calcul des PDF, ce qui permet de s'affranchir de la recherche d'un Δt adapté.

Le deuxième avantage est la réduction du nombre de données : pour $\Delta\varepsilon=0.02$ retenu précédemment, on obtient 50 valeurs de PDF, alors que les 4 valeurs des moments statistiques permettent à elles seules de synthétiser l'allure de ces courbes.

Enfin, un troisième avantage est que ces grandeurs statistiques peuvent être calculées directement à partir du signal brut donné par la sonde, ce qui permet de s'affranchir des éventuels problèmes de filtrage. En effet, la traduction du signal brut en signal échelon, ou le choix de seuils permettant de définir dans quelle phase on se trouve, est indispensable pour pouvoir calculer le temps passé dans chaque phase et donc les PDF. Cela induit une certaine marge d'erreur, car les bulles qui n'ont pas été complètement « enregistrées » par la sonde sont ignorées comme l'illustre la Figure 5.19.

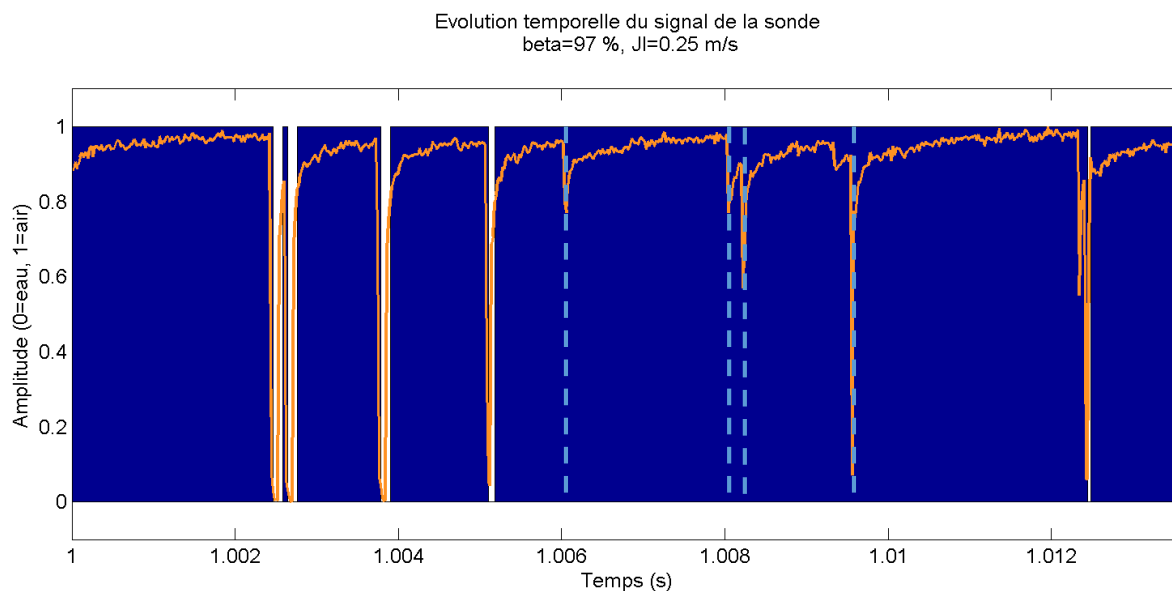


Figure 5.19 : Différences entre le signal brut (ligne claire) et le signal filtré (échelon)

Les pointillés - - - correspondent aux frontières des bulles ignorées par le filtrage

A moins de visualiser le tracé complet du signal pour tous les points, ce qui serait très fastidieux, il paraît difficile de fixer de façon parfaite les seuils de détection pour les PDF.

Les Figure 5.20 et Figure 5.21 montrent la superposition des valeurs des moments statistiques pour tous les points de mesure (regroupés arbitrairement en abscisses par numéro de point croissant de 1 à 110, les barres verticales marquant les changements de titre volumique), calculées d'une part sur le signal brut, et d'autre part sur le signal filtré.

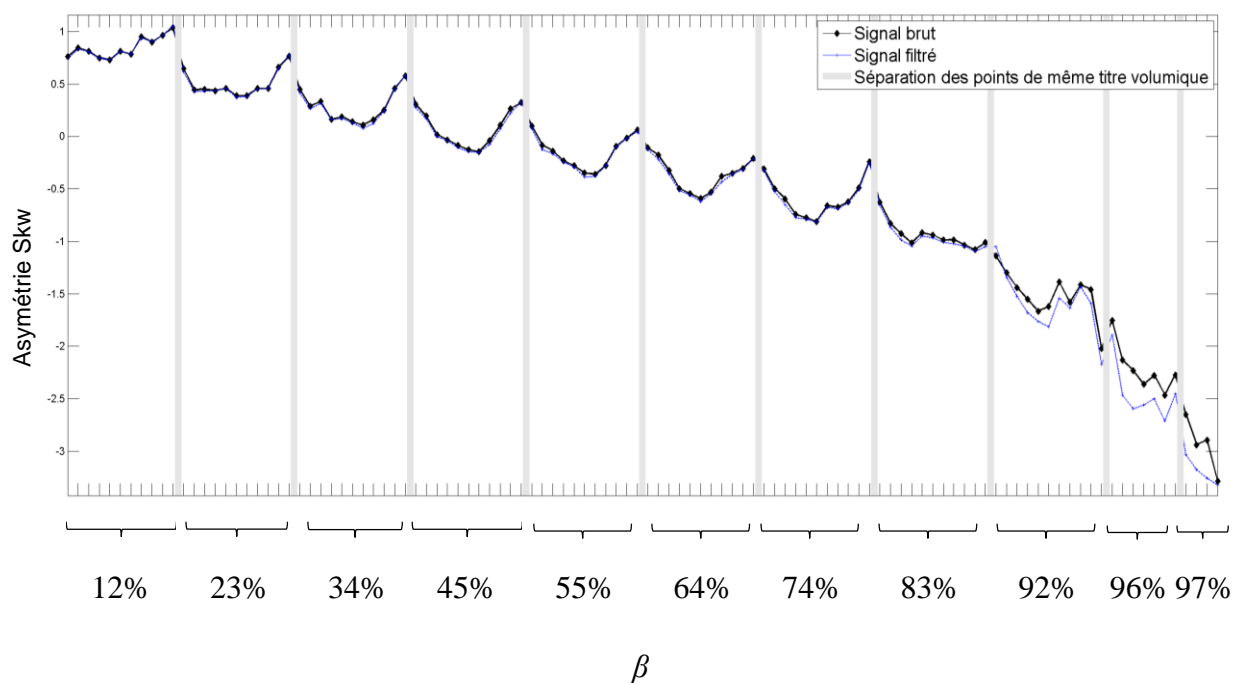


Figure 5.20 : Comparaison de l'asymétrie pour le signal brut (foncé) et le signal filtré (clair)

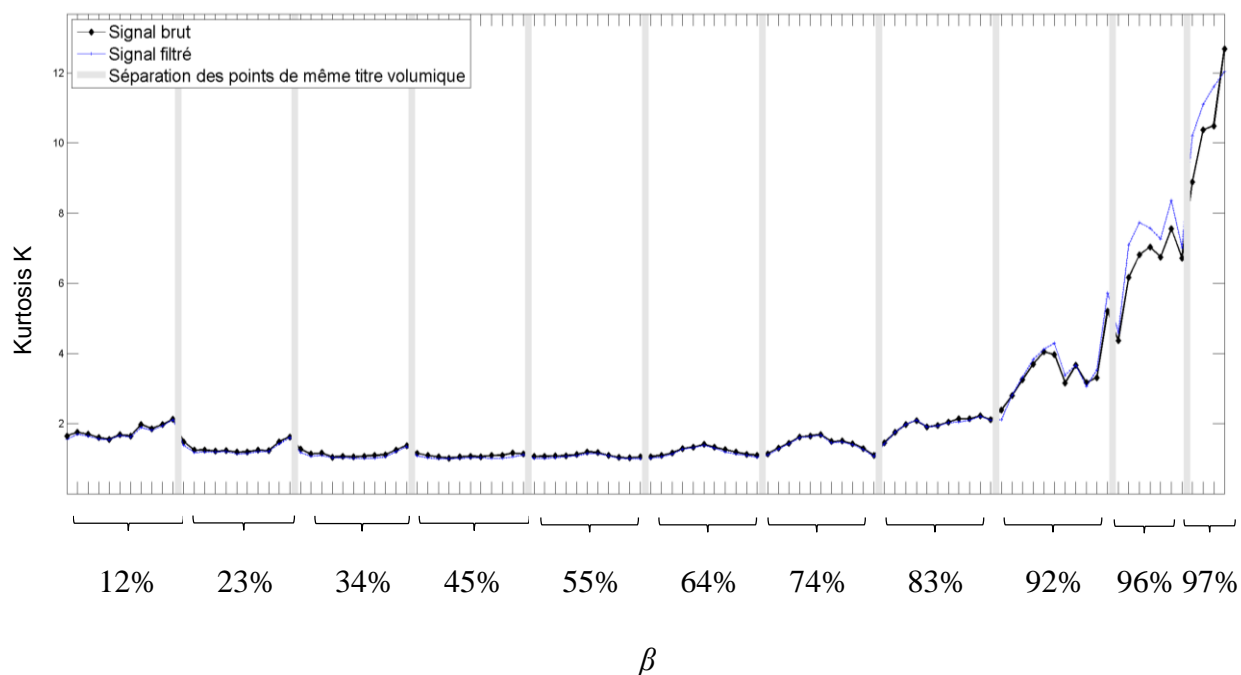


Figure 5.21 : Comparaison du kurtosis pour le signal brut (foncé) et le signal filtré (clair)

On observe bien une différence due au filtrage, qu'on retrouve principalement dans l'asymétrie et le kurtosis. Puisque ce dernier traduit l'amplitude des pics, on s'attendait bien à avoir des valeurs de K plus faibles pour le signal brut que pour le signal filtré (qui surestime le pic pour $\varepsilon=1$ puisqu'une succession de bulles est interprétée comme une structure gazeuse beaucoup plus grande).

On peut cependant constater que les deux signaux sont très bien superposés tant que le titre volumique ne dépasse pas 90%, et même après, l'écart n'est pas si grand. Cela suggère qu'une identification des configurations basée sur les moments statistiques du signal filtré, ou sur les PDF (qui commettent une erreur comparable) ne devrait pas donner des résultats fondamentalement différents de celle basée sur le signal brut.

Conclusion de la comparaison

Pour les raisons énoncées dans cette section, on préférera utiliser comme indicateurs des configurations de l'écoulement les 4 premiers moments statistiques du signal brut plutôt que les valeurs de PDF.

C'est de plus ce qui préconisé par Rosa et al. dans la section 5.3 (Rosa et al., 2010) :

- De par leur nature même, les 4 premiers moments statistiques synthétisent parfaitement l'allure des PDF, et peuvent donc fournir la même information à la machine d'apprentissage, mais de façon plus condensée et sans nécessiter de post-traitement du signal.
- En utilisant les moments statistiques comme données d'entrée pour leur intelligence artificielle (réseaux de neurones à couches multiples), une identification correcte à 100% de leurs points de mesure a été obtenue, ce qui valide leur utilisation.

CHAPITRE 6 UTILISATION D'UNE MACHINE À VECTEURS DE SUPPORT POUR OBTENIR UNE CARTE D'ÉCOULEMENT PLUS OBJECTIVE

Les résultats obtenus lors de l'analyse des signaux de sonde optique, en particulier les quatre premiers moments statistiques, suggèrent d'eux-mêmes des regroupements en plusieurs catégories. On pouvait en effet déjà différencier manuellement des types de profils et définir des critères de classification. Cependant, si certains profils sont faciles à classer, d'autres semblent se situer dans des zones de transition entre deux groupes, et le groupe qui leur serait attribué dépend trop de la façon de définir les critères de classification, qui serait ici subjective. C'est pourquoi on fait appel à la capacité de classification partiellement autonome d'un algorithme d'apprentissage supervisé : une machine à vecteurs de support.

6.1 Machine learning

Ce champ d'étude de l'intelligence artificielle consiste à construire une « machine » qui peut, de façon plus ou moins autonome, manipuler un ensemble de données pour en « apprendre » les structures sous-jacentes et les regrouper selon les similarités implicites observées. Un des domaines d'application du *machine learning* est la classification de données, dans des cas complexes où un humain n'est pas toujours capable d'effectuer des regroupements objectifs, ou encore lorsque les différences entre les classes ne sont pas directement observables.

6.1.1 Principe général de l'apprentissage supervisé

On parle d'apprentissage supervisé lorsqu'on a déjà identifié dans les données :

- un nombre de *classes*, c'est-à-dire de groupes possédant des caractéristiques différenciables de celles des autres groupes.
- des *données de référence* dont on connaît la classe (ou la probabilité d'appartenir à une classe donnée, on parle alors d'apprentissage supervisé probabiliste).

En indiquant à l'algorithme de classification ce nombre de classes d'une part, et quelles sont les données de référence et leurs classes d'autre part, la machine est alors capable d'apprendre à classer les données de façon à respecter la structure sous-jacente qui fait correspondre une classe donnée à chaque donnée de référence. Celles-ci constituent les *données d'entraînement*, à partir desquelles la machine construit un *modèle de classement*. Cet *entraînement* de la machine est la première phase à réaliser.

La deuxième phase est la phase de *classification* à proprement parler : grâce au modèle de classement qu'elle a créé, la machine est capable d'attribuer une classe à de nouvelles données en se basant uniquement sur leurs caractéristiques connues, en les comparant aux données d'entraînement. On parle aussi de prédiction de classes, et ces nouvelles données constituent les *données de test*.

6.1.2 Machine à Vecteurs de Support (SVM)

Une telle machine est une application de la théorie d'apprentissage statistique supervisé développée entre autres par Vapnik à partir des années 1960 (Cortes & Vapnik, 1995 ; Vapnik, 2013). A chaque point de mesure est associé un vecteur, dont les composantes contiennent des valeurs numériques liées implicitement aux classes, sur lesquelles on se basera pour classer les points.

L'apprentissage est basé sur la recherche d'un hyperplan permettant de séparer linéairement les données d'entraînement de façon optimale dans un espace image de grande dimension, pour une frontière donnée. La dénomination SVM vient du fait que lors de la recherche de cet hyperplan séparateur, on n'utilise pas tous les points de référence, mais on recherche parmi eux les points qui sont les plus proches de la frontière et qui contribuent donc le plus à la définition de l'hyperplan. C'est pourquoi on les appelle des vecteurs de support.

Les SVM sont aussi appelées « classificateurs à marge maximale » (*Large margin classifiers*), car on définit l'hyperplan optimal comme étant celui qui assure la marge maximale entre l'hyperplan et les données de chaque classe les plus proches. Cela permet de séparer les données au mieux. La Figure 6.1 représente des données d'entraînement de deux classes dans l'espace

image, et montre deux exemples d'hyperplans qui ne sont pas optimaux, contrairement à celui de la Figure 6.2, qui se situe au milieu des hyperplans passant par les données les plus proches (ces données seront des vecteurs de support), et qui définit ainsi la marge maximale. Intuitivement, c'est l'hyperplan qu'on aurait envie de tracer pour séparer les données.

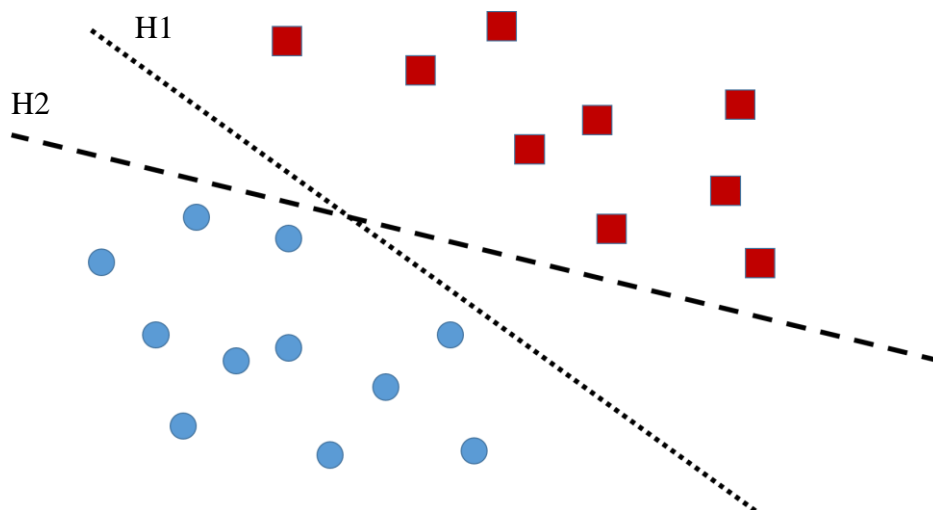


Figure 6.1 : Deux exemples d'hyperplans séparateurs, pas optimaux

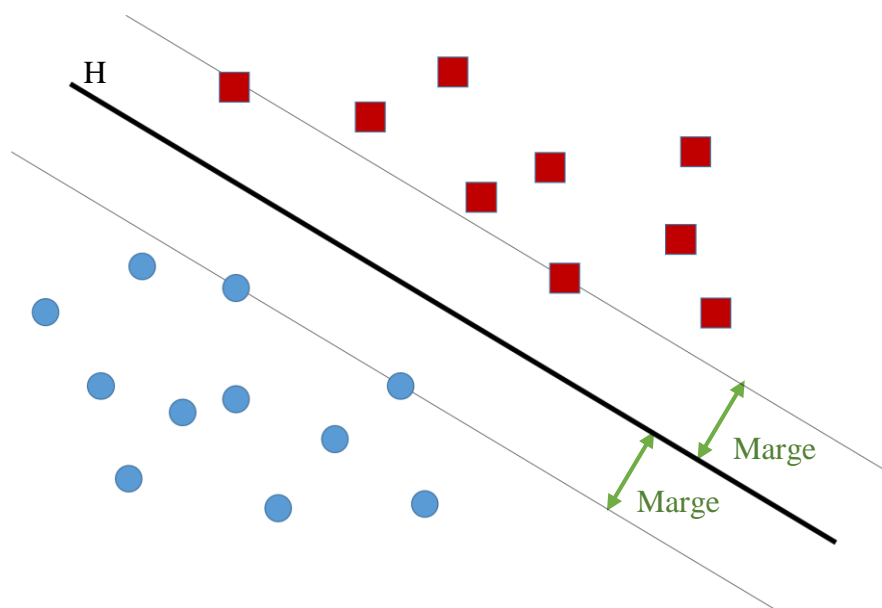


Figure 6.2 : Hyperplan optimal H définissant une marge maximale

Mathématiquement, on cherche à maximiser la marge, il s'agit donc de résoudre un problème d'optimisation, à l'issue duquel on obtient un modèle permettant de prédire la classe de nouveaux points pour lesquels on dispose seulement des valeurs des composantes des vecteurs associés.

Ci-dessous sont présentés les éléments théoriques principaux sur lesquels est basée la SVM, pour lesquels on s'est appuyé sur une étude théorique détaillée réalisée au sein de la Chaire (Komgom Nguerpjop, 2006).

On note x_i les vecteurs associés aux i points de mesure (avec $i = 1, 2, \dots, n$ et $x_i \in \mathbb{R}^n$) et y_i les classes associées (avec $y_i = -1$ ou $+1$ selon la classe). L'hyperplan séparant les deux classes est défini par l'équation :

$$\omega \cdot x + b = 0 \quad (6.1)$$

où ω est le vecteur normal à l'hyperplan, b la distance de l'hyperplan à l'espace d'origine, et x un vecteur de valeurs réelles, comme illustré à la Figure 6.3.

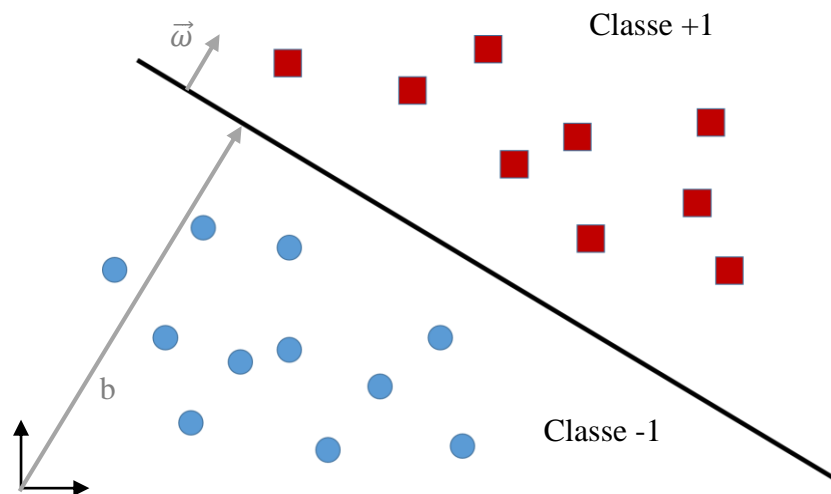


Figure 6.3 : Définition de l'hyperplan

Pour obtenir l'hyperplan optimal, on veut maximiser la marge, ce qui revient à minimiser la quantité $\frac{1}{2} \|\omega\|^2$. Lorsque les données sont linéairement séparables dans l'espace d'origine, la contrainte à respecter est :

$$y_i (\langle \omega, x_i \rangle + b) \geq 1, i = 1, 2, \dots, n \quad (6.2)$$

où $\langle ., . \rangle$ est le produit scalaire dans l'espace de départ.

Les paramètres ω et b de l'hyperplan optimal sont donc les solutions d'un problème d'optimisation dont la formulation lagrangienne primale est :

$$L_P = \frac{1}{2} \langle \omega, \omega \rangle - \sum_{i=1}^n \alpha_i [y_i (\langle \omega, x_i \rangle + b) - 1] \quad (6.3)$$

où $\alpha_i \geq 0$ sont les coefficients de Lagrange.

En différenciant cette équation par rapport à ω et b qu'on fixe à zéro, en utilisant une condition de stationnarité puis en substituant les résultats dans la formulation primale du Lagrangien, on obtient la formulation duale suivante :

$$L_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \quad (6.4)$$

Les coefficients de Lagrange mesurent en fait la sensibilité de la solution par rapport à chaque point x_i . En respectant les conditions du théorème de Karush Kuhn Tuckner, qui permet de garantir que le problème d'optimisation n'a qu'une unique solution, seuls les points x_i tels que $\alpha_i \neq 0$ ont un impact sur la solution et permettent de définir l'hyperplan optimal : ce sont les vecteurs de support. On note α_i^* ces coefficients de Lagrange non nuls, qui correspondent donc aux valeurs optimales et aux vecteurs de support, et on a alors l'expression du vecteur normal à l'hyperplan optimal :

$$\omega^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (6.5)$$

Et la fonction de décision f qui permet de classer les vecteurs de part et d'autre de l'hyperplan, en fonction du signe associé, est alors :

$$f(x) = \text{sgn}(\langle \omega^*, x \rangle + b^*) = \text{sgn} \left(\sum_{i=1}^n \alpha_i^* y_i \langle x_i, x \rangle + b^* \right) \quad (6.6)$$

où sgn dénote la fonction signe (qui associe +1 à une quantité positive, -1 à une quantité négative, 0 à une quantité nulle), et b^* correspond à la valeur optimale de la distance entre l'hyperplan et l'origine de l'espace.

La force des SVM réside notamment dans leur capacité à classer des données qui ne sont pas linéairement séparables au départ. On utilise pour cela une fonction noyau K (*Kernel*), qui correspond à un produit scalaire ϕ dans un espace image de plus grande dimension, où les données sont linéairement séparables. L'existence d'une telle fonction est assurée lorsque les conditions du théorème de Mercer sont respectées.

Pour deux vecteurs x_1, x_2 appartenant à l'espace de départ X , K est la fonction noyau telle que :

$$K(x_1, x_2) = \langle \phi(x_1), \phi(x_2) \rangle$$

avec $\phi : X \rightarrow F$

$$x \mapsto \phi(x) \quad (6.7)$$

avec F l'espace image dans lequel les données seront linéairement séparables.

On n'a pas besoin de connaître le produit scalaire ϕ , il suffit d'effectuer la transformation par la fonction noyau, ce qui n'augmente pas la complexité des calculs.

Pour des données non linéairement séparables dans l'espace de départ, la fonction de décision est donc finalement donnée par :

$$f(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i^* y_i K(x_i, x) + b^* \right) \quad (6.8)$$

Ainsi, en résolvant le problème d'optimisation (Eq. (6.4)) pour les n données d'entraînement, on obtient les paramètres correspondants à l'hyperplan optimal : α_i^* et b^* , qui permettent ensuite de calculer les valeurs de la fonction de décision f pour les autres points : +1 ou -1, ce qui revient à les classer de part et d'autre de l'hyperplan, et donc à prédire leur classe en se ramenant à l'espace de départ.

Multi-classification

Une SVM est donc au départ un classificateur binaire, mais il existe plusieurs méthodes permettant de prendre en charge un plus grand nombre de classes, en se ramenant à des classificateurs binaires intermédiaires. On peut par exemple traiter successivement toutes les paires de classes appropriées et regrouper les résultats (*Pair-wise multiclassification*), ou encore considérer successivement chaque classe et définir la frontière entre cette classe et le regroupement de toutes les autres, puis recouper les résultats pour obtenir une classification globale (*One Versus Rest multiclassification*).

6.2 Application à l'identification des configurations d'écoulement

Pour l'implémentation de la SVM, un programme en libre accès, largement utilisé, a été employé : LIBSVM (Chang & Lin, 2011). Les références détaillées des ressources liées aux SVM sont données en Annexe D.

Même si des bibliothèques comme LIBSVM dispensent du codage détaillé de la SVM, il reste à bien utiliser les fonctions fournies et notamment à choisir des paramètres adaptés au cas que l'on traite, sans quoi la pertinence des classifications obtenues peut être remise en cause. Les différents paramètres et méthodes employés pour définir une SVM adaptée sont détaillés dans les paragraphes suivants, et regroupés dans le Tableau 6.2 récapitulatif.

6.2.1 Choix des paramètres généraux de la SVM

Dans le cas de cette étude, les données sont au nombre de 110; pour chacun de ces points, on dispose de 4 grandeurs caractéristiques (*features*) : les 4 premiers moments statistiques calculés directement à partir du signal brut de la sonde optique, qui synthétisent l'information liée à la configuration de l'écoulement comme montré au Chapitre 4. Ces grandeurs sont les valeurs de la moyenne, l'écart-type, l'asymétrie et le kurtosis du signal. Les vecteurs représentant les 110 points de mesure sont donc initialement de dimension 4 :

$$x_i = \begin{Bmatrix} M(i) \\ Std(i) \\ Skw(i) \\ K(i) \end{Bmatrix}, i = 1, 2, \dots, 110 \quad (6.9)$$

Le type de SVM utilisé est « C-SVC » c'est-à-dire une classification à vecteurs de support qui utilise le paramètre C , qui quantifie l'importance accordée aux erreurs de classification, et permet d'ajuster la complexité de la fonction de décision et donc des frontières obtenues.

Étant donné l'ordre de grandeur du set de données et la dimension des vecteurs associés, on choisit comme fonction noyau un noyau gaussien (aussi appelé *Radial Basis Function*). C'est également le noyau par défaut employé par LIBSVM, car c'est un bon compromis vis-à-vis du degré de complexité de la SVM ainsi définie. Ce noyau peut en effet gérer le cas d'une relation non-linéaire entre les caractéristiques des vecteurs et les classes observées, contrairement à un noyau linéaire. Pour deux vecteurs x_1 et x_2 , cette fonction noyau est définie par :

$$K(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2} \quad (6.10)$$

Le paramètre γ sera choisi par la suite.

Dans cette étude, une pré-classification manuelle des données a permis d'identifier trois classes différentes – cf section suivante 6.2.2. Il faudra donc employer une procédure de multi-classification, on choisit ici l'approche *One Versus Rest* (OVR) qui consiste à réaliser l'entraînement et la classification des données d'une classe à la fois face à l'ensemble des autres classes réunies, qui simule une deuxième classe.

Les paramètres C et γ seront obtenus lors de la procédure de *cross validation*, qui sera détaillée dans la section 6.2.3.

De plus, avant d'utiliser une fonction noyau, en particulier le noyau gaussien, il est impératif d'effectuer une mise à l'échelle des valeurs contenues dans les vecteurs (*Scaling*), sans quoi les différences d'ordre de grandeur qui peuvent exister d'une composante à l'autre peuvent biaiser la classification. Par exemple, si on néglige cette étape avec les données utilisées, la 4^{ème} composante des vecteurs, le kurtosis du signal, va majoritairement influencer la classification car l'amplitude de l'intervalle dans lequel évoluent ses valeurs est bien plus grand que celles des autres grandeurs caractéristiques – cf Tableau 6.1. Les valeurs prises par la fonction noyau dépendraient donc essentiellement du kurtosis, de façon artificielle.

Tableau 6.1 : Valeurs initiales des moments (minimales et maximales)

| | M | Std | Skw | K |
|-----|------|-------|-------|-------|
| Min | 0.25 | 0.16 | -3.29 | 1.02 |
| Max | 0.85 | 0.41 | 1.04 | 12.68 |

On choisit donc de ramener toutes les valeurs dans l'intervalle $[0;1]$ avant de faire quoi que ce soit d'autre (en retirant simplement la valeur minimale sur l'ensemble des 110 points, puis en divisant les valeurs obtenues par leur valeur maximale, et ce successivement pour chaque composante des vecteurs : M , Std , Skw et K). Il est important d'appliquer le même procédé à toutes les valeurs d'une composante donnée, pour tous les sets (training, test) s'ils sont déjà définis.

Enfin, le choix des données d'entraînement et de test, par la suite appelées respectivement *training set* et *test set*, sera fait de façon aléatoire au sein de chaque classe. Parmi les données dont on connaît la classe, l'idée est de respecter une proportion commune de la littérature de 70% entre le training set et le test set (qui permet d'estimer la capacité de généralisation de la SVM). Comme les points attribués aux classes 1 et 3 sont peu nombreux par rapport à ceux de la classe 2, on choisit aléatoirement *au sein de chaque classe* 70% des points, qui seront affectés au training set, et les 30% restants (environ) seront affectés au test set. Cela permet de s'assurer que les trois classes seront bien représentées au sein de chaque set comme l'illustre la Figure 6.4.

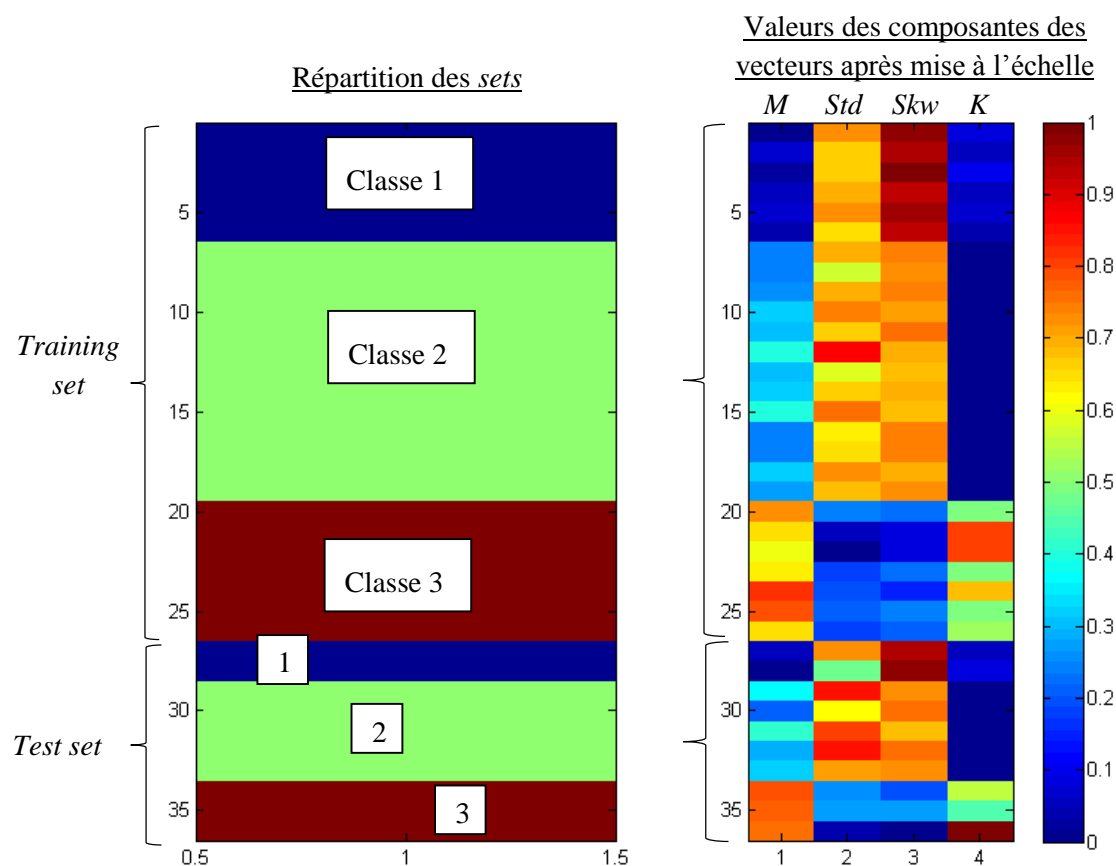


Figure 6.4 : Exemple de répartition en *training* et *test sets*,
et valeurs des composantes (*features*) correspondantes

(NB : les indices en ordonnée ne représentent pas le numéro des points de mesure).

Tableau 6.2 : Récapitulatif des paramètres de la SVM

| | |
|--|---|
| Nombre total de données considérées | 110, points de mesure présentés au Chapitre 3 |
| Composantes des vecteurs (<i>features</i>) | 4 moments statistiques : M , Std , Skw , K (<i>sauf mention du contraire</i>) |
| Type de SVM | C-SVC : <i>Support Vector Classification</i> , utilisation du paramètre C : paramètre de régularisation / importance accordée aux erreurs de classification |
| Fonction noyau | Noyau Gaussien (<i>Radial Basis Function</i>), paramètre γ $K(x_1, x_2) = e^{-\gamma \ x_1 - x_2\ ^2}$ |
| Mode de classification | OVR Multi-classification, identification des données d'une classe par rapport à toutes les autres (<i>One Versus Rest</i>) |
| Choix des paramètres C et γ | Lors de la procédure de <i>3-fold cross validation</i> |
| Mise à l'échelle des données | On ramène toutes les valeurs de chaque composante des vecteurs dans l'intervalle $[0 ; 1]$ |
| Ratio <i>training/test sets</i> parmi les données de classe connue | 70% / 30% |

6.2.2 Pré-classification des données

Comme la SVM est une forme d'apprentissage supervisé, il est nécessaire de définir un certain nombre de points de référence, parmi lesquels seront choisies les données d'entraînement grâce auxquelles on pourra construire le modèle de classification. Pour ces points, la configuration ne peut être définie que manuellement, donc avec une part de subjectivité.

Mais en prenant des points de référence loin des frontières supposées, dont la configuration est clairement définie, le fait d'utiliser ensuite la SVM pour classer les points des zones de transition paraît être une solution déjà plus objective qu'une classification manuelle de l'ensemble des points. De façon générale, pour que la classification proposée ait un sens physique, il est inévitable à un moment ou à un autre de relier les mesures quantitatives (quelles qu'elles soient) à l'observation de la structure réelle de l'écoulement, qui reste une donnée qualitative comportant donc une part de subjectivité.

Afin de garantir toutefois la plus grande objectivité possible, les points auxquels une classe a été attribuée, qui seront considérées comme des points de référence par la suite, respectent les conditions énumérées dans le Tableau 6.3.

Tableau 6.3 : Conditions respectées par les points de référence

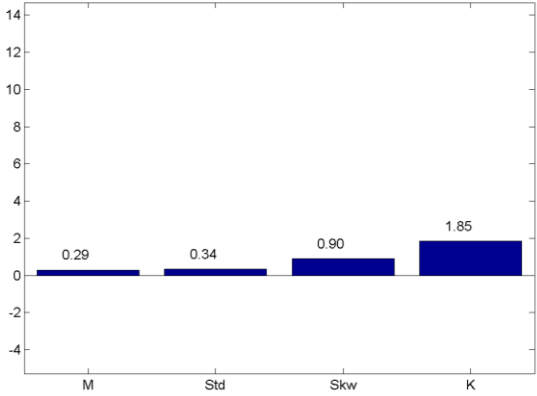
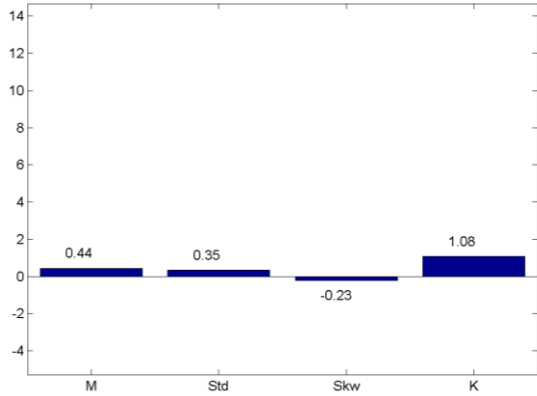
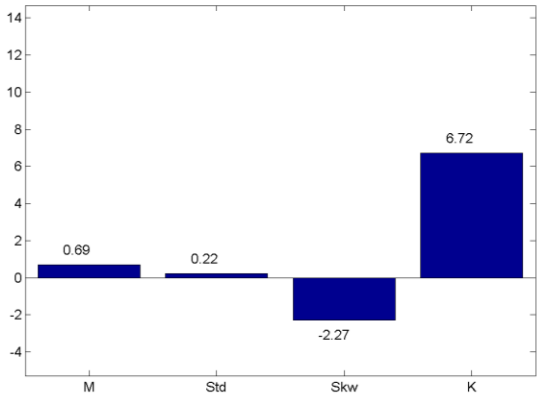
| | |
|--|--|
| <p><u>Condition 1</u> : parfait accord entre les configurations qui seraient attribuées en se basant sur les différentes observations ci-contre, indépendamment les unes des autres. L'ordre dans lequel elles sont citées est décroissant en terme de fiabilité estimée.</p> | <p>A) Profil des 4 moments statistiques du signal de la sonde : regroupement manuel de certains points dont l'allure des moments se distingue facilement des autres. On s'est notamment appuyé sur les allures typiques rapportées par Rosa et al. (2010) qui présentent des similitudes importantes avec les valeurs obtenues ici.</p> |
| | <p>B) Structure instantanée observée sur les photos de l'écoulement prises par caméra haute vitesse : taille, forme et position relative des structures des deux phases. On s'appuie aussi sur les observations disponibles dans la littérature, qui sont comparées avec les présentes données.</p> |
| | <p>C) Structure dynamique observée sur les vidéos de l'écoulement créées à partir des images : sens de l'écoulement, présence ou non d'oscillations locales ou globales.</p> |
| | <p>D) Observations visuelles en direct lors des mesures, notamment sur le sens de l'écoulement et le degré d'agitation.</p> |
| <p><u>Condition 2</u> : les points qui respectent la condition 1 doivent également se trouver loin des frontières supposées, par rapport aux zones de transition estimées d'après les moments statistiques, et à la succession attendue des configurations lorsque le débit de gaz (et dans une moindre mesure, le débit de liquide) augmente, par rapport à la littérature.</p> | |

Définition de profils typiques d'après l'allure des moments statistiques

Comme on l'avait vu dans la section 5.2, on peut distinguer trois types de profils différents sur l'ensemble des courbes représentant les 4 moments statistiques du signal brut de la sonde, pour chaque point. Ce sont principalement les valeurs de l'asymétrie et du kurtosis qui permettent de distinguer les trois groupes, comme l'illustre le Tableau 6.4.

Afin de visualiser les regroupements retenus, la Figure 6.5 présente l'ensemble de ces profils dans un tableau, dans le même ordre que sur les cartes tracées dans le repère des vitesses superficielles. On rappelle que les profils des moments statistiques sont donnés en Annexe C pour l'ensemble des points. Les profils repérés ici sont ceux des points qui respectent les deux conditions précédemment fixées : l'ensemble des observations visuelles pour un point donné est en accord avec le regroupement effectué d'après les profils des moments, et les points se trouvent loin des zones incertaines qui semblent être des transitions entre deux profils bien définis.

Tableau 6.4 : Types de profils de moments statistiques

| Groupe 1 | <p>Moments statistiques du signal de la sonde beta=12 %, JI=0.60 m/s</p>  <p>A bar chart with a y-axis from -4 to 14 and an x-axis with categories M, Std, Skw, and K. The bars are blue and have the following values: M=0.29, Std=0.34, Skw=0.90, and K=1.85.</p> <table border="1"><thead><tr><th>Moment</th><th>Value</th></tr></thead><tbody><tr><td>M</td><td>0.29</td></tr><tr><td>Std</td><td>0.34</td></tr><tr><td>Skw</td><td>0.90</td></tr><tr><td>K</td><td>1.85</td></tr></tbody></table> | Moment | Value | M | 0.29 | Std | 0.34 | Skw | 0.90 | K | 1.85 |
|----------|--|--------|-------|---|------|-----|------|-----|-------|---|------|
| Moment | Value | | | | | | | | | | |
| M | 0.29 | | | | | | | | | | |
| Std | 0.34 | | | | | | | | | | |
| Skw | 0.90 | | | | | | | | | | |
| K | 1.85 | | | | | | | | | | |
| Groupe 2 | <p>Moments statistiques du signal de la sonde beta=55 %, JI=0.25 m/s</p>  <p>A bar chart with a y-axis from -4 to 14 and an x-axis with categories M, Std, Skw, and K. The bars are blue and have the following values: M=0.44, Std=0.35, Skw=-0.23, and K=1.08.</p> <table border="1"><thead><tr><th>Moment</th><th>Value</th></tr></thead><tbody><tr><td>M</td><td>0.44</td></tr><tr><td>Std</td><td>0.35</td></tr><tr><td>Skw</td><td>-0.23</td></tr><tr><td>K</td><td>1.08</td></tr></tbody></table> | Moment | Value | M | 0.44 | Std | 0.35 | Skw | -0.23 | K | 1.08 |
| Moment | Value | | | | | | | | | | |
| M | 0.44 | | | | | | | | | | |
| Std | 0.35 | | | | | | | | | | |
| Skw | -0.23 | | | | | | | | | | |
| K | 1.08 | | | | | | | | | | |
| Groupe 3 | <p>Moments statistiques du signal de la sonde beta=96 %, JI=0.40 m/s</p>  <p>A bar chart with a y-axis from -4 to 14 and an x-axis with categories M, Std, Skw, and K. The bars are blue and have the following values: M=0.69, Std=0.22, Skw=-2.27, and K=6.72.</p> <table border="1"><thead><tr><th>Moment</th><th>Value</th></tr></thead><tbody><tr><td>M</td><td>0.69</td></tr><tr><td>Std</td><td>0.22</td></tr><tr><td>Skw</td><td>-2.27</td></tr><tr><td>K</td><td>6.72</td></tr></tbody></table> | Moment | Value | M | 0.69 | Std | 0.22 | Skw | -2.27 | K | 6.72 |
| Moment | Value | | | | | | | | | | |
| M | 0.69 | | | | | | | | | | |
| Std | 0.22 | | | | | | | | | | |
| Skw | -2.27 | | | | | | | | | | |
| K | 6.72 | | | | | | | | | | |

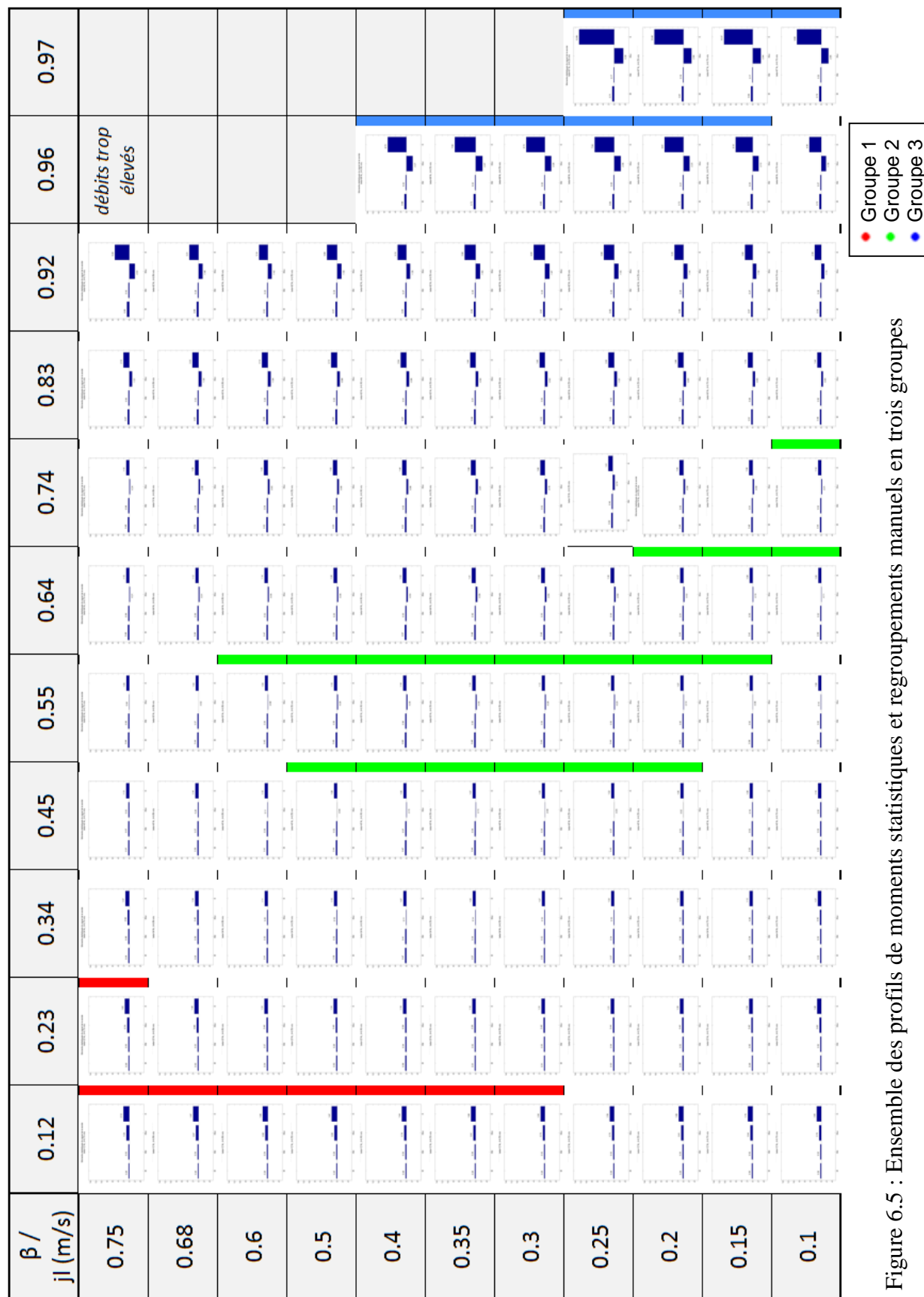


Figure 6.5 : Ensemble des profils de moments statistiques et regroupements manuels en trois groupes

Points de référence retenus pour chaque configuration

D'après la correspondance entre les photos prises comme références pour chacune des trois configurations définies à la section 4.2 et les groupes associés à ces mêmes points d'après les moments statistiques, on peut attribuer les points du groupe 1 à la configuration à bulles, ceux du groupe 2 à la configuration intermittente, et ceux du groupe 3 à la configuration semi-annulaire, telles que définies ci-dessus.

On obtient ainsi la liste de points de référence présentée dans le Tableau 6.5, qui est en accord avec la succession des configurations que l'on observe dans la littérature.

Tableau 6.5 : Points de référence pour les trois configurations

| # des points | Titre volumique - modèle homogène | | | | | | | | | | | |
|----------------|-----------------------------------|------|------|------|------|------|------|------|------|------|------|------|
| | β | 0.12 | 0.23 | 0.34 | 0.45 | 0.55 | 0.64 | 0.74 | 0.83 | 0.92 | 0.96 | 0.97 |
| j_t (m/s) | 0.75 | 11 | 22 | 33 | 44 | 55 | 66 | 77 | 88 | 99 | - | - |
| | 0.68 | 10 | 21 | 32 | 43 | 54 | 65 | 76 | 87 | 98 | - | - |
| | 0.60 | 9 | 20 | 31 | 42 | 53 | 64 | 75 | 86 | 97 | - | - |
| | 0.50 | 8 | 19 | 30 | 41 | 52 | 63 | 74 | 85 | 96 | - | - |
| | 0.40 | 7 | 18 | 29 | 40 | 51 | 62 | 73 | 84 | 95 | 106 | - |
| | 0.35 | 6 | 17 | 28 | 39 | 50 | 61 | 72 | 83 | 94 | 105 | - |
| | 0.30 | 5 | 16 | 27 | 38 | 49 | 60 | 71 | 82 | 93 | 104 | - |
| | 0.25 | 4 | 15 | 26 | 37 | 48 | 59 | 70 | 81 | 92 | 103 | 110 |
| | 0.20 | 3 | 14 | 25 | 36 | 47 | 58 | 69 | 80 | 91 | 102 | 109 |
| | 0.15 | 2 | 13 | 24 | 35 | 46 | 57 | 68 | 79 | 90 | 101 | 108 |
| | 0.10 | 1 | 12 | 23 | 34 | 45 | 56 | 67 | 78 | 89 | 100 | 107 |



Config. à bulles



Config. intermittente



Config. semi-annulaire

La configuration de ces points est dorénavant supposée connue et représentative de la réalité, afin de pouvoir choisir parmi eux des points d'entraînement pour la SVM. Inversement, on veut pouvoir attribuer une configuration à tous les autres points de mesure grâce au modèle final de cette SVM, ce qui permettra finalement d'obtenir une carte d'écoulement plus objective.

6.2.3 Implémentation d'une SVM

La procédure employée est la suivante, qui s'inspire des recommandations des sources consultées :

- 1) Mise à l'échelle des données
- 2) Répartition aléatoire des données dont on connaît la classe en *training* et *test sets*, avec un ratio de 70%
- 3) *3-fold cross validation* par répartition aléatoire des données du *training set* en 3 groupes qu'on classe successivement les uns par rapport aux autres, pour différentes valeurs pour C et γ . On garde ensuite le couple (C, γ) optimal qui permet de classer correctement 100% des données lors de cette étape.
- 4) Entraînement de la SVM pour ce couple (C, γ) sur l'ensemble du *training set*
- 5) Classification des données du *test set* par la SVM ainsi obtenue et estimation de la capacité de généralisation de cette SVM
- 6) Classification des données de classe inconnue par le modèle final retenu

A cause du caractère aléatoire de la répartition initiale en *training* et *test sets*, on peut obtenir des classifications finales différentes, même en choisissant des paramètres identiques pour les autres étapes. Une classification moyenne globale sur un certain nombre d'exécutions des étapes 2 à 6 sera donc nécessaire pour obtenir un résultat indépendant de ce choix initial des *sets*. Ces exécutions seront repérées par un numéro, le numéro 1 correspondant à la première exécution du programme après démarrage de Matlab (ou réinitialisation du générateur de nombres aléatoires, ce qui revient ici au même).

Le fonctionnement de la SVM est ici illustré pour les 4 premières exécutions. La mise à l'échelle des données est réalisée une fois pour toutes au début, comme présenté plus haut. Les 4 composantes des vecteurs représentant les 110 points de mesure varient donc toutes entre 0 et 1.

6.2.3.1 Répartition aléatoire en *training* et *test sets*

Parmi les points de référence définis à la section 6.2.2, on choisit donc aléatoirement environ 70% des points au sein de chaque classe pour constituer les données d'entraînement du *training set*. Les données restantes sont affectées au *test set*.

La Figure 6.6 montre la répartition aléatoire obtenue pour la 1^{ère} exécution, où les cercles représentent les données du *training set*, et les losanges celles du *test set*, avec la couleur correspondant à leur « vraie » classe.

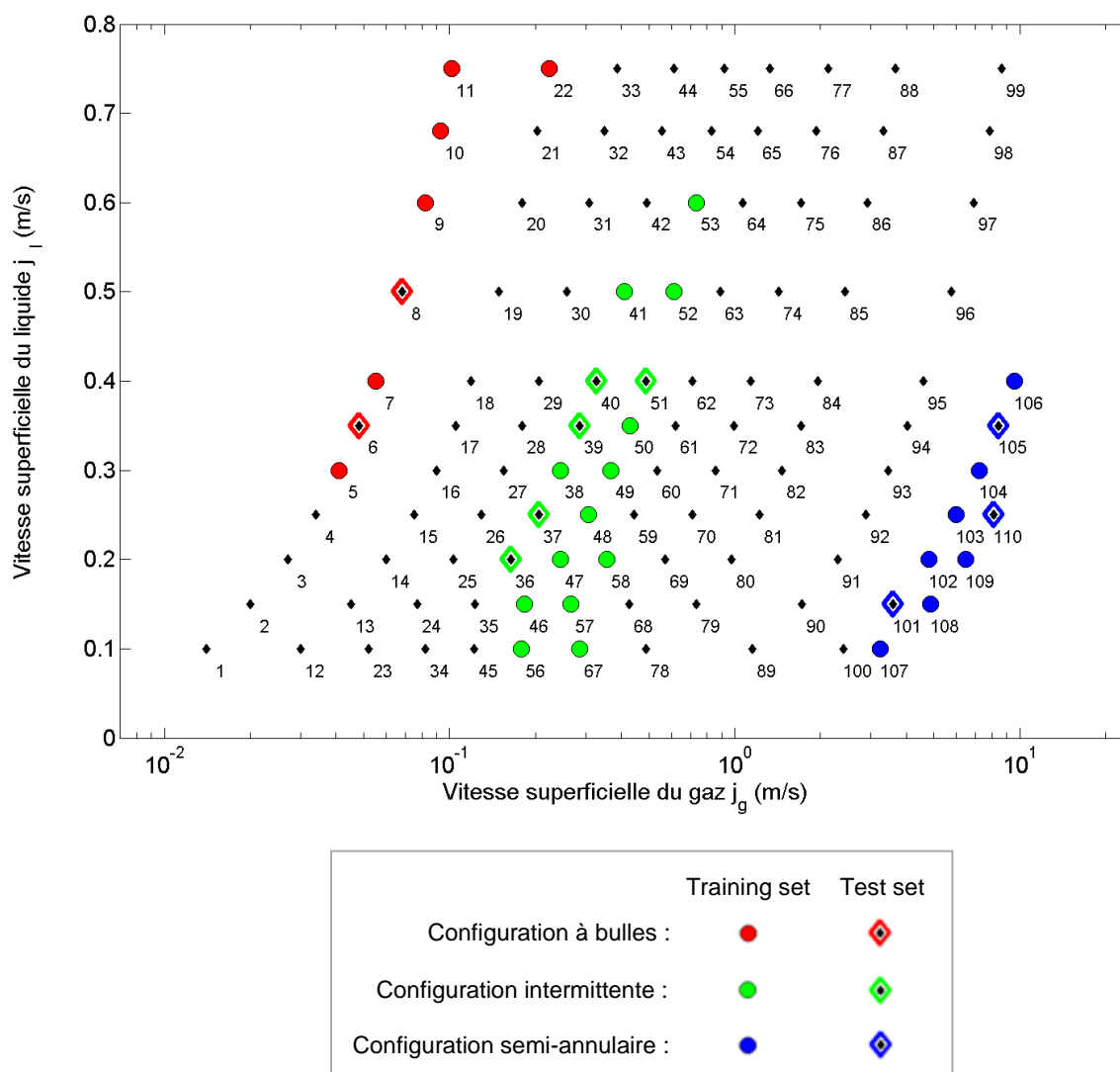


Figure 6.6 : Répartition aléatoire en *training* et *test sets*, exécution 1

6.2.3.2 Choix des paramètres C et γ lors de la *cross validation*

Les paramètres C et γ doivent être choisis par l'utilisateur. Il est crucial d'optimiser ce choix, car sinon les résultats, notamment la classification finale, donnés par la SVM peuvent être mauvais. On peut se reporter à l'annexe E qui donne les résultats détaillés obtenus avec les valeurs par défaut, pour illustrer les conséquences de l'absence d'optimisation.

Signification de C et γ et influence sur les frontières obtenues

Il faut savoir qu'un modèle de SVM construit à partir des données d'entraînement est le meilleur qui est trouvé *pour les paramètres C et γ choisis*. Le paramètre C est l'importance accordée aux erreurs de classification, et le paramètre γ permet d'ajuster les valeurs de la fonction-noyau, et donc de se placer dans un espace-image ajusté aux données. Concrètement, ces deux paramètres ont le même effet sur la frontière obtenue et la classification des données.

Si C et/ou γ sont trop élevés, la SVM va essayer d'éviter à tout prix les erreurs de classification et on va donc obtenir une frontière complexe pour que toutes les données d'entraînement soient bien classées. Le problème est alors que cette SVM aura une faible capacité de généralisation : comme la frontière est compliquée, il y a un risque élevé que de nouveaux points situés trop près de cette frontière soient classés assez arbitrairement d'un côté ou de l'autre. Dans ce cas, la SVM tient trop compte des points « aberrants » (*outliers*) c'est-à-dire statistiquement « loin » des autres points, c'est pourquoi ce problème est appelé *overfitting*. Il est illustré plus en détails à la fin de l'Annexe E pour l'espace image. Dans l'espace de départ, l'*overfitting* se traduirait par une frontière trop complexe comme celle de la Figure 6.7 où on voit que la SVM tient trop compte des points aberrants.

Inversement, si C et/ou γ sont trop faibles, la frontière obtenue sera trop simple par rapport aux données et la classification des données d'entraînement comportera trop d'erreurs, comme l'illustre la Figure 6.8 : on parle alors d'*underfitting*.

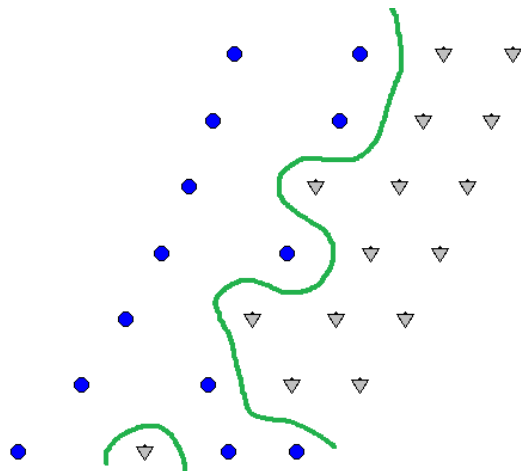


Figure 6.7 : Illustration de l'*overfitting* dans l'espace de départ

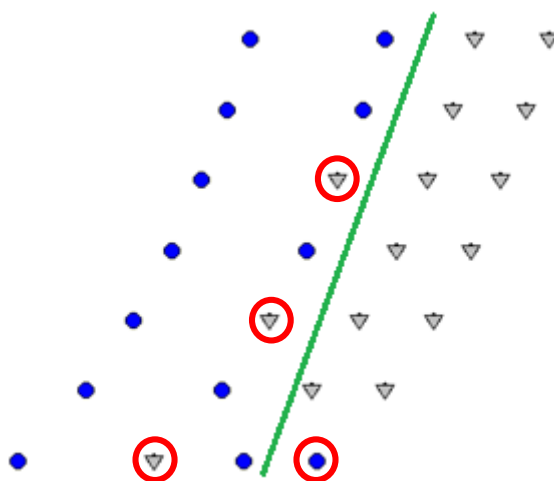


Figure 6.8 : Illustration de l'*underfitting* dans l'espace de départ

Dans ce cas, les données d'entraînement qui sont entourées sont mal classées par la SVM.

Il faut donc ajuster les valeurs de C et γ aux données étudiées, ainsi la valeur optimale de ces paramètres ne peut être connue à l'avance.

En calculant le pourcentage d'exactitude de la classification des données d'entraînement pour différentes valeurs de C et γ , on peut garder les couples qui optimisent cette exactitude. C'est l'objectif du processus de *cross validation*.

Sélection de paramètres par 3-fold cross validation

L'idée de la *cross validation* (CV) est d'utiliser un ensemble intermédiaire de données non utilisées dans l'entraînement de la SVM pour sélectionner les paramètres C et γ les plus appropriés, mais qui soit également distinct du *test set* sans quoi l'estimation de la capacité à généraliser effectuée par la suite serait biaisée. Une des méthodes préconisées par Hsu et al. (2003) est d'effectuer une *n-fold cross validation*.

Le principe est de générer artificiellement de nouveaux sous-ensembles de données « inconnues » au sein du *training set*. Plus exactement, on divise aléatoirement le *training set* en n sous-ensembles de taille égale. On considère successivement chaque sous-ensemble, dont on prédit les classes en utilisant une SVM entraînée sur les données des $n-1$ autres sous-ensembles, pour des valeurs données de C et γ . On peut ainsi prédire une classification de chacun des points du *training set*, à la fois comme données d'entraînement et comme s'il s'agissait de nouvelles données, tout en connaissant leurs vraies classes, grâce auxquelles on peut estimer l'exactitude globale de ce processus. Lorsque la valeur de cette exactitude est élevée, cela signifie donc qu'on a bien classé les données à la fois en tant que données d'entraînement (ce qui permet d'éviter l'*underfitting*) et en tant que pseudo-nouvelles données (ce qui permet d'éviter l'*overfitting*, en assurant une bonne capacité de généralisation de la frontière obtenue à de nouvelles données).

Au lieu de tester indépendamment différentes valeurs pour les paramètres C et γ , on effectue directement le processus de *cross validation* pour plusieurs couples (C, γ) , en définissant un intervalle de valeurs pour chaque paramètre, ce qui donne finalement une grille de couples possibles. Une méthode usuelle est de faire varier C et γ de façon exponentielle, par exemple entre 2^{-10} et 2^5 . Pour chaque couple (C, γ) , on effectue une procédure de *n-fold cross validation* afin de calculer le pourcentage d'exactitude correspondant. En affinant les intervalles de la grille selon les résultats obtenus, on peut finalement choisir un couple (C, γ) qui donne la plus grande exactitude possible.

C'est la méthode qui sera mise en pratique ici, avec $n=3$, étant donné le nombre limité de données d'entraînement pour certaines classes.

La Figure 6.9 donne un exemple de grille de *cross validation* où la couleur de chaque carreau est proportionnelle au pourcentage d'exactitude obtenu lors de la *cross validation* du *training set*.

Les données d'entraînement sont celles qui ont été choisies aléatoirement lors de l'exécution 1, et C et γ varient entre 2^{-10} et 2^{10} pour avoir une première idée de la répartition (les coordonnées des axes sont ici les valeurs des puissances de 2 – NB : les axes n'ont pas la même origine).

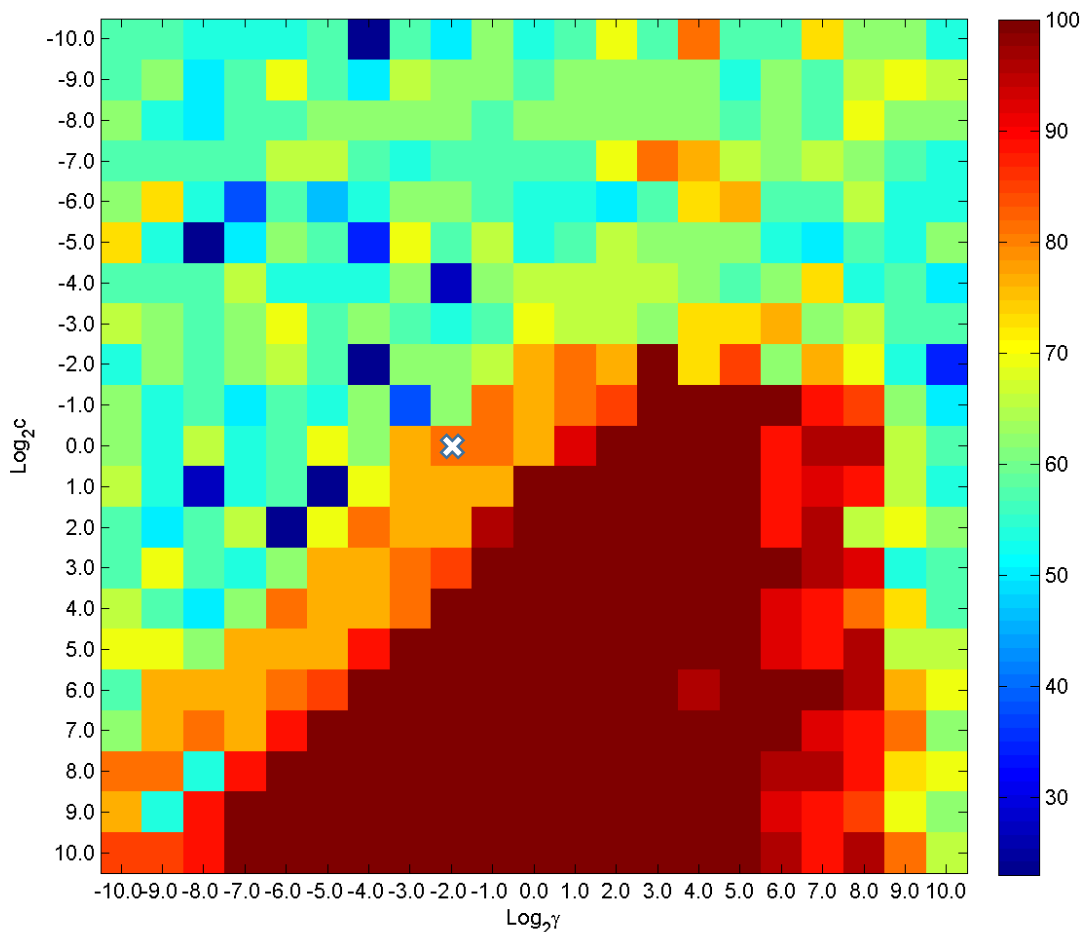


Figure 6.9 : Exemple de grille de CV obtenue pour l'exécution 1 :

Pourcentage d'exactitude de chaque 3-fold CV sur le *training set* en fonction de (C, γ)

La croix blanche repère le couple des valeurs par défaut pour $n_{feat}=4$ ($\log_2(C)=0$ et $\log_2(\gamma)=-2$), pour lequel l'exactitude globale de la CV vaut 80.8%. On voit qu'un certain nombre d'autres couples permettent d'obtenir une exactitude de 100% sur la CV, et que le choix des valeurs par défaut n'est donc pas optimal (cf Annexe E), d'où la nécessité de suivre cette procédure.

Il reste alors à choisir les paramètres finaux parmi les couples dont l'exactitude vaut 100%.

Choix final d'un couple (C, γ)

On remarque sur les premières exécutions que la grille de CV a toujours la même forme de pointe, comme on peut le voir sur la Figure 6.9, même en élargissant l'intervalle des valeurs de C et γ à $[2^{-20}; 2^{20}]$. Pour une exécution donnée, il y a donc de nombreux couples (C, γ) tels que l'exactitude globale du processus de CV soit égale à 100%. Le critère retenu pour guider le choix final de C et γ est le fait qu'on veut être sûr d'éviter l'*overfitting* car on ne s'attend pas des frontières très complexes entre les configurations, et on veut préserver la capacité de généralisation de la SVM afin d'assurer une bonne classification des points inconnus. On va donc chercher à minimiser les valeurs de C et γ parmi les couples qui donnent 100% sur la CV, et ce pour chaque exécution, afin d'obtenir une classification qu'on espère optimale pour chaque répartition aléatoire en *training* et *test sets*.

Pour une exécution donnée, la grille de CV obtenue repose sur une répartition aléatoire des données d'entraînement en sous-ensembles de CV, classés les uns par rapport aux autres. Afin de garder une certaine sécurité, on ne choisira donc pas les couples donnant 100% qui sont situés sur la « frontière » de la pointe. Pour ce faire, on part de la plus petite valeur de C donnant 100%, à partir de laquelle on passe en revue tous les couples donnant 100% en augmentant γ pour un C donné, puis à nouveau C , jusqu'à trouver le premier couple tel que tous les couples voisins donnent 100%. Pour la première exécution, il s'agit du couple $(2^0=1 ; 2^4=16)$ tel qu'illustré Figure 6.10. Cela permet d'assurer une certaine marge au cas où un changement de répartition initiale dans la CV causerait une diminution de l'exactitude d'un couple donné, ce qui concernerait davantage les points de la frontière.

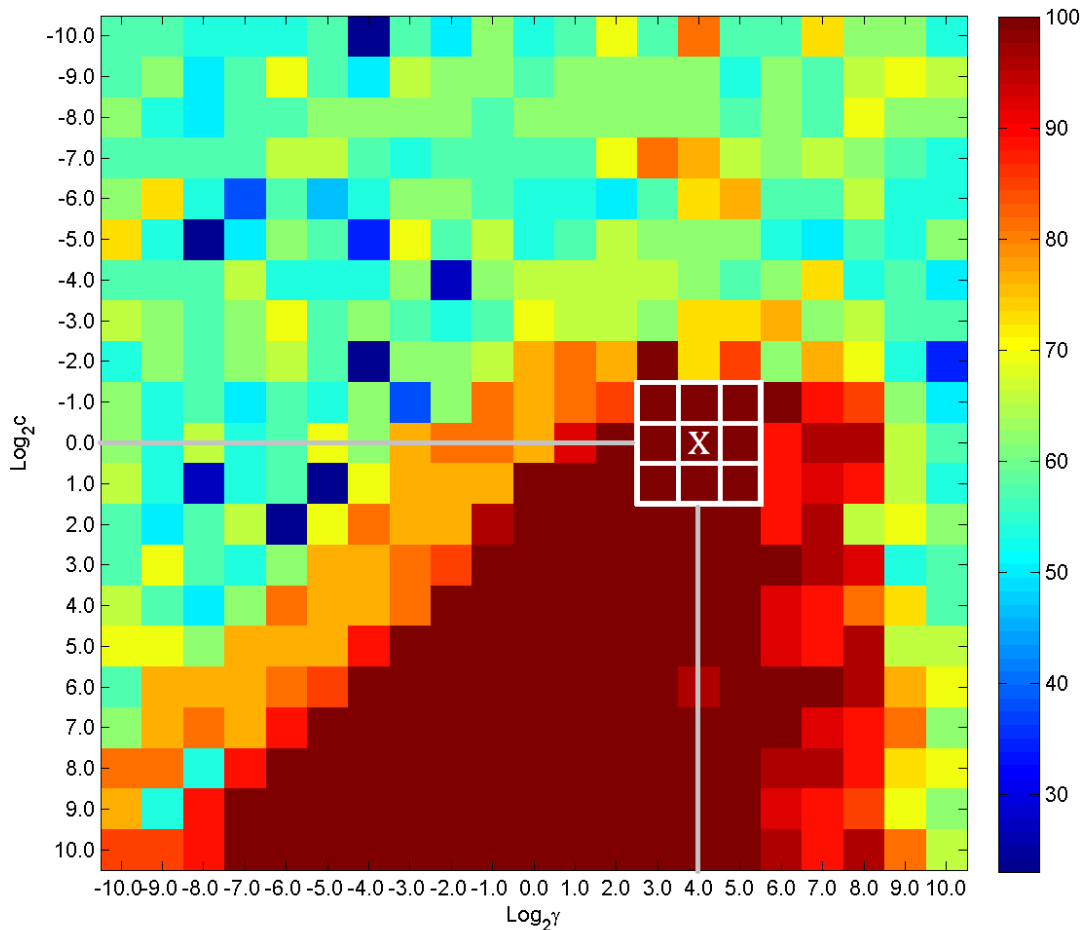


Figure 6.10 : Illustration du choix de C et γ pour une exécution donnée (#1)

6.2.3.3 Entraînement de la SVM et classification d'après le modèle obtenu

Pour une exécution donnée, on réalise donc l'entraînement de la SVM avec les données d'entraînement choisies, dont les valeurs des composantes sont transformées par la fonction-noyau gaussienne avec le couple (C, γ) retenu, afin de se placer dans un espace image dans lequel on peut séparer linéairement ces données d'entraînement. On fournit également à la SVM la classe qui a été attribuée à chaque donnée d'entraînement.

En résolvant le problème d'optimisation de la marge (distance entre l'hyperplan séparateur et les vecteurs de support de chaque classe), on obtient un modèle contenant les paramètres qui définissent cet hyperplan séparateur, qui est donc optimal pour *ces* données d'entraînement et *ces* paramètres C et γ .

Comme on est dans un cas de multi-classification, ce procédé d'apprentissage d'une frontière binaire est répété pour chaque classe selon une procédure *One Versus Rest* : on cherche la frontière entre la classe 1 et une classe artificielle composée de {2+3}, puis entre 2 et {1+3} et ainsi de suite. D'un point de vue technique, on entraîne donc en réalité 3 SVM. On construit ensuite un modèle global optimal à partir de ces trois modèles.

On utilise alors ce modèle global pour classer les données du *test set*, qui sont des données inconnues pour la SVM. Mais comme ces données ont été choisies parmi les données de référence, on connaît leur « vraie » classe, et on peut comparer la vraie classe avec celle qui est prédite par la SVM. Ces données permettent ainsi de tester la SVM et d'estimer sa performance de façon non biaisée (car le modèle construit ne s'est pas basé sur ces données de test, donc les considère comme des données complètement nouvelles).

On a constaté que l'exactitude de la classification du *test set* valait 100% pour toutes les exécutions effectuées d'après cette procédure. On peut donc estimer que les SVM obtenues ont donc une bonne capacité à généraliser et que les classifications proposées pour les points inconnus seront fiables.

Enfin, on peut obtenir grâce à la SVM ainsi construite une prédiction des données de configuration inconnue. Cela donne par exemple, pour la 1^{ère} exécution du programme, les prédictions de la Figure 6.11. Les données d'entraînement sont représentées par des cercles, celles du *test set* par de petits losanges pleins, et la prédiction des points de classe inconnue est donnée par les losanges extérieurs plus grands.

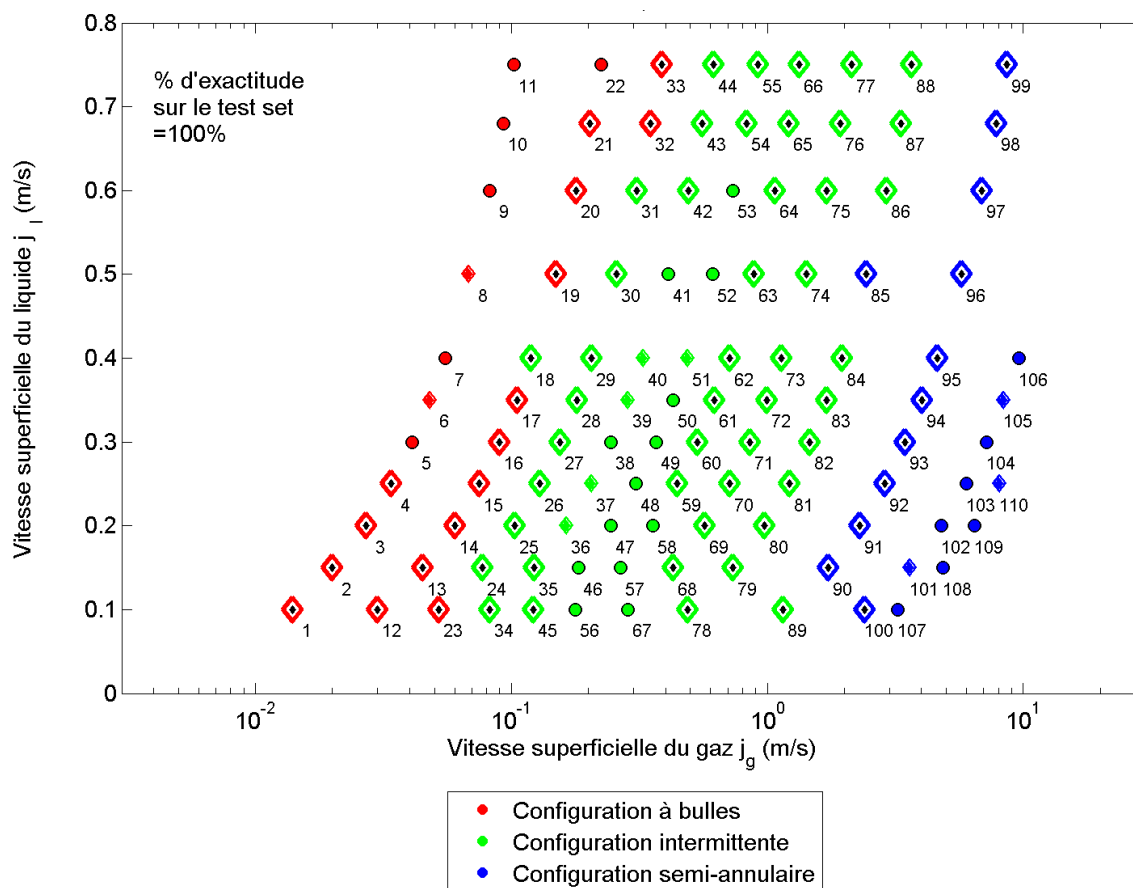


Figure 6.11 : Prédiction des configurations de l'ensemble des points pour l'exécution 1, $C=1, \gamma=16$

6.2.3.4 Classification finale

On obtient ainsi une classification par exécution, correspondant à un couple (C, γ) supposé optimal, et à une répartition initiale donnée. Afin d'avoir une classification finale qui soit indépendante de cette répartition initiale, on effectue une moyenne sur un grand nombre d'exécutions, ici 100 exécutions, en calculant simplement quelle classe est la plus attribuée à chaque point, et la probabilité d'obtenir cette classe moyenne sur les 100 exécutions.

Cela donne pour les 100 premières exécutions la classification finale de la Figure 6.12, où la vraie classe des points de référence (*training* et *test sets*) est représentée par les cadres carrés et la classe prédite pour les autres points est représentée par un carré plein dont la taille est

proportionnelle à la probabilité de la classe moyenne sur les 100 exécutions, qui donne la couleur du carré, l'autre classe parfois prédite étant représentée par le liseré coloré. Par exemple, la configuration prédite du point 20 est en moyenne « à bulles », mais il y a eu une ou plusieurs exécutions pour lesquelles ce point a été prédit en configuration intermittente. Une bordure noire signifie que le point est toujours classé dans la même configuration quel que soit le numéro d'exécution du programme.

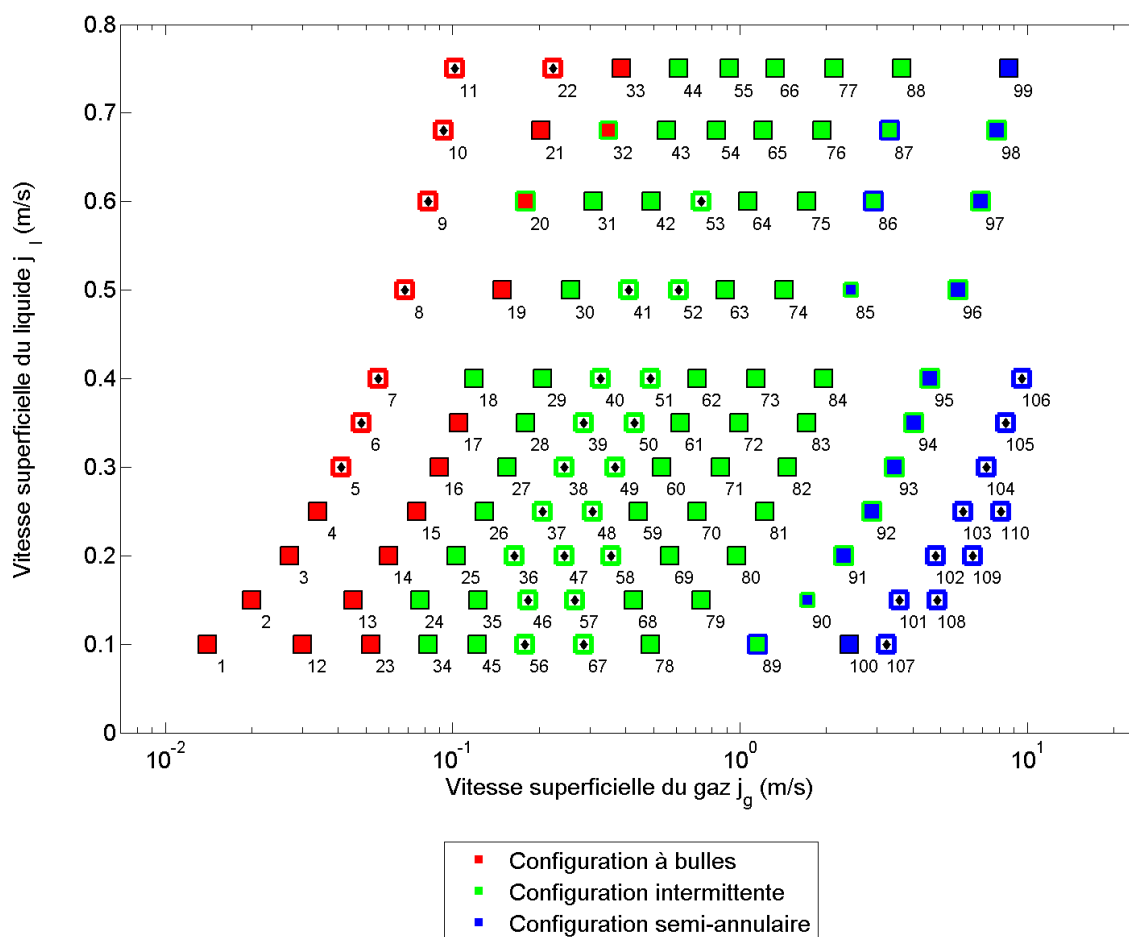


Figure 6.12 : Classification finale moyenne sur les 100 premières exécutions ($n_{feat}=4$)

6.3 Analyse des classifications obtenues et choix final des paramètres

Outre le choix des paramètres C et γ pour chaque exécution, on peut faire varier les composantes des vecteurs (*features*, qui contiennent les données d'entrée de la SVM), et notamment leur nombre, n_{feat} . Les classifications moyennes obtenues pour $n_{feat}=4$ et $n_{feat}=2$ seront ici comparées.

6.3.1 Classification obtenue pour $n_{feat}=4$

La classification moyenne obtenue sur les 100 premières exécutions, en gardant tous les moments statistiques comme composantes des vecteurs, est celle de la Figure 6.12 ci-dessus.

Cette classification moyenne est plus pertinente que celle qu'on aurait obtenue pour les paramètres par défaut de C et γ (cf Annexe E), principalement au niveau de la configuration à bulles. L'allure des moments statistiques ainsi que les observations visuelles suggèrent en effet un rattachement à cette configuration des points 1 à 4, et 12 à 21 dans une moindre mesure, mais de façon moins marquée que pour les points 5 à 11, c'est pourquoi on avait préféré laisser la SVM prédire leur classification.

On retrouve le fait que c'est principalement l'augmentation du débit d'air qui génère le changement de configuration, avec une exception marquée pour les points de la configuration à bulles de vitesse superficielle du liquide élevée (0.68 m/s et 0.75 m/s) : #32 et 33. On observait aussi cette disparité expérimentalement.

On constate que les frontières se situent globalement entre les « lignes » $\beta=23\%$ et $\beta=34\%$ pour les configurations à bulles et intermittente, et $\beta=83\%$ et $\beta=96\%$ pour les configurations intermittente et semi-annulaire. La ligne $\beta=92\%$ semble être située dans la zone de transition entre ces deux configurations bien qu'elle soit majoritairement prédite en configuration semi-annulaire dans ces conditions.

Bien qu'il soit situé près de la transition entre les configurations à bulles et intermittente, le point #18 est toujours prédit en configuration intermittente contrairement à ses voisins, ce qui est un peu étrange, mais explicable par sa proximité avec la frontière.

6.3.2 Classification obtenue pour $n_{feat}=2$

Dans l'utilisation des SVM, il est intéressant d'essayer de réduire le nombre de composantes des vecteurs, car cela signifie qu'on ne retient que les grandeurs qui synthétisent le mieux l'information liée au phénomène étudié, ici à la configuration de l'écoulement diphasique. Il existe différentes méthodes d'optimisation de ces composantes, comme la recherche de combinaisons plus synthétiques à partir des grandeurs physiques, ce qui permet de réduire leur nombre. Mais une des premières possibilités est d'effectuer directement l'entraînement et la classification sur un nombre réduit de composantes.

Dans le contexte étudié, si l'on revient au sens physique des moments statistiques et leur équivalent observable sur l'allure des courbes de PDF (cf Figure 5.15 à Figure 5.17), on peut voir que certains moments ont un impact plus significatif que d'autres.

En effet, la configuration attribuée à un écoulement donné est fortement liée à la probabilité qu'une phase ou l'autre de l'écoulement soit majoritaire (ou au contraire, au fait d'avoir une probabilité semblable, en configuration intermittente). Concernant les graphes des fonctions de densité de probabilité (PDF) du taux de vide, d'après Jones et Zuber (1975), ce qui permet de différencier les configurations est la présence et la position des pics : on s'attend à un seul pic de densité de probabilité proche des faibles valeurs du taux de vide pour la configuration à bulles, un seul pic proche des valeurs élevées du taux de vide pour la configuration annulaire, et donc deux pics pour ces deux extrêmes, d'amplitude moindre, pour la configuration intermittente. Or il se trouve que les moments statistiques qui expriment ces différences d'allure sont l'asymétrie et le kurtosis, qui sont dénommés « paramètres de forme », contrairement à la moyenne et à l'écart-type qui sont des paramètres d'échelle. En effet, on rappelle que l'asymétrie quantifie l'absence de symétrie du signal, et le kurtosis la présence et l'amplitude des pics dans le signal. Inversement, la moyenne reflète davantage l'influence de l'augmentation continue des débits d'air et d'eau et donc l'évolution du titre volumique.

Dans le cas présent, la sonde optique donne un signal qui indique les changements de phase et la proportion de chaque phase dans le temps. Le fait de quantifier l'asymétrie et la présence de pics dans ce signal va donc suffire à traduire cette prédominance (ou non) d'une phase sur l'autre, et permettre d'identifier la configuration, tout comme les graphes de PDF, mais avec les avantages présentés à la section 5.3.

Pour cette raison, on s'est intéressé à la classification moyenne obtenue dans les mêmes conditions que précédemment, mais en ne conservant que l'asymétrie et le kurtosis du signal brut comme composantes des vecteurs :

$$x_{i,2} = \left\{ \begin{array}{c} Skw(i) \\ K(i) \end{array} \right\}, i = 1, 2, \dots, 110 \quad (6.11)$$

La classification moyenne obtenue, par la procédure décrite pour le choix de C et γ pour chaque exécution, pour les 100 premières exécutions, est donnée Figure 6.13.

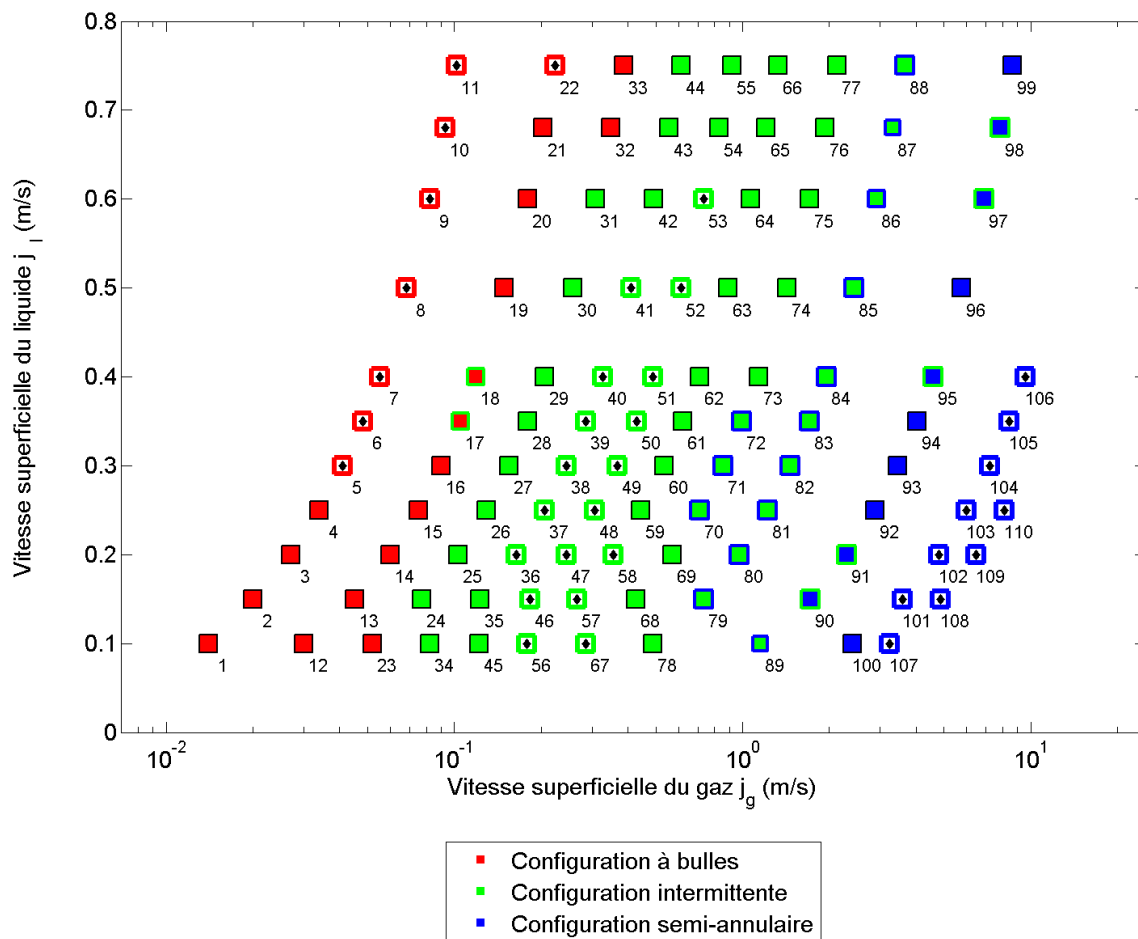


Figure 6.13 : Classification finale, moyenne sur les 100 premières exécutions ($n_{feat}=2$)

L'attribution des points de la configuration à bulles paraît plus satisfaisante, il y a encore 32% des cas où le point 18, et le 17 à présent, sont prédits en configuration intermittente, ce qui dénote probablement une plus grande proximité avec la frontière.

Les zones de transition sont situées aux mêmes endroits que précédemment, sauf qu'on observe quelques cas assez rares où les points 70 à 72 et 79 à 84 sont prédits dans la configuration semi-annulaire (respectivement 2% des cas pour les premiers et entre 3 et 15% des cas pour les seconds). Ce qui n'était pas le cas pour $n_{feat}=4$.

En observant les valeurs de C et γ retenues pour ces couples, on s'aperçoit qu'il s'agit de valeurs de γ bien supérieures aux autres ($\gamma_{max}=128$), à cause de la procédure qui minimise d'abord C , et de l'évolution en puissances de 2. On s'écarte donc du but recherché qui était de minimiser le couple (C,γ) . Pour $n_{feat}=4$, la valeur maximale retenue était $\gamma=16$. En ne retenant que les couples (C,γ) tels que $\gamma_{max}=16$, on obtient la classification présentée Figure 6.14. Il faut noter que la moyenne est alors calculée sur 66 exécutions au lieu de 100, mais on obtient un résultat plus cohérent vis-à-vis de la transition à bulles/intermittente.

Finalement, si l'on compare cette classification avec celle de la Figure 6.12, obtenue dans les mêmes conditions (si ce n'est le nombre d'exécutions et de composantes), on obtient ici une meilleure classification moyenne pour la configuration à bulles, et une frontière plus marquée entre les lignes à $\beta=83\%$ et $\beta=92\%$ pour la transition entre les configurations à bulles et intermittente. Toutefois, certains points à $\beta=92\%$ sont encore classés en configuration intermittente (pour 5% à 21% des cas selon les points), ce qui confirme une certaine proximité avec la frontière.

On gardera donc comme **classification finale** celle de la Figure 6.14, obtenue pour un vecteur réduit à deux composantes, Skw et K , l'asymétrie et le kurtosis, qui synthétisent à eux seuls le lien avec la configuration de l'écoulement. Les paramètres internes à la SVM, C et γ , ont été optimisés pour chaque exécution et prennent respectivement les valeurs $\{1;2\}$ et $\{8;16\}$ selon les exécutions du programme.

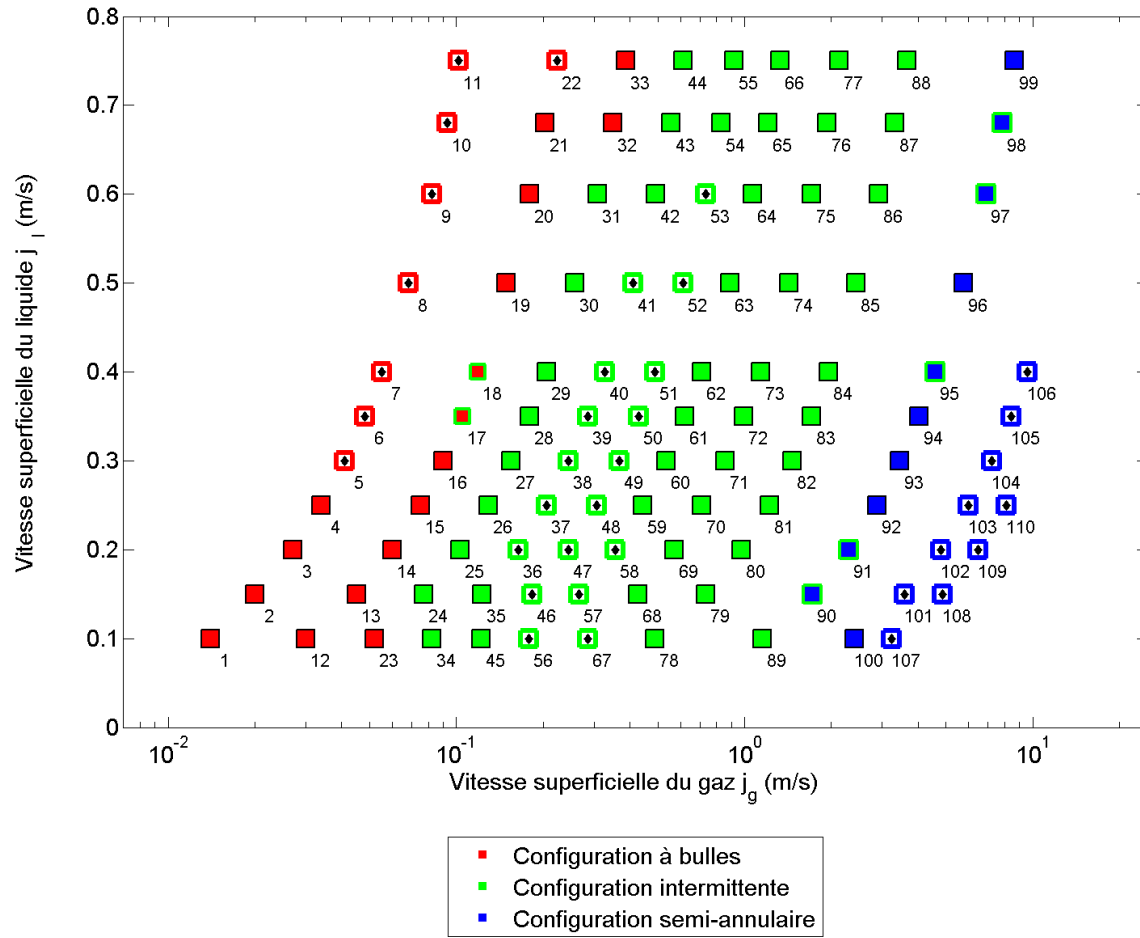


Figure 6.14 : Classification finale, moyenne sur les 100 premières exécutions telles que $\gamma_{\max}=16$
 $(n_{feat}=2)$

CHAPITRE 7 CARTE D'ÉCOULEMENT OBTENUE ET COMPARAISON AVEC LA LITTÉRATURE

A partir des moments statistiques du signal donné par la sonde optique, on a donc pu entraîner et utiliser une SVM pour prédire une classification moyenne des points de mesures, jugée satisfaisante à l'issue des différentes procédures d'optimisation des paramètres internes (C et γ) et externes (nombre de composantes n_{feat}) de la SVM. Ce chapitre présente les frontières de la carte d'écoulement que l'on peut déduire de cette classification, et compare ces frontières avec celles que l'on trouve dans la littérature.

7.1 Carte d'écoulement obtenue

La Figure 7.1 présente l'estimation que l'on peut faire des frontières à partir de la prédiction moyenne finale (celle de la Figure 6.14).

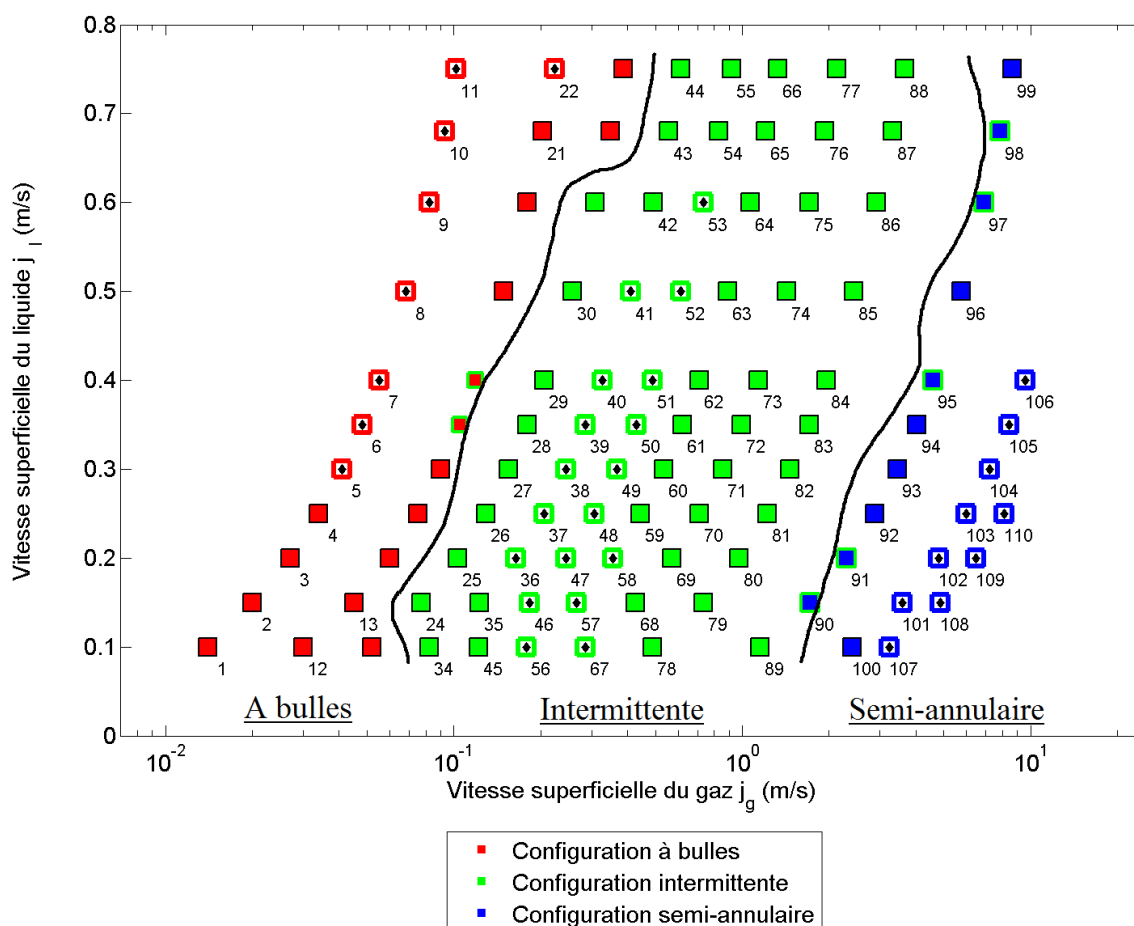


Figure 7.1 : Carte d'écoulement finale et frontières des configurations proposées

On a choisi de ne pas lisser les frontières afin de tenir compte de la probabilité des prédictions moyennes de la SVM. Il faut noter que ces frontières dépendent forcément de la position des points de mesure et sont donc relatives, tout en donnant une bonne idée de l'emplacement des transitions.

7.2 Comparaison avec la littérature

Étant donné que notre section d'essai est une portion de faisceau de tubes, c'est-à-dire ni une conduite verticale, ni un faisceau à proprement parler, on veut ici comparer la classification obtenue avec les cartes d'écoulement qu'on retrouve dans la littérature, principalement les cartes pour les faisceaux de tubes, mais aussi la carte de Taitel et al. pour une conduite verticale.

7.2.1 Comparaison avec les cartes obtenues en faisceaux de tubes

On reprend la comparaison effectuée par Ribatski et Thome (2007), qui compare d'une part les cartes d'écoulement basées sur des observations visuelles : Grant et Chisholm (1979) et Xu et al. (1998), et d'autre part celles qui sont basées sur des mesures objectives : Ulbrich et Mewes (1994) et Noghrehkar et al. (1999). La Figure 7.2 et la Figure 7.3 montrent la superposition de ces frontières avec la classification finale prédite par la SVM.

Configuration à bulles

Les points attribués à cette configuration par la SVM ne peuvent pas vraiment être comparés avec la carte de Grant & Chisholm car il n'est pas prudent de tenter de prolonger les frontières, mais dans tous les cas les points les plus proches de la zone « A bulles » de Grant ne sont pas du tout prédits dans cette configuration par la SVM. On n'observe pas d'accord entre la frontière prédite par la SVM et la frontière de Xu et al. De même, beaucoup de points qui seraient attribués à la configuration à bulles par les cartes d'Ulbrich & Mewes, et Noghrehkar et al., ont été prédits en configuration intermittente.

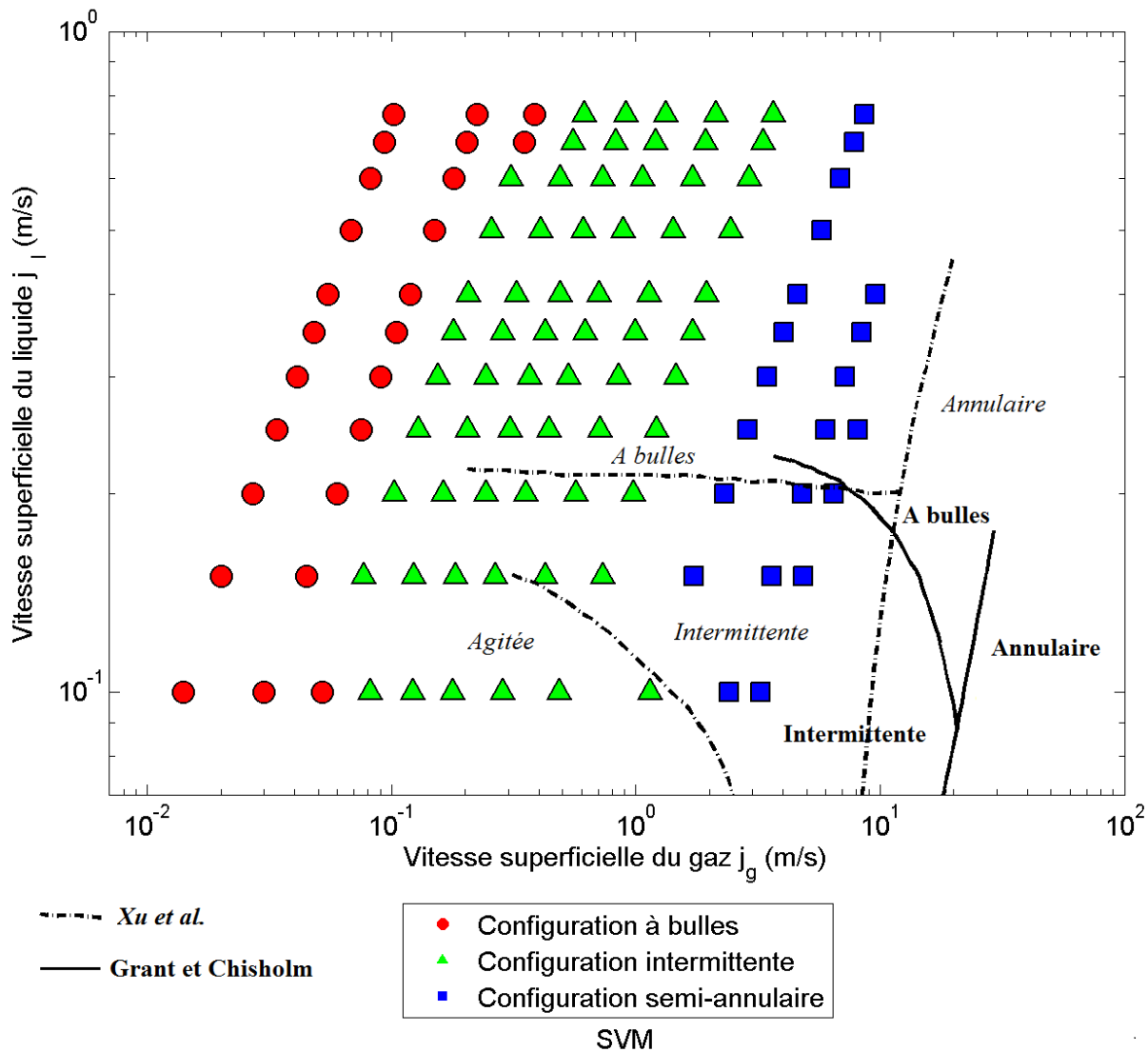


Figure 7.2 : Comparaison avec les frontières de Grant & Chisholm et Xu et al.

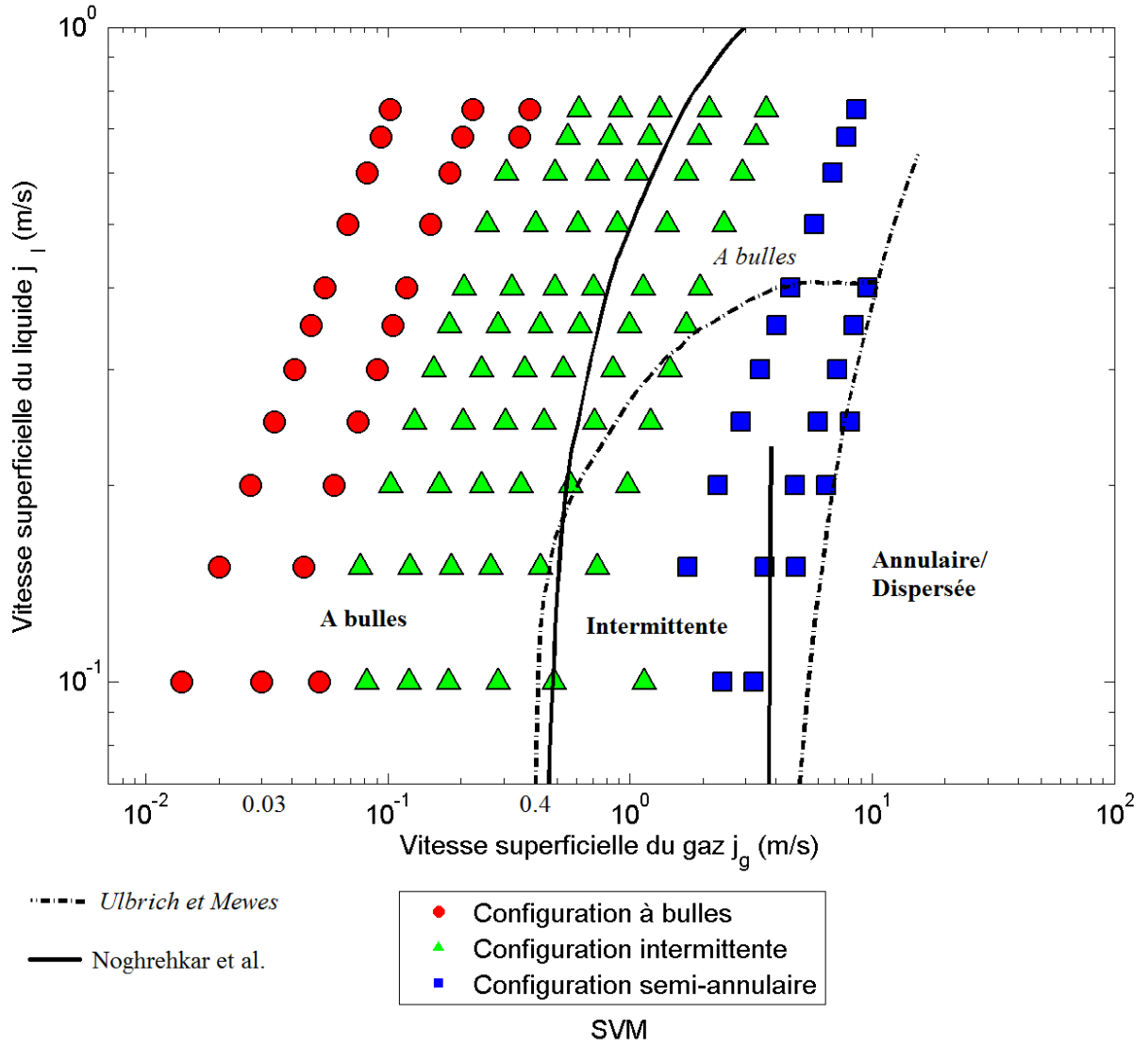


Figure 7.3 : Comparaison avec les frontières de Ulbrich & Mewes, et Noghrehkar et al.

Configurations intermittente et semi-annulaire

Pour certains auteurs, notamment Grant & Chisholm et Xu et al., la configuration intermittente correspond en fait davantage à la configuration semi-annulaire de cette étude qu'à la configuration qu'on a ici nommée intermittente. Les Figure 7.4 et Figure 7.5 (b), montrent en effet la succession de portions « à bulles » et de portions « annulaires », qu'on a définie ici comme étant caractéristique de la configuration semi-annulaire. Xu et al. décrivent aussi l'existence de « pulsations » dans cette configuration, ce qui correspond aux observations réalisées.



Figure 7.4 : Schéma de la configuration intermittente de Grant et Chisholm
(citée par Ulbrich et Mewes)

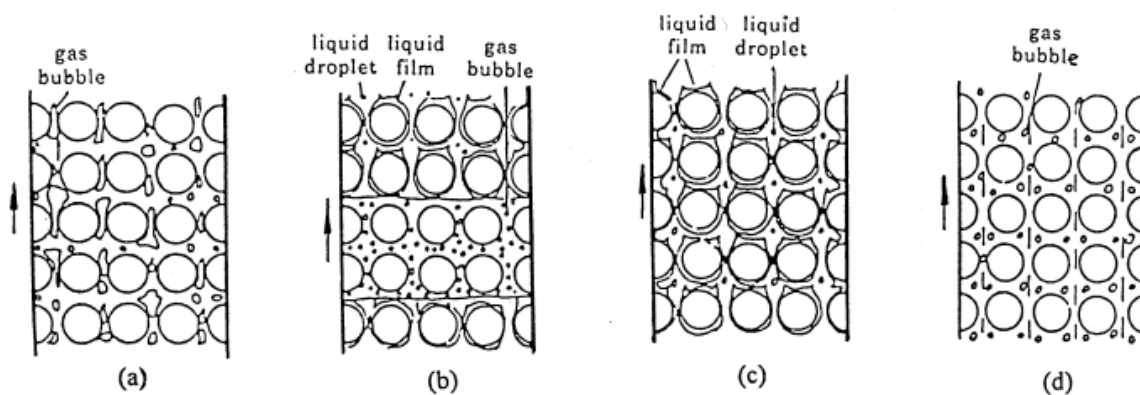


Figure 7.5 : Schémas des configurations décrites par Xu et al.
a) agitée, b) intermittente, c) annulaire et d) à bulles

Une fois cette correspondance établie, on constate un bon accord entre la configuration semi-annulaire de cette étude et la configuration intermittente de Grant & Chisholm et Xu et al., mais seulement pour les basses vitesses de liquide ($j_l < 0.15$ m/s) car on retrouve ensuite le désaccord lié à la configuration à bulles.

De façon similaire, les configurations intermittente et intermittente-dispersée de Ulbrich et Mewes (Tableau 4.2 et Figure 4.4) semblent correspondre globalement au regroupement des configurations intermittente et semi-annulaire de cette étude. Les points attribués à la configuration intermittente par Ulbrich et Mewes correspondent bien à des points des deux configurations équivalentes de cette étude, mais la réciproque n'est pas vraie, toujours à cause de la frontière avec la configuration à bulles. Noghrehkar et al. ne décrivent malheureusement pas les caractéristiques de leurs configurations dans leur article, ce qui limite la comparaison, mais on observe le même problème de réciprocity que pour Ulbrich et Mewes au niveau des frontières proposées.

De même, physiquement, la configuration agitée de Xu et al. - Figure 7.5 (a) -correspond assez bien à la configuration intermittente de cette étude, et à basse vitesse de liquide ($j_l < 0.15$ m/s), la prédiction de la SVM concorde assez bien avec la frontière proposée par Xu et al.

Frontière entre les configurations semi-annulaire et annulaire

C'est pour cette frontière qu'on a le meilleur accord entre les différentes cartes de la littérature comme l'illustre la Figure 7.6. Dans cette étude, la configuration annulaire n'a pas été atteinte, ce qui est en désaccord avec la frontière de Noghrehkar et al., et assez probablement avec celle de Ulbrich et Mewes. Il faudrait disposer de mesures à des titres volumiques plus élevés pour pouvoir conclure quant aux frontières proposées par Grant & Chisholm et Xu et al.

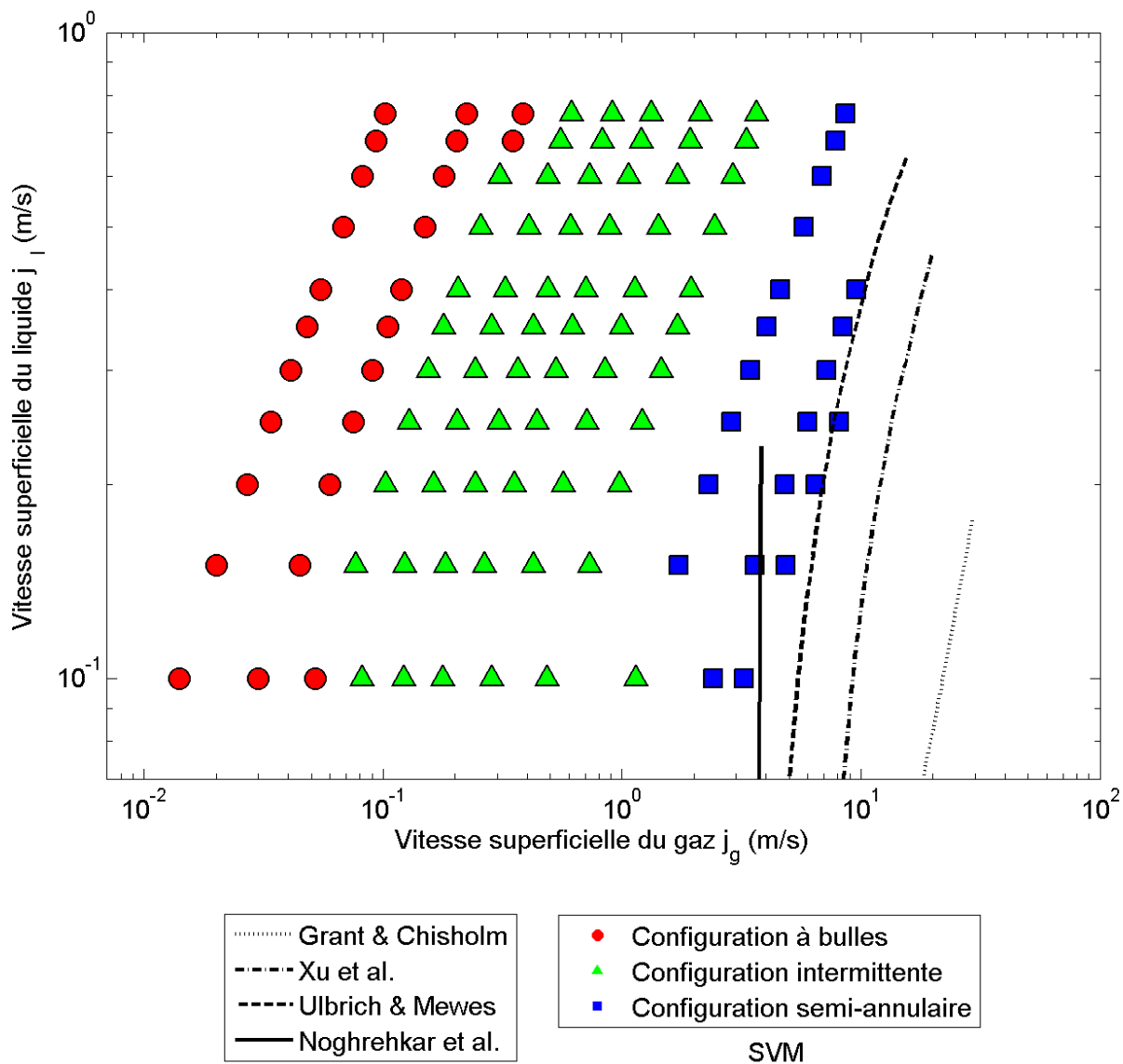


Figure 7.6 : Comparaison des frontières proposées pour la configuration annulaire

7.2.2 Comparaison avec la carte de Taitel et al.

La Figure 7.7 donne la superposition entre la classification prédite par la SVM et les frontières proposées par Taitel et al. (1980), pour un diamètre équivalent à la largeur entre les parois sur lesquelles sont fixés les demi-tubes (4.92 cm, se reporter à l'annexe A).

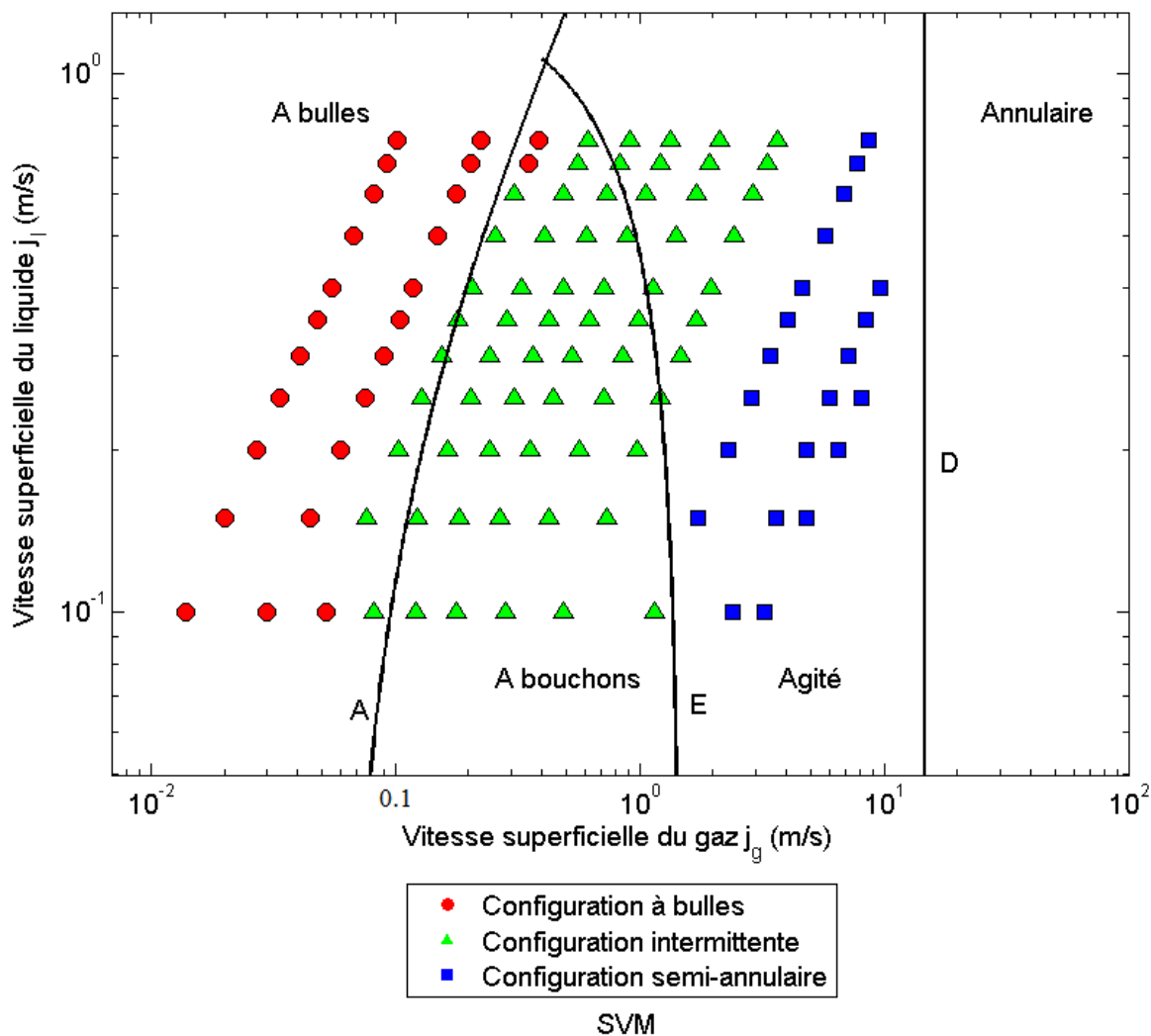


Figure 7.7 : Comparaison avec la carte de Taitel et al. (frontières A, E et D)

A cause de la présence des demi-tubes, on ne s'attend pas à obtenir des bouchons aussi bien définis que dans un tube, mais on observe à la place des bulles plus grosses qui finissent par constituer des poches gazeuses qui occupent quasiment tout l'espace entre les demi-tubes, et de longueur plusieurs fois le diamètre des demi-tubes. On peut les considérer comme étant les équivalents des bouchons pour les faisceaux de tubes, comme le suggèrent Ueno et al. (1995) dont l'étude était menée dans le même type de section aux demi-tubes. La configuration à bouchons observée dans un tube pourra donc être comparée à la configuration intermittente de cette étude.

Frontière à bulles/intermittente

On observe un bon accord au niveau de la frontière entre les configurations à bulles et intermittente/à bouchons, bien meilleur que précédemment en faisceaux de tubes. Il faut toutefois noter que les bulles ont des tailles et formes plus variées dans un faisceau de tubes ou une section qui s'en rapproche que pour un écoulement à l'intérieur d'un tube. Dans leur étude de l'ébullition en faisceau de tubes, van Rooyen et al. (2012) notent plusieurs différences observées entre les tubes et les faisceaux de tubes, et font remarquer notamment que, dans les faisceaux, les bulles sont soumises à des compressions et expansions cycliques selon leur position par rapport aux tubes et l'espace plus ou moins réduit dans lequel elles évoluent.

Frontière intermittente/semi-annulaire

La majorité des points attribués à la configuration intermittente par la SVM le sont aussi par la frontière définie par Taitel et al., mais la forme globale de la frontière ne semble pas correspondre. Comme précédemment, c'est pour les basses vitesses de liquide, mais pour une plage plus grande, ($j_l < 0.25$ m/s) qu'on a un bon accord.

Ici encore, les titres volumiques étudiés ne sont pas assez élevés pour permettre une comparaison des frontières entre les configurations semi-annulaire et annulaire.

7.2.3 Synthèse des comparaisons

On peut conclure suite à cette comparaison que, comme on pouvait s'y attendre, on n'obtient un parfait accord ni avec les cartes données pour les faisceaux de tubes ni avec la carte de Taitel et al. pour un écoulement interne.

La comparaison avec la carte de Taitel et al. se révèle intéressante pour la frontière à bulles/intermittente (frontière A chez Taitel et al.) et pour les basses vitesses superficielles de liquide : $j_l < 0.25$ m/s pour la frontière intermittente/semi-annulaire (frontière E chez Taitel et al.). En revanche, au-delà de cette vitesse, la partie supérieure de la frontière E n'est pas valable, et pour les vitesses superficielles de liquide plus élevées, la courbure des frontières proposées par Noghrehkar et al. est plus appropriée (bien que leur emplacement ne corresponde pas). De façon générale, les cartes pour les faisceaux ont une orientation globale des frontières plus plausible.

Au-delà des différences inévitables dues aux montages expérimentaux eux-mêmes, la section d'essai étudiée présente en effet des caractéristiques communes à chaque type.

Le fait de n'avoir qu'un seul chemin d'écoulement peut expliquer une certaine ressemblance avec l'écoulement à l'intérieur d'un tube. Dans un faisceau de tubes, des interactions sont possibles entre les bulles circulant au départ dans deux chemins d'écoulement parallèles, dans la zone située à l'arrière des tubes, ce qui n'est pas le cas dans la présente section d'essai.

En revanche, la présence des demi-tubes sur les parois impacte fortement la forme de l'écoulement et la nature physique des configurations observées, qui se rapprochent beaucoup plus de ce qu'on observe dans les faisceaux que des configurations décrites à l'intérieur d'un tube.

Cette section d'essai pourrait ainsi servir de référence lors de l'identification des configurations dans un faisceau de tubes, en permettant de séparer les effets dus à la présence de chemins d'écoulement multiples et leurs interactions, et ceux dus à la forme du chemin d'écoulement à cause de la disposition des tubes, puisqu'on observe déjà cet effet ici avec les demi-tubes.

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

Ce projet avait pour objet l'étude des configurations d'un écoulement diphasique air-eau dans une section d'essai se rapprochant d'un faisceau de tubes. L'objectif principal était d'obtenir une identification fiable des configurations d'écoulement et plus particulièrement de leurs frontières. Étant donné le potentiel des machines d'apprentissage supervisé démontré dans plusieurs études concernant l'identification des configurations pour les écoulements internes, le moyen permettant d'atteindre cet objectif était d'utiliser une machine à vecteurs de support (SVM) pour classer les données provenant de mesures par sonde optique, pour un éventail assez large de vitesses de gaz et de liquide.

La carte d'écoulement obtenue à partir de la classification finale prédite par la SVM est jugée satisfaisante, et la comparaison avec la littérature confirme des ressemblances dans les frontières à la fois avec l'écoulement à l'intérieur d'un tube et avec ceux qu'on observe dans les faisceaux de tubes. Cela correspond bien au fait qu'on a à la fois une conduite présentant un seul chemin d'écoulement (d'où les similitudes avec l'écoulement interne), mais comportant des demi-tubes sur les parois (d'où les similitudes avec les faisceaux).

La méthode d'identification proposée est transférable aux faisceaux de tubes, et constitue un progrès du point de vue de l'objectivité de la démarche. L'observation visuelle depuis l'extérieur du faisceau est toujours nécessaire pour décrire et nommer les configurations observées, ainsi que pour obtenir des points de référence pour entraîner la SVM. Cependant, la classification finale repose sur des mesures objectives obtenues au centre de la section d'essai ce qui évite les problèmes liés aux parois relevés par Noghrehkar et al. (1999). De plus, le calcul des moments statistiques est directement possible à partir du signal de la sonde, et ils synthétisent de façon très condensée les caractéristiques de la distribution du taux de vide, qui est elle-même porteuse de l'information reliée à la configuration de l'écoulement. En particulier, l'asymétrie et le kurtosis du signal de la sonde optique semblent suffire comme données d'entrées de la SVM. Lorsque les paramètres internes de la SVM sont bien optimisés, celle-ci permet alors d'obtenir une classification plus objective, particulièrement au niveau des transitions entre les configurations.

Recommandations

1. Approfondissement de l'étude des configurations dans la section d'essai aux demi-tubes

- a. La carte d'écoulement obtenue est cohérente avec la littérature, même si on observe d'inévitables divergences, notamment liées aux conditions expérimentales. Elle donne une classification plausible pour cette section d'essai, bien plus objective que si on avait classé manuellement tous les points, ce qui serait difficile et assez subjectif pour les zones de transition. Cependant, elle dépend assez fortement des points de référence choisis au départ, sur lesquels toute la construction de la SVM est basée.

Il serait donc intéressant d'étudier la classification prédite par une SVM basée sur d'autres points, mais il faudrait pour cela obtenir davantage de points de mesure susceptibles de servir de points de référence (donc loin des zones de transition identifiées ici), surtout pour des titres volumiques faibles (pour avoir plus de points en configuration à bulles) et élevés (idem pour la configuration semi-annulaire).

- b. D'autre part, pour les points de mesure réalisés, avec $\beta \leq 97\%$, on n'a pas observé la configuration annulaire. Il serait intéressant d'essayer d'augmenter encore le titre volumique pour voir si on peut observer cette configuration, et mettre à jour la SVM pour chercher la frontière entre les configurations semi-annulaire et annulaire.
- c. Cela amène à un troisième point : les frontières qui ont été définies dépendent aussi de la précision du « maillage » des points de mesure, et de leur orientation sous forme de lignes sur lesquelles on augmente la vitesse de liquide pour un titre volumique donné. Afin d'obtenir des frontières plus précises, on pourrait réaliser un plus grand nombre de mesures au niveau des zones de transition, disposées de façon plus aléatoire, et là encore, se servir de la SVM pour classer ces points de façon objective.
- d. Concernant la SVM, une autre façon de choisir les paramètres C et γ optimaux pour la SVM à l'issue de la *cross validation* ou d'effectuer la multi-classification (*pairwise classification* au lieu de la classification *One versus Rest*), pourrait être étudiée. Il existe également des méthodes d'optimisation du choix des composantes des vecteurs comme les algorithmes génétiques, qui peuvent améliorer encore les résultats obtenus.

- e. Enfin, on pourrait aussi comparer les résultats obtenus à partir des mêmes données mais avec d'autres types d'apprentissage supervisé, comme les réseaux de neurones, utilisés pour un écoulement interne par Rosa et al. (2010).

2. *Adaptation de cette méthode d'identification, basée sur l'apprentissage supervisé, aux faisceaux de tubes*

La méthode d'identification des configurations d'écoulement diphasique obtenue ici, par mesure du signal donné par une sonde optique, calcul des moments statistiques de ce signal, et utilisation d'une machine d'apprentissage supervisé, ici une machine à vecteurs de support, devrait être transposable au cas du faisceau de tubes, sans difficultés du point de vue de la programmation, car les différents paramètres peuvent facilement être ajustés à de nouvelles données ou à un nombre différent de configurations.

Le plus compliqué sera sans doute la définition des configurations de points de référence car les observations visuelles sont plus compliquées à analyser dans ce cas. D'autre part, ce montage plus complexe nécessitera probablement des mesures par sonde optique en différents points, situés dans le chemin d'écoulement ou à d'autres endroits stratégiques par rapport à ce qu'on observe visuellement. On pourra s'appuyer pour cela sur les observations détaillées et vidéos réalisées au sein de la Chaire par Perrot (2011), dans un faisceau triangulaire normal.

BIBLIOGRAPHIE

- Bell, K. J., Taborek, J., & Fenoglio, F. (1970). Interpretation of horizontal in-tube condensation heat transfer correlations with a two-phase flow regime map. In *Chem. Eng. Progr., Symp. Ser. 66: No. 102* (pp. 150–165). Heat Transfer Research, Inc., Alhambra, Calif.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), 27.
- Chung, H. J., & Chu, I. (2006). Fluid-Elastic Instability of Rotated Square Tube Array in Air-Water Two-Phase Cross-Flow. *Nuclear Engineering and Technology*, 38(1), 69.
- Coquand, N. (2005). *Design d'une nouvelle boucle d'essai diphasique pour le laboratoire*. Rapport de stage, École Polytechnique de Montréal.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- Costigan, G., & Whalley, P. B. (1997). Slug flow regime identification from dynamic void fraction measurements in vertical air-water flows. *International Journal of Multiphase Flow*, 23(2), 263–282.
- Faure, B. (2009). *A Damping Model for Two-phase Cross-flow : Experimental Study*. Rapport de stage, École Polytechnique de Montréal.
- Girardeau, M. (2011). *Forces fluctuantes générées par un écoulement diphasique interne : mécanisme et modélisation*. Mémoire de maîtrise, École Polytechnique de Montréal.
- Girardeau, M., Pettigrew, M. J., & Mureithi, N. W. (2011). Two-Phase Flow Excitation Forces on a Vertical U-Bend Tube. In *ASME 2011 Pressure Vessels and Piping Conference* (pp. 103–111). American Society of Mechanical Engineers.
- Grant, I. D. R., & Chisholm, D. (1979). Two-phase flow on the shell-side of a segmentally baffled shell-and-tube heat exchanger. *Journal of Heat Transfer*, 101(1), 38–42.

- Grant, I. D. R., & Murray, I. (1972). *Pressure drop on the shell-side of a segmentally baffled shell-and-tube heat exchanger with vertical two-phase flow*. Report NEL-500, National Engineering Laboratory.
- Hong, W. P., & Liu, Y. (2010). Investigation of gas-liquid two-phase flow patterns based on high-speed imaging methods. In *Proceedings of 2010 3rd International Congress on Image and Signal Processing, CISP 2010* (Vol. 1, pp. 154–156). IEEE.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification.
- Jayanti, S., & Hewitt, G. F. (1992). Prediction of the slug-to-churn flow transition in vertical two-phase flow. *International Journal of Multiphase Flow*, 18(6), 847–860.
- Jayanti, S., Hewitt, G. F., Low, D. E. F., & Hervieu, E. (1993). Observation of flooding in the Taylor bubble of co-current upwards slug flow. *International Journal of Multiphase Flow*, 19(3), 531–534.
- Jones, O. C., & Zuber, N. (1975). The interrelation between void fraction fluctuations and flow patterns in two-phase flow. *International Journal of Multiphase Flow*, 2(3), 273–306.
- Kaichiro, M., & Ishii, M. (1984). Flow regime transition criteria for upward two-phase flow in vertical tubes. *International Journal of Heat and Mass Transfer*, 27(5), 723–737.
- Kampa, K. (2013). libsvm for MATLAB - Guide by Kittipat Kampa (Integrated Brain Imaging Center, UW Medical Center, Seattle). Retrieved from https://sites.google.com/site/kittipat/libsvm_matlab
- Komgom Nguempjop, C. (2006). *Two-phase flow identification using Support Vector Machine*. Rapport de projet, Ecole Polytechnique de Montréal.
- Kondo, M., & Nakajima, K. (1980). Experimental investigation of air-water two phase upflow across horizontal tube bundles: Part 1, flow pattern and void fraction. *Bulletin of JSME*, 23(177), 385–393.

- Mureithi, N., & Masabarakiza, C. (2011). Determination of flow structure in a tube array by Particle Image Velocimetry. In *Proceedings of 2011 ASME Pressure Vessels & Piping Conference* (pp. 1–12 (PVP2011–57918)). Baltimore, Maryland, U.S.A.
- Ng, A. (Stanford U. (2015). Machine learning - online course by Coursera. Retrieved from <https://www.coursera.org/learn/machine-learning/>
- Noghrehkar, G. R., Kawaji, M., & Chan, a. M. C. (1999). Investigation of two-phase flow regimes in tube bundles under cross-flow conditions. *International Journal of Multiphase Flow*, 25(5), 857–874.
- Perrot, E. (2011). *Forces d'excitation vibratoire dans un faisceau de tubes en configuration triangulaire normale soumis à un écoulement diphasique*. Mémoire de maîtrise, École Polytechnique de Montréal.
- Pettigrew, M. J., Taylor, C. E., Fisher, N. J., Yetisir, M., & Smith, B. A. W. (1998). Flow-induced vibration: recent findings and open questions. *Nuclear Engineering and Design*, 185(2), 249–276.
- Ribatski, G., & Thome, J. R. (2007). Two-Phase Flow and Heat Transfer across Horizontal Tube Bundles- A Review. *Heat Transfer Engineering*, 28(6), 508–524.
- Rosa, E. S., Salgado, R. M., Ohishi, T., & Mastelari, N. (2010). Performance comparison of artificial neural networks and expert systems applied to flow pattern identification in vertical ascendant gas–liquid flows. *International Journal of Multiphase Flow*, 36(9), 738–754.
- Taitel, Y., Bornea, D., & Dukler, A. E. (1980). Modelling flow pattern transitions for steady upward gas- liquid flow in vertical tubes. *AIChE Journal*, 26(3), 345–354.
- Tan, C., Dong, F., & Wu, M. (2007). Identification of gas/liquid two-phase flow regime through ERT-based measurement and feature extraction. *Flow Measurement and Instrumentation*, 18(5-6), 255–261.

- Ueno, T., Leung, W.H., Ishii, M. (1995). Local measurement in two-phase flow across a horizontal tube bundle. In *Proceedings of the 2nd International Conference on Multiphase Flow, Kyoto, Japan, 3-7 April* (pp. 89–95).
- Ulbrich, R., & Mewes, D. (1994). Vertical, upward gas-liquid two-phase flow across a tube bundle. *International Journal of Multiphase Flow*, 20(2), 249–272.
- Van Rooyen, E., Agostini, F., Borhani, N., & Thome, J. R. (2012). Boiling on a tube bundle: part I—flow visualization and onset of dryout. *Heat Transfer Engineering*, 33(11), 913–929.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.
- Wallis, G. B. (1969). *One-dimensional two-phase flow* (Vol. 1). McGraw-Hill New York.
- Xu, G. P., Tso, C. P., & Tou, K. W. (1998). Hydrodynamics of two-phase flow in vertical up and down-flow across a horizontal tube bundle. *International Journal of Multiphase Flow*, 24, 1317–1342.

ANNEXE A – DÉTAILS DU MONTAGE EXPÉRIMENTAL

Dimensions de la section d'essai

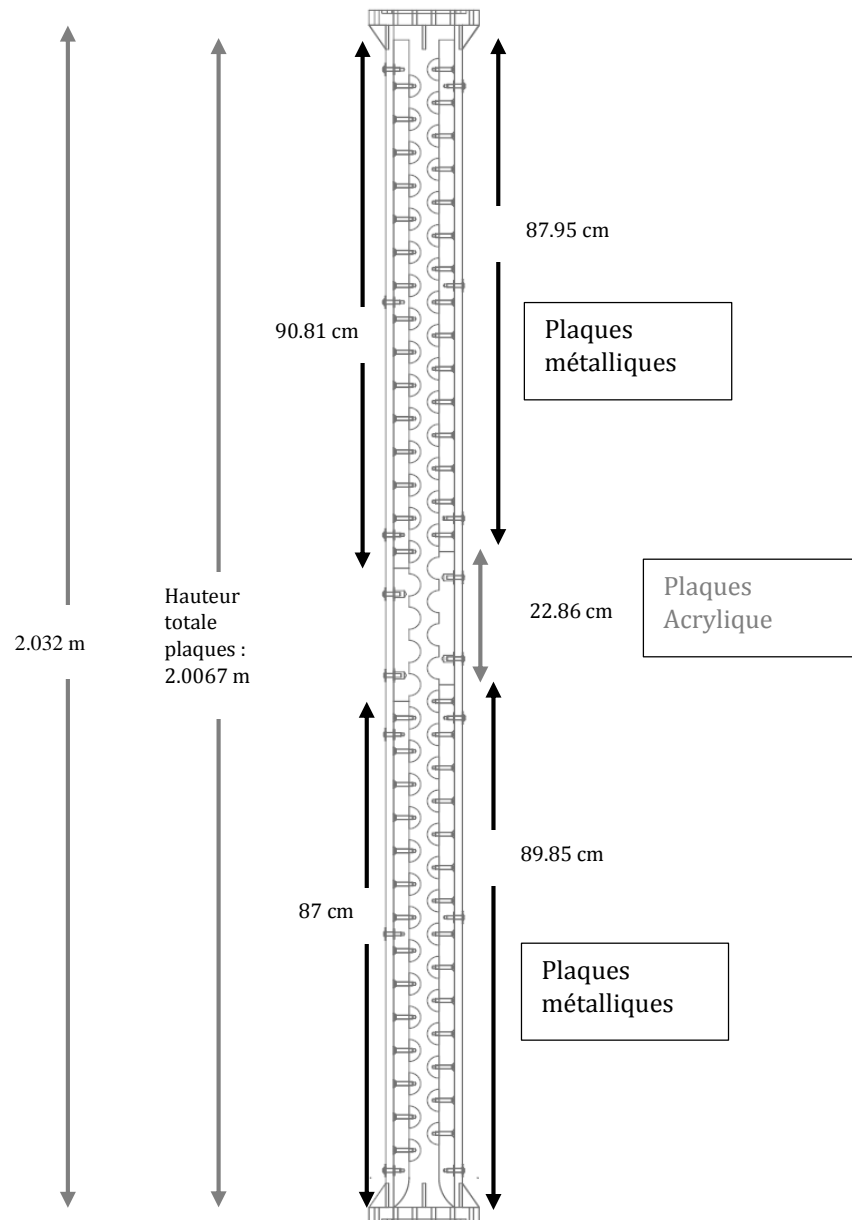


Figure A. 1 : Dimensions générales de la section d'essai

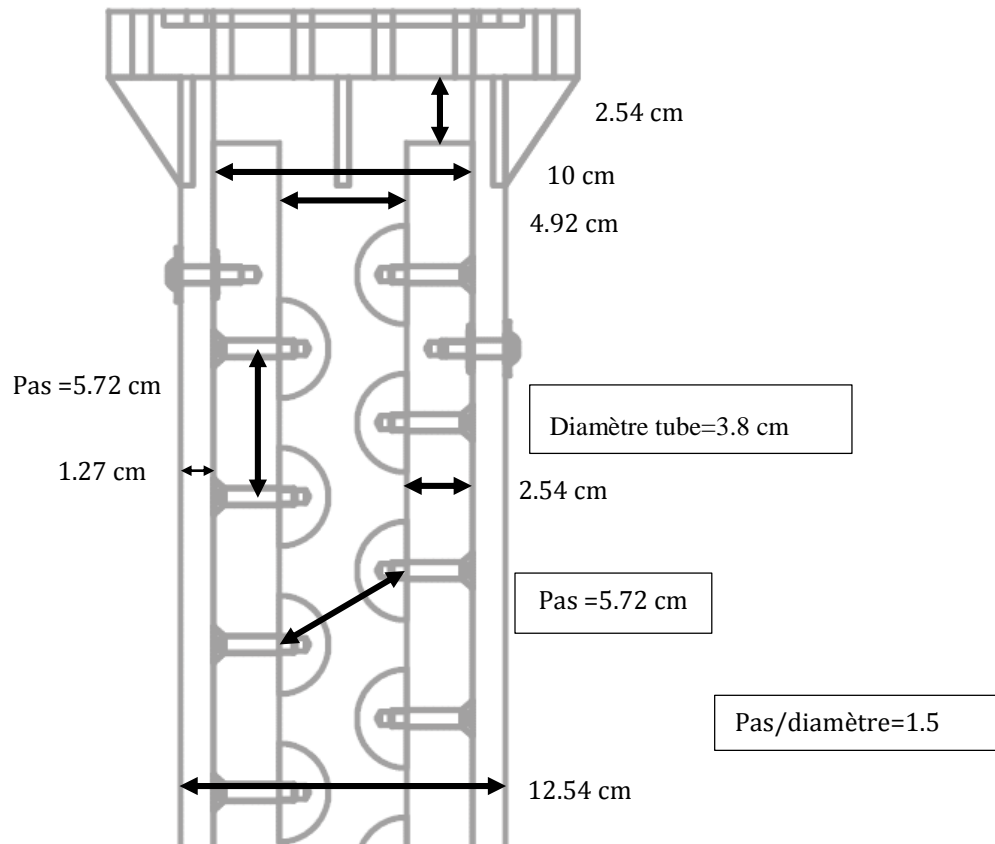


Figure A. 2 : Détails des dimensions des plaques et tubes métalliques

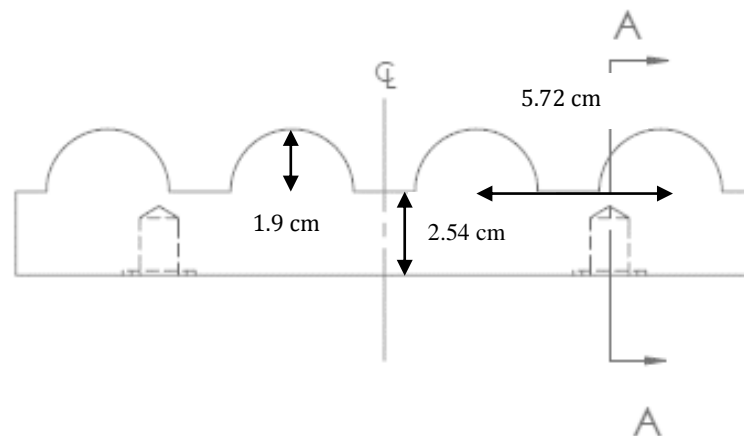


Figure A. 3 : Dimensions des plaques acryliques

Débitmètres et pompes

Les Tableau A. 1 et Tableau A. 2 regroupent les caractéristiques des composants des circuits d'air et d'eau. Les débits d'air indiqués ici correspondent aux conditions standard de pression et de température : $T_0 = 20^\circ C$ et $P_0 = 14.7 \text{ psi} = 101353 \text{ Pa}$, dans lesquelles on a : $1 \text{ SCFM} \simeq 1 \text{ CFM}$. La température de la salle est maintenue autour de T_0 .

Le débitmètre Aalborg a été utilisé pour mesurer le débit d'air, sauf pour les débits supérieurs à $0.01583 \text{ m}^3/\text{s}$ (950 LPM), pour lesquels on a utilisé la plaque à orifice Rosemount.

Tableau A. 1 : Caractéristiques des composants des circuits d'air et d'eau

| Composant | Fluide | Marque | Modèle | Plage d'opération | Précision et résolution |
|------------------------|--------|---|--------------|--|---|
| Circuit d'air comprimé | Air | Fourni par la centrale thermique de l'École Polytechnique | | Débit d'air constant jusqu'à $300 \text{ CFM} = 0.1416 \text{ m}^3/\text{s}$ | - |
| Débitmètre massique | Air | Aalborg | GFC 77 | $0 - 1000 \text{ LPM}$ $0 - 0.0167 \text{ m}^3/\text{s}$ | $\pm 1.5\%$ de la valeur maximale : $\pm 15 \text{ LPM}$ $\pm 0.00025 \text{ m}^3/\text{s}$ |
| Plaque à orifice | Air | Rosemount | S015T1S0512D | $10 - 50 \text{ CFM}$ $0.00472 - 0.0236 \text{ m}^3/\text{s}$ | $\pm 1.5\%$ de la valeur mesurée (sur la plage d'opération) |
| Débitmètre massique | Eau | Krohne | IFM4010/D/6 | $35.4-1413.7 \text{ LPM}$ $0.0006-0.0236 \text{ m}^3/\text{s}$ | $\pm 0.5\%$ de la valeur mesurée (sur la plage d'opération) |

Tableau A. 2 : Caractéristiques des pompes

| Composant | Fluide | Marque et modèle | Puissance et Nombre de tours par minute (RPM) | Plage d'opération | |
|--------------------|--------|---|---|------------------------------|--|
| Pompe à engrenages | Eau | OBERDORFER PUMPS INC. | A 60 Hz, 2 HP (1.5 kW) | Fréquence (Hz) | Débit mesuré |
| | | | 1750 RPM | 0.3 | 0.00233 L/s 2.3×10^{-6} m^3/s |
| | | | | 60 | 1.27 L/s 1.27×10^{-3} m^3/s |
| Variateur associé | - | SMVector Lenze AC Tech | - | | |
| Pompe de la boucle | Eau | Goulds Pumps, ITT Industries Inc. | A 60 Hz, 2 HP (1.5 kW) | 0-10 L/s $0 - 0.01 m^3/s$ | |
| | | Model SSH 5SH2G6B0 | 1740 RPM | | |
| Variateur associé | - | Modèle Aquavar, Goulds Pumps, ITT Industries Inc. | - | | |

La pompe à engrenages a été utilisée pour tous les points de mesure sauf ceux dont le débit était supérieur à 1.27 L/s, pour lesquels la deuxième pompe a été utilisée.

Injection d'air

La configuration initiale de la boucle d'essai prévoyait une injection du mélange diphasique avant le divergent relié au coude , cf Figure A. 4.

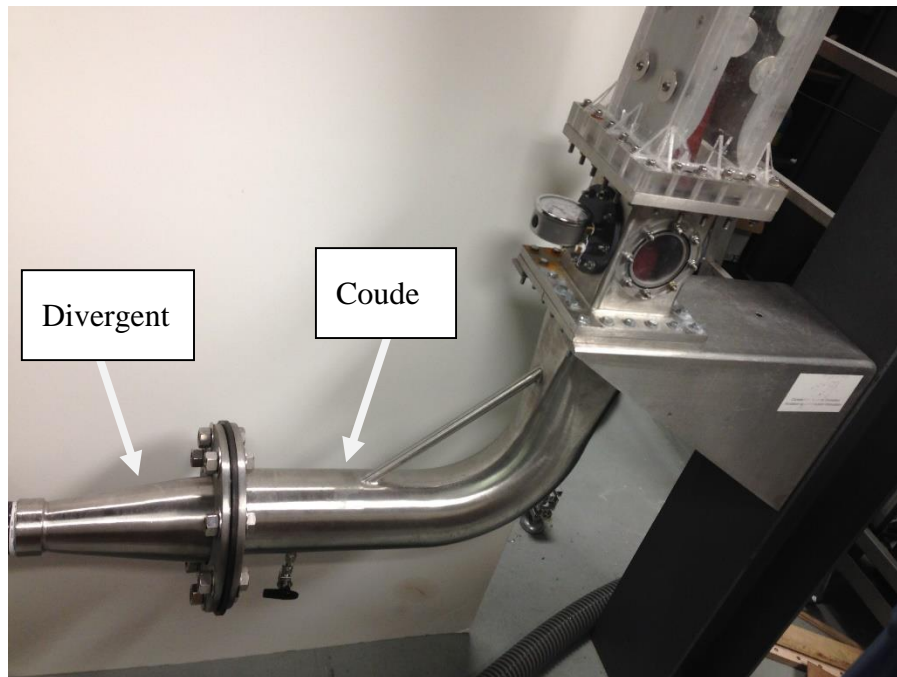


Figure A. 4 : Divergent et coude

Il s'est avéré lors des premiers tests de fonctionnement que ceci créait un phénomène périodique parasite dans la section d'essai : lorsque le débit d'air était trop faible par rapport au débit d'eau, à cause du volume d'eau important au niveau du coude, l'air ne pouvait pas entraîner l'eau vers le haut de façon constante, et on observait une intermittence dans l'injection du mélange.

Le connecteur a donc été modifié pour obtenir le système présenté Figure A. 5.

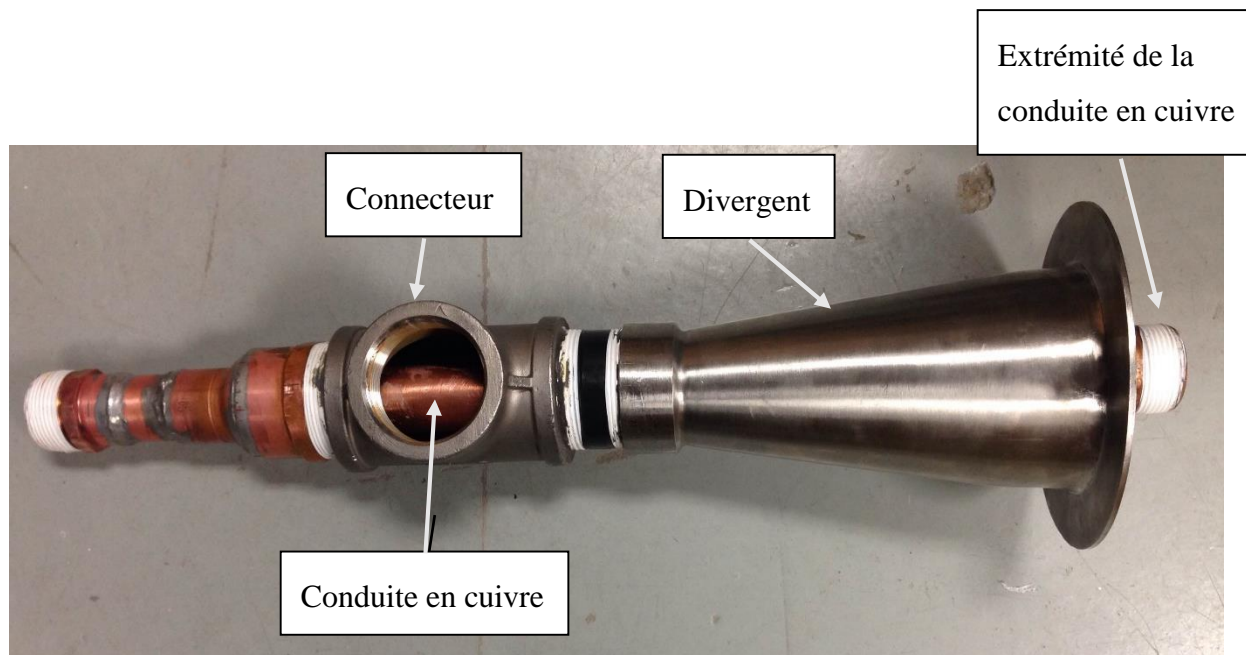


Figure A. 5 : Nouveau connecteur pour l'injection d'air et eau

Une conduite en cuivre a été insérée à l'intérieur du connecteur, et traverse le divergent jusqu'à l'extrémité indiquée par la flèche sur la Figure A. 5, à laquelle un tuyau flexible a été connecté (par le biais d'un Quick Connect résistant à l'eau) afin d'injecter l'air au-dessus du coude, directement dans la section d'essai. Il aurait été préférable de trouver un moyen d'insérer un mélangeur, mais cette solution s'est toutefois avérée satisfaisante. En effet, l'instrument de mesure est placé à plus d'un mètre haut-dessus de ce point d'injection, on a donc supposé que la configuration des tubes était suffisamment dense pour assurer un mélange convenable des deux phases. Cela a été vérifié expérimentalement, par observation directe en fonctionnement et grâce aux images prises par caméra haute vitesse, qui montrent bien un mélange homogène lorsqu'on est dans les conditions adéquates.

La Figure A. 6 montre les arrivées finales d'air et d'eau.

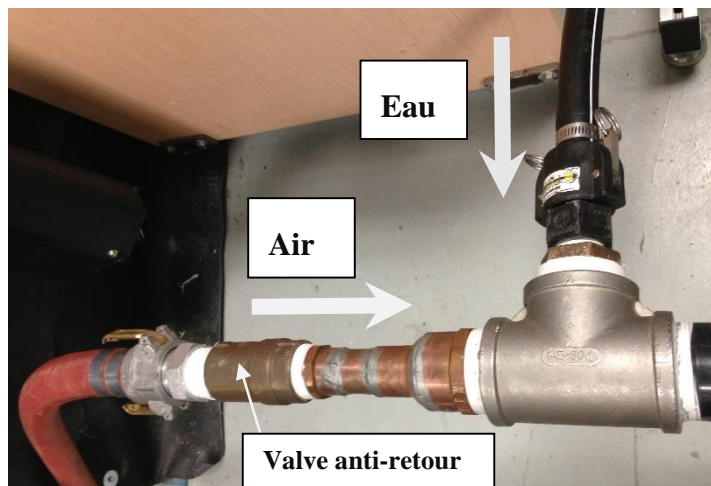


Figure A. 6 : Entrées finales d'air et d'eau

Procédure de réglage des débits

Débit d'eau

En utilisant la pompe à engrenages, dont le principe de fonctionnement est de pousser mécaniquement l'eau à une fréquence donnée, le débit obtenu est constant. Un étalonnage précis de la pompe en situation réelle (c'est-à-dire à l'extrémité des tuyaux utilisés) a permis d'obtenir la relation linéaire entre la fréquence de la pompe et le débit obtenu au point d'injection de l'eau (cf Figure A. 6), en tenant compte directement des pertes de charge dans le circuit. Pour la plupart des points, le débit était donc imposé directement par le biais de la fréquence de la pompe.

Pour les points dont le débit était supérieur aux capacités de cette pompe, la deuxième pompe (Goulds Pumps) a été utilisée, et dans ce cas le débit était donné par le débitmètre Krohne. On se trouvait alors dans la plage de débits pour lesquels l'erreur de mesure est $\pm 0.5\%$ de la valeur mesurée (cf Annexe A).

Débit d'air

Le circuit présenté à la Figure A. 7, qui met en parallèle plusieurs conduites, permet d'utiliser le débitmètre Aalborg pour la plupart des points de mesure, ou de régler la valeur des hauts débits grâce aux vannes de réglage grossier et fin.

La plaque à orifice utilisée pour la mesure des hauts débits est directement installée sur le circuit d'air du laboratoire.

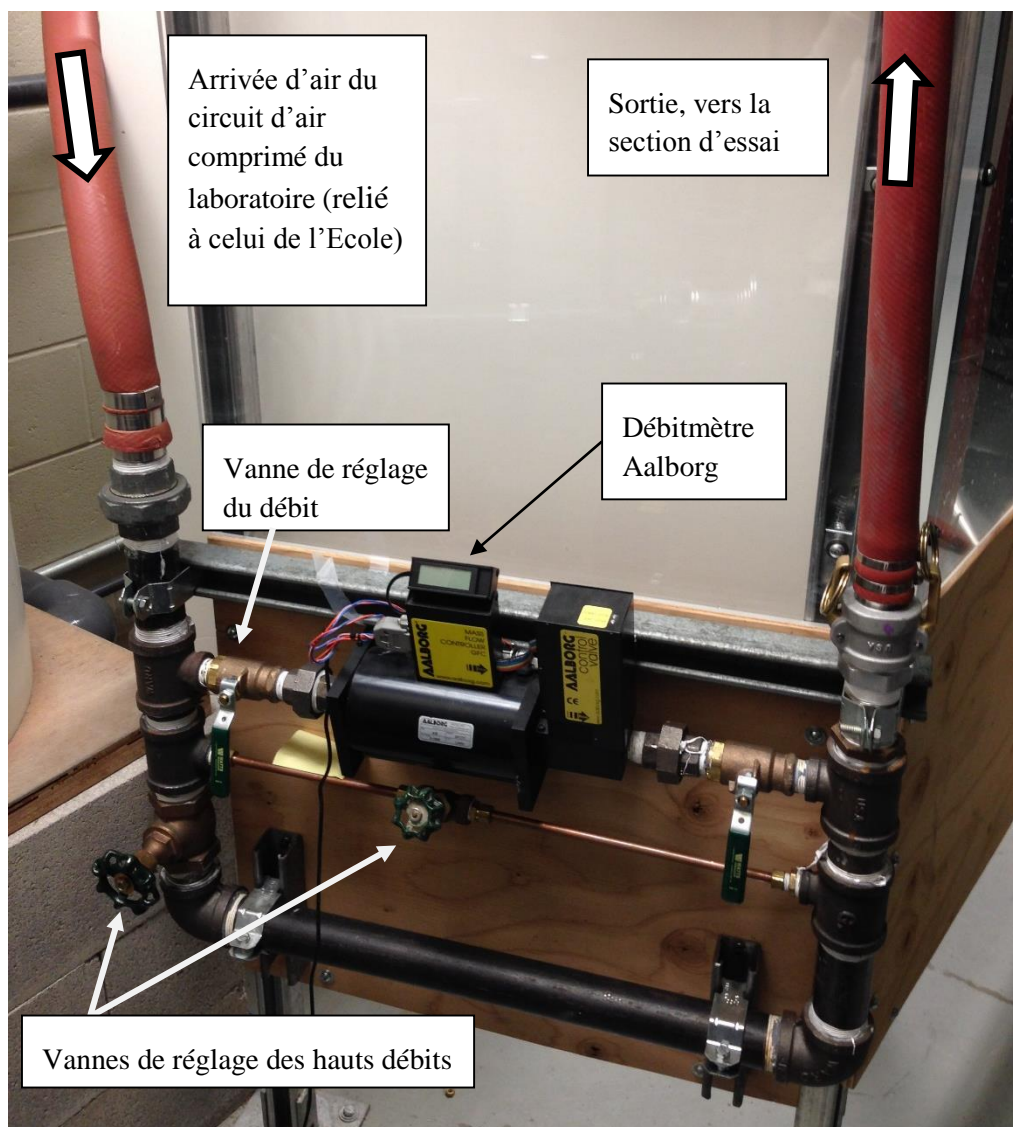


Figure A. 7 : Vannes de réglage du débit d'air et débitmètre Aalborg

Correction du débit d'air

Dans le cas du débit d'air, une correction par itération est nécessaire. En effet, la valeur du débit affichée par le débitmètre utilisé correspond à une mesure réalisée à la pression standard P_o . Mais ce qu'on veut imposer, la vitesse superficielle j_g , correspond au débit d'air au niveau de l'aire de passage de l'écoulement entre les tubes, A , définie par la longueur de passage L_p et la profondeur L_{pr} de la section d'essai :

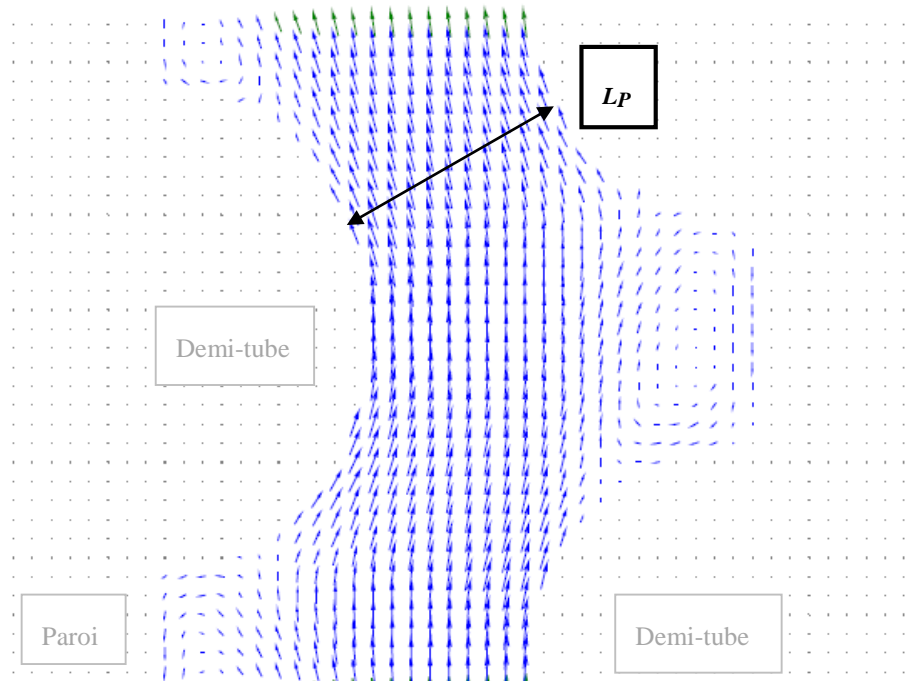


Figure A. 8 : Longueur de passage L_p (rappel de la Figure 3.5)

On a donc : $A = L_p \times L_{pr}$ d'où $A = 0.0019 \text{ m}^2$.

Or à cet endroit, à l'intérieur de la section d'essai, la pression n'est pas égale à la pression standard P_0 . L'air étant un gaz compressible, son volume et donc le débit mesuré dépendent de la pression. D'après la loi des gaz parfaits, à température constante : $P.Q = \text{constante}$.

On effectue donc le réglage du débit d'air par correction itérative grâce à la pression relative, P_{rel} , mesurée au milieu de la section d'essai (juste en-dessous de la sonde optique) :

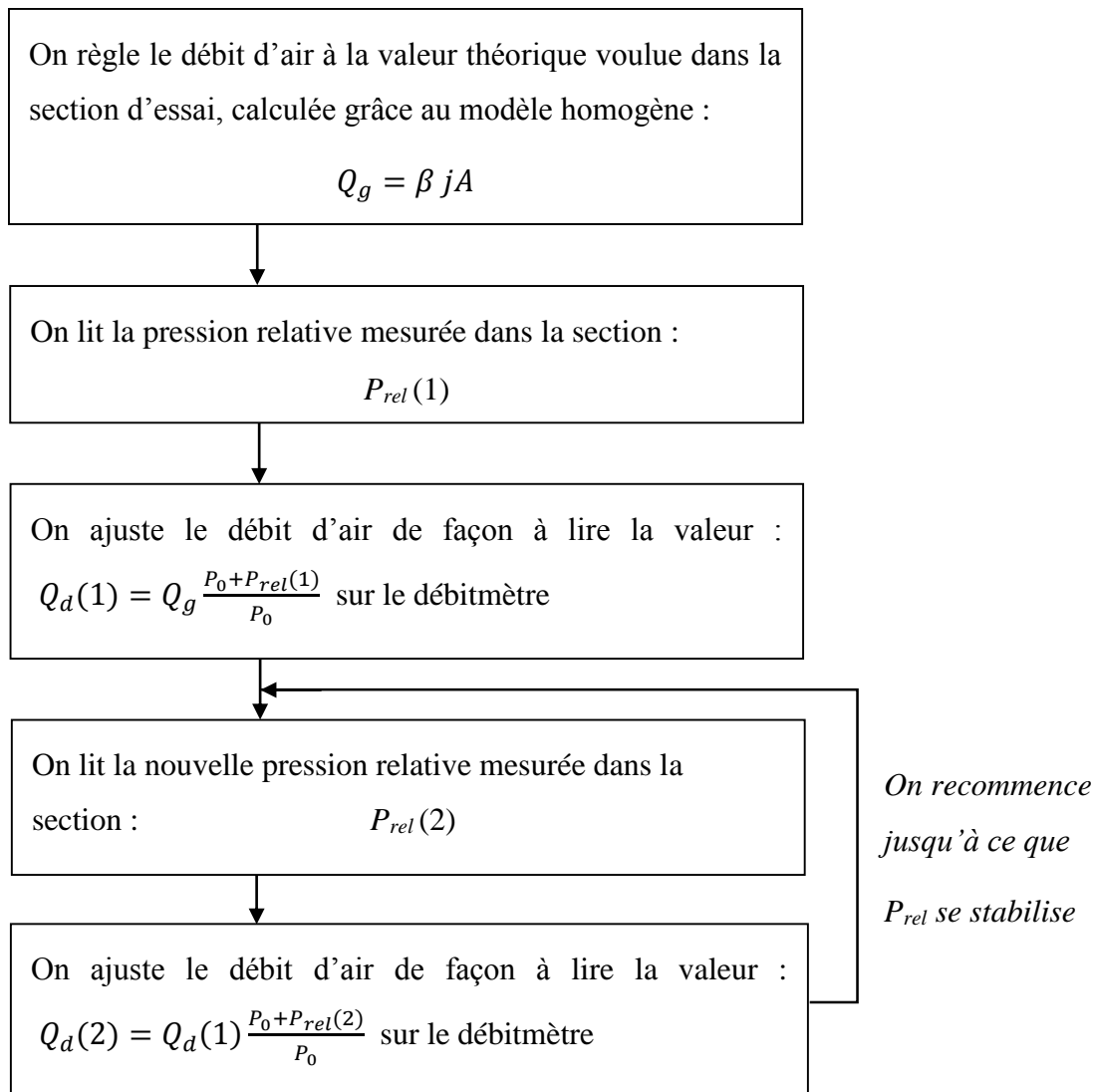


Figure A. 9 : Procédure itérative de correction du débit d'air

Une fois la pression stabilisée, au bout de quelques itérations en général, on peut considérer qu'on a alors le débit d'air voulu Q_g au niveau de la zone de mesure dans la section d'essai.

Sonde optique

Fabrication de la sonde optique

La fibre optique utilisée, multimodes à saut d'indice, présente les caractéristiques suivantes :

Tableau A. 3 : Diamètres de la fibre optique utilisée

| | |
|--|-------------------|
| Diamètre du cœur | 100 μm |
| Diamètre de la gaine | 110 μm |
| Diamètre de l'enveloppe protectrice (gaine extérieure) | 125 μm |

Une attaque chimique par acide fluorhydrique permet tout d'abord d'obtenir une forme conique à l'extrémité de la fibre optique, afin de mieux percer les bulles pour mesurer le changement de phase :

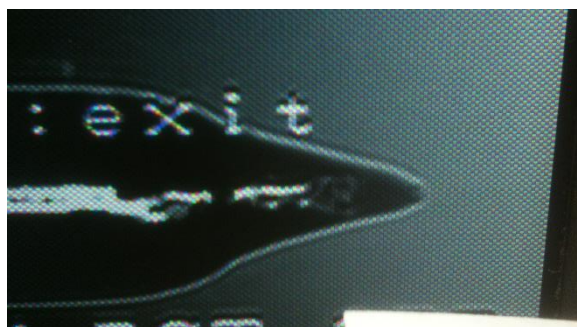


Figure A. 10 : Exemple de cône obtenu à l'extrémité de la sonde

La fibre est ensuite placée et collée dans le support coudé qui permet son insertion au cœur de la section d'essai :



Figure A. 11 : Support de la fibre optique

L'autre extrémité de la fibre est ensuite fusionnée avec un connecteur grâce à la fusionneuse Fujikura (*Arc fusion splicer FSM-30S*). Celui-ci permet de la raccorder au système Ozoptics, comportant un émetteur et amplificateur laser, lui-même relié à l'oscilloscope qui permet de contrôler le signal de sortie, notamment de le centrer :

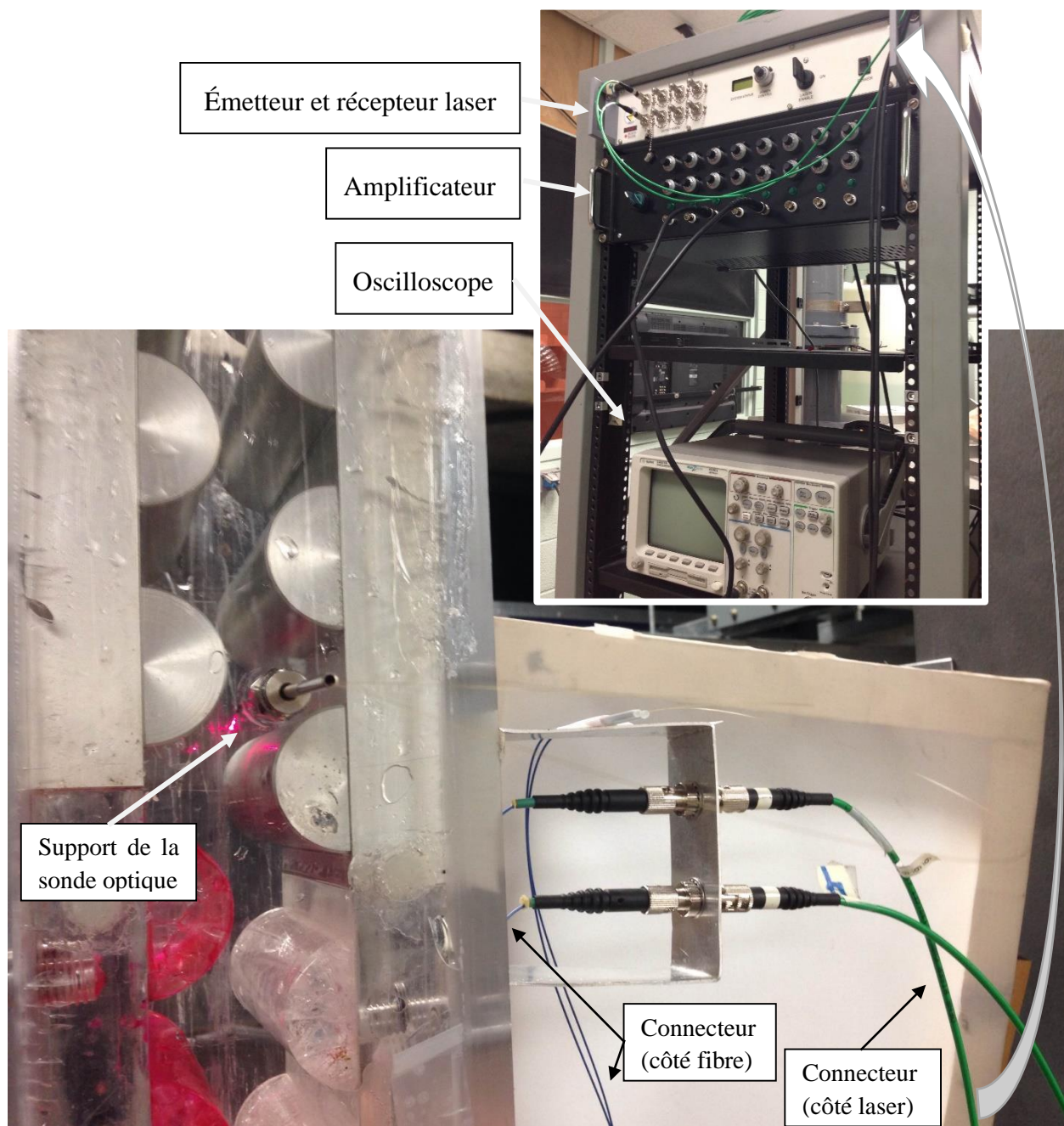


Figure A. 12 : Système d'acquisition du signal de sonde optique

Le signal électrique reçu par l'oscilloscope est ensuite transmis au programme LabVIEW installé sur un ordinateur, par le biais d'un boîtier d'acquisition analogique National Instruments.

La sonde est située à 1.15 m au-dessus de la base de la section d'essai. Dans la direction de la profondeur, l'extrémité de la sonde optique est positionnée au milieu de la section d'essai. Dans le plan perpendiculaire à la section de passage de l'écoulement, comme le montre la Figure A.9, l'extrémité de la sonde est positionnée dans le chemin d'écoulement, rappelé ci-dessous :

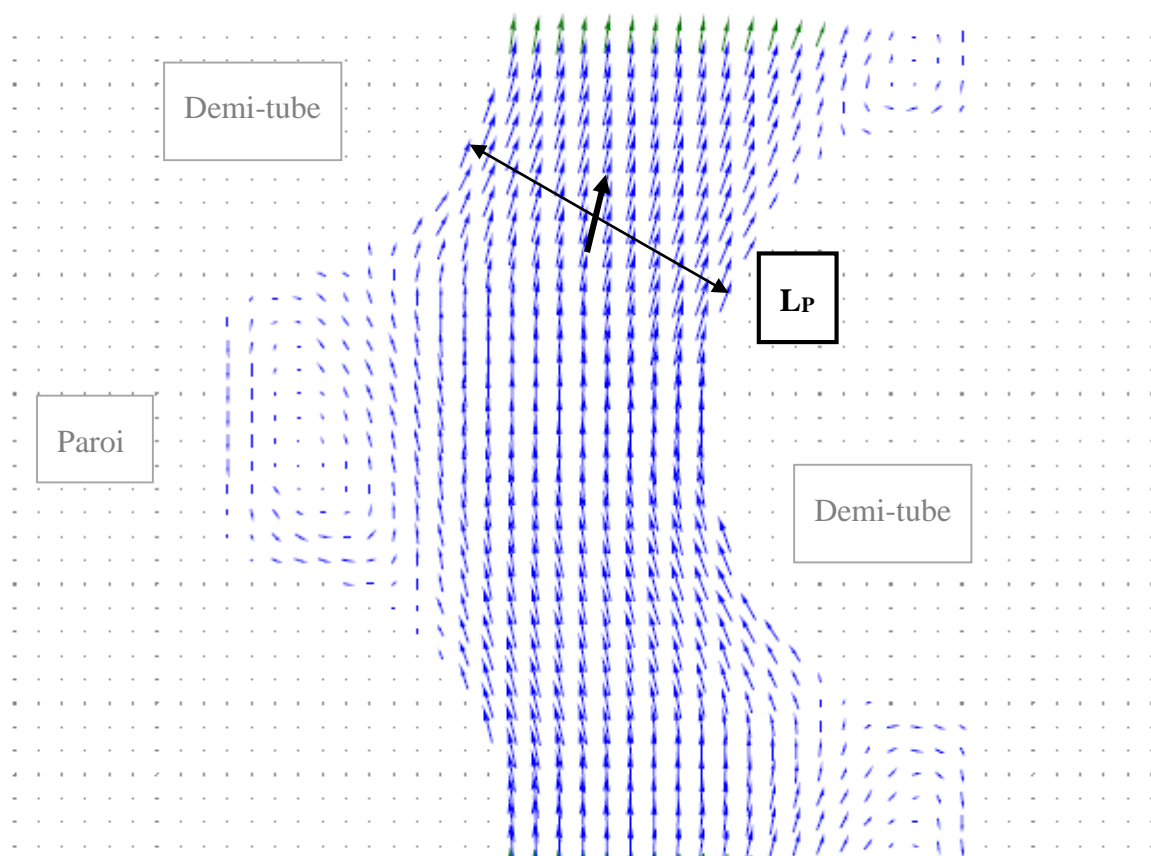


Figure A. 13 : Rappel du champ des vitesses moyennes, pour $jp=0.3$ m/s, $\beta=10\%$., d'après Mureithi & Masabarakiza, 2011

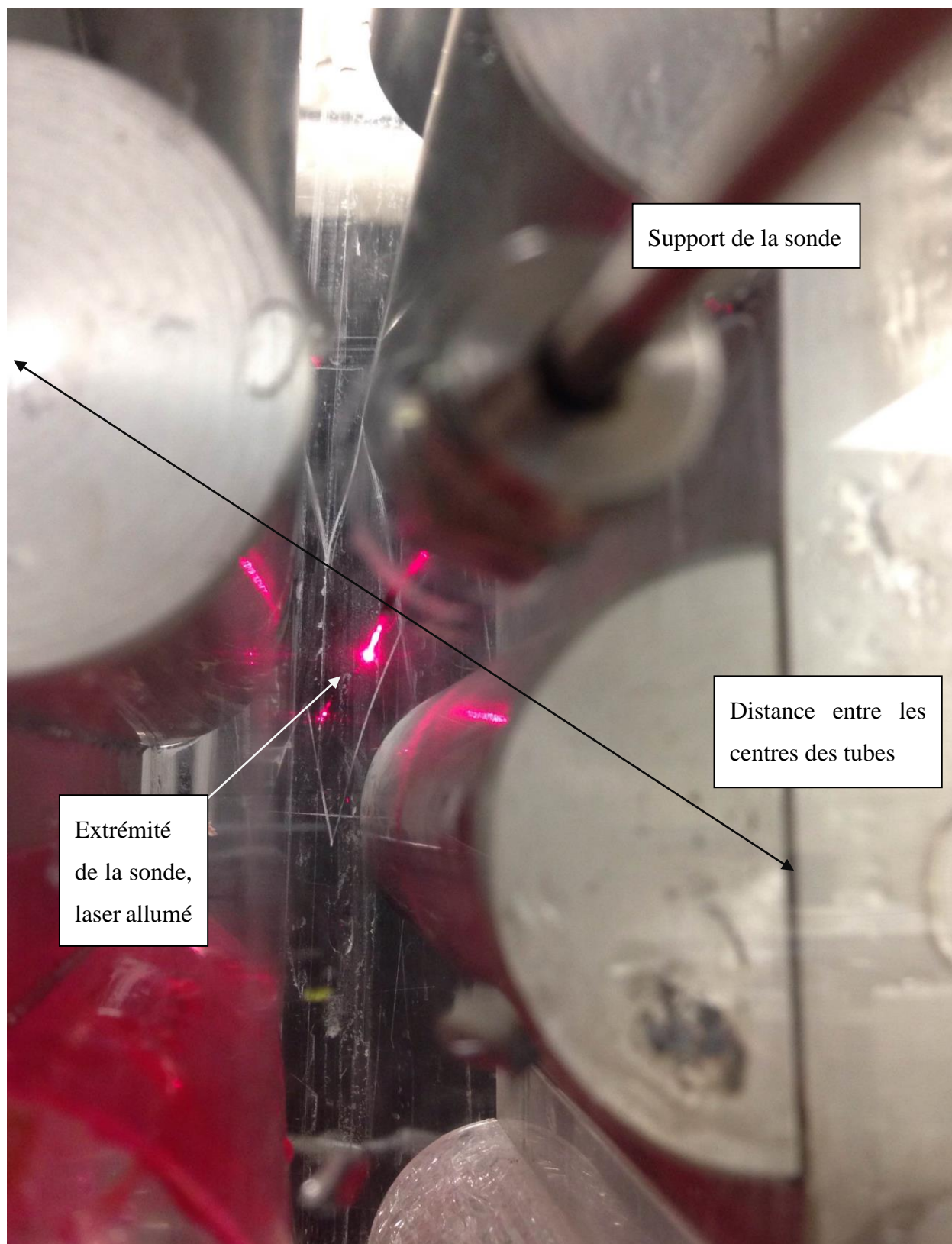


Figure A. 14 : Position de la sonde dans le chemin d'écoulement

Capteur de pression

En plus du trou permettant d'insérer une sonde optique (cf section 0), 4 trous permettent d'insérer des capteurs de pression et sont situés en bas, au milieu, aux trois-quarts et en haut de la section d'essai.

Pour ce projet, la pression relative moyenne a été mesurée au milieu de la section, près de la zone de mesure par sonde optique, pour corriger le débit d'air comme expliqué en sous-section 3.1.5. Le capteur utilisé est le modèle PX409-015GUSBH de la marque Omega. Il permet de mesurer les fluctuations de la pression relative entre 0 et 15 PSIG, avec une précision de $\pm 0.08\%$ BSL, grâce au logiciel USBH Application.



Figure A. 15 : Capteur de pression

ANNEXE B – TRACÉ DES PDF DU TAUX DE VIDE

Tableau B. 1 : PDF pour $j_i=0.1$ m/s

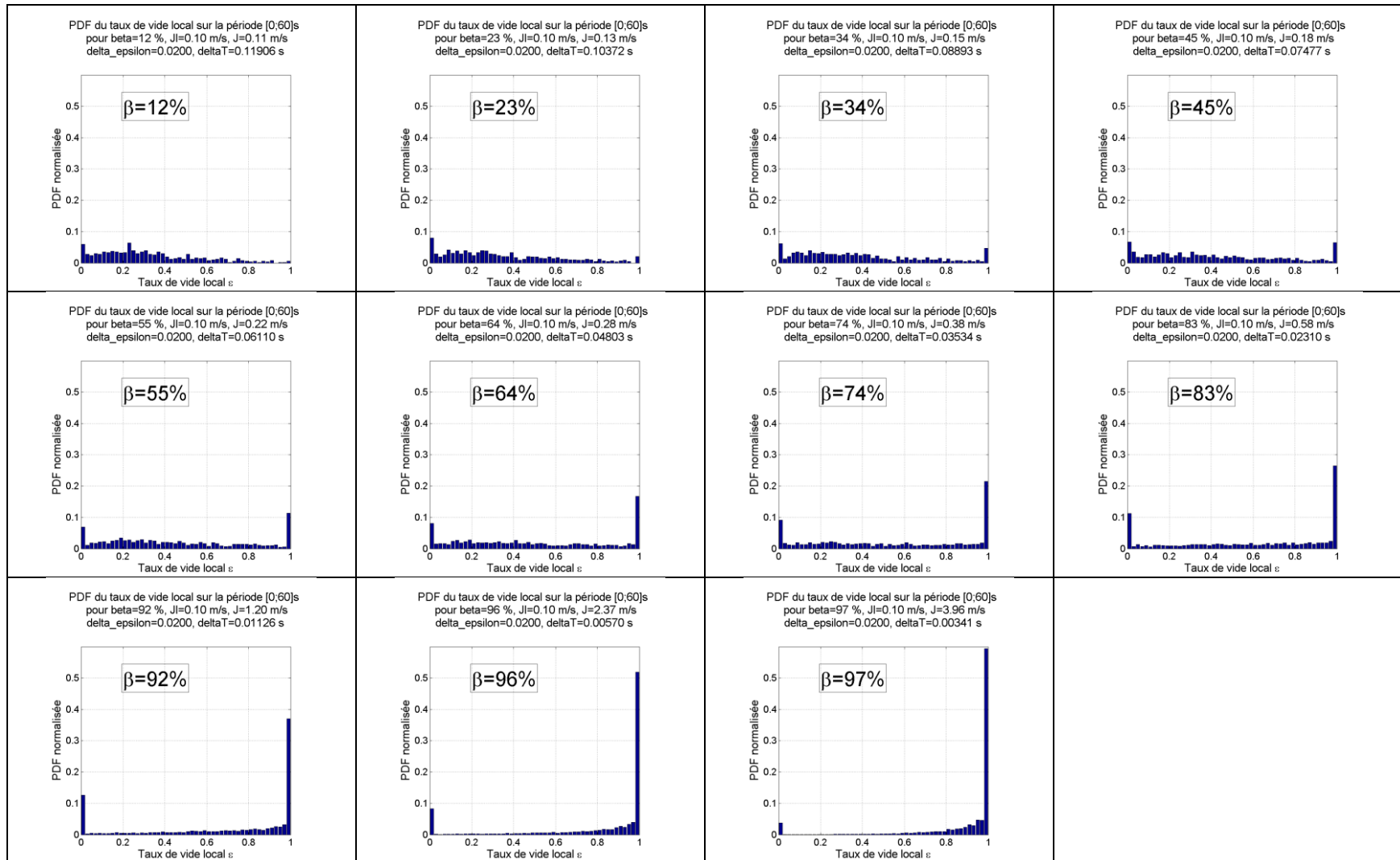


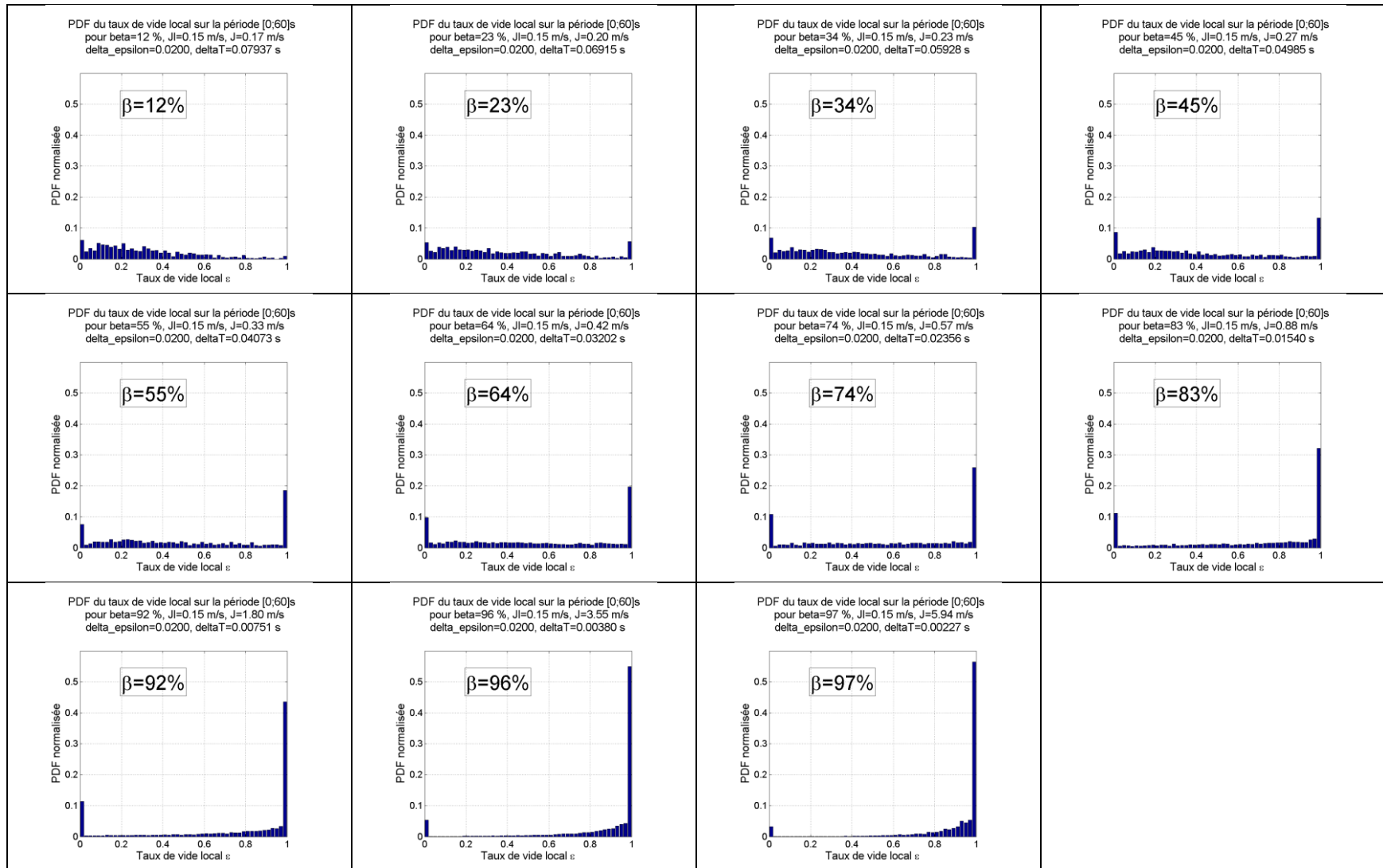
Tableau B. 2 : PDF pour $j_l = 0.15$ m/s

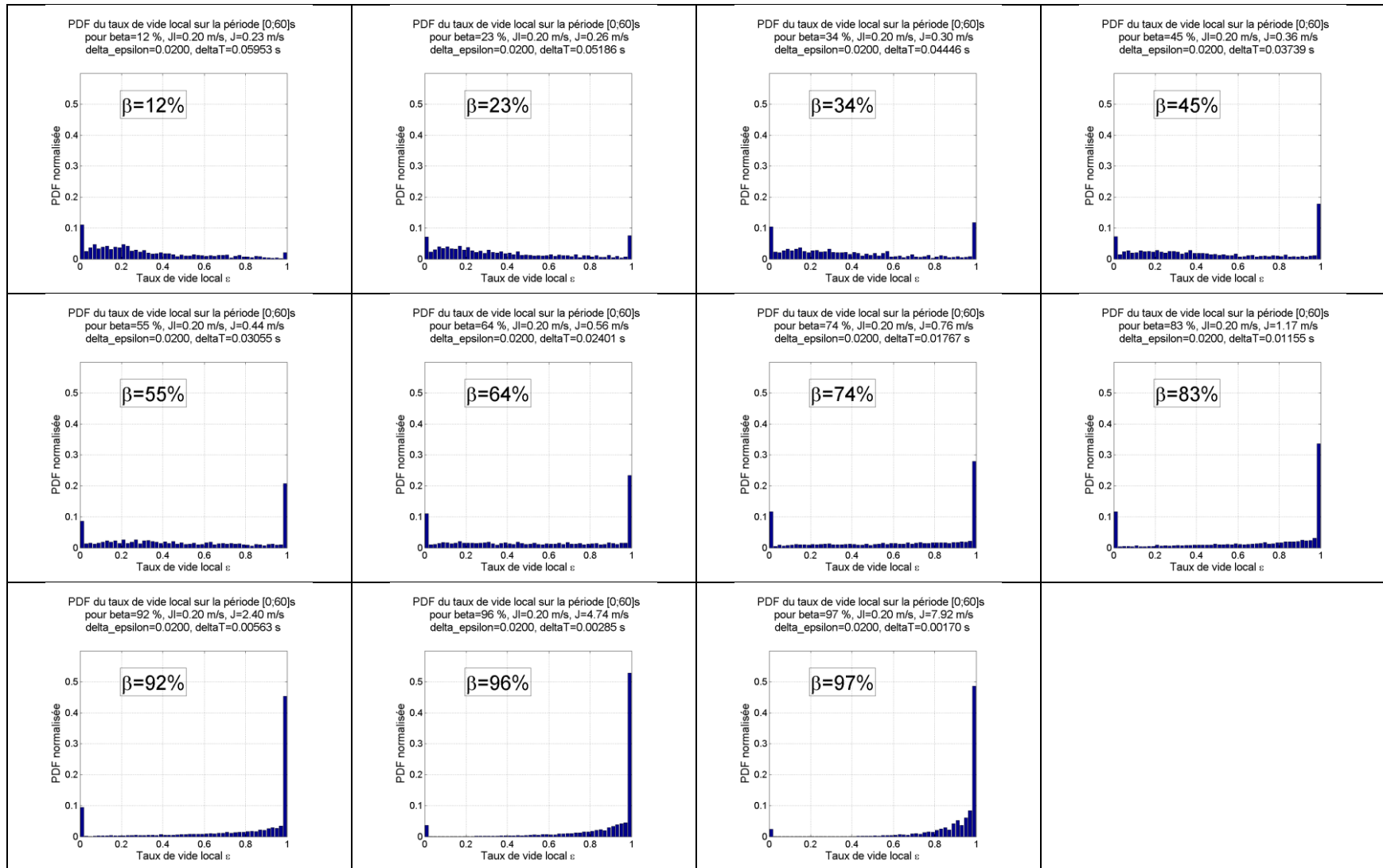
Tableau B. 3 : PDF pour $j_l = 0.2$ m/s

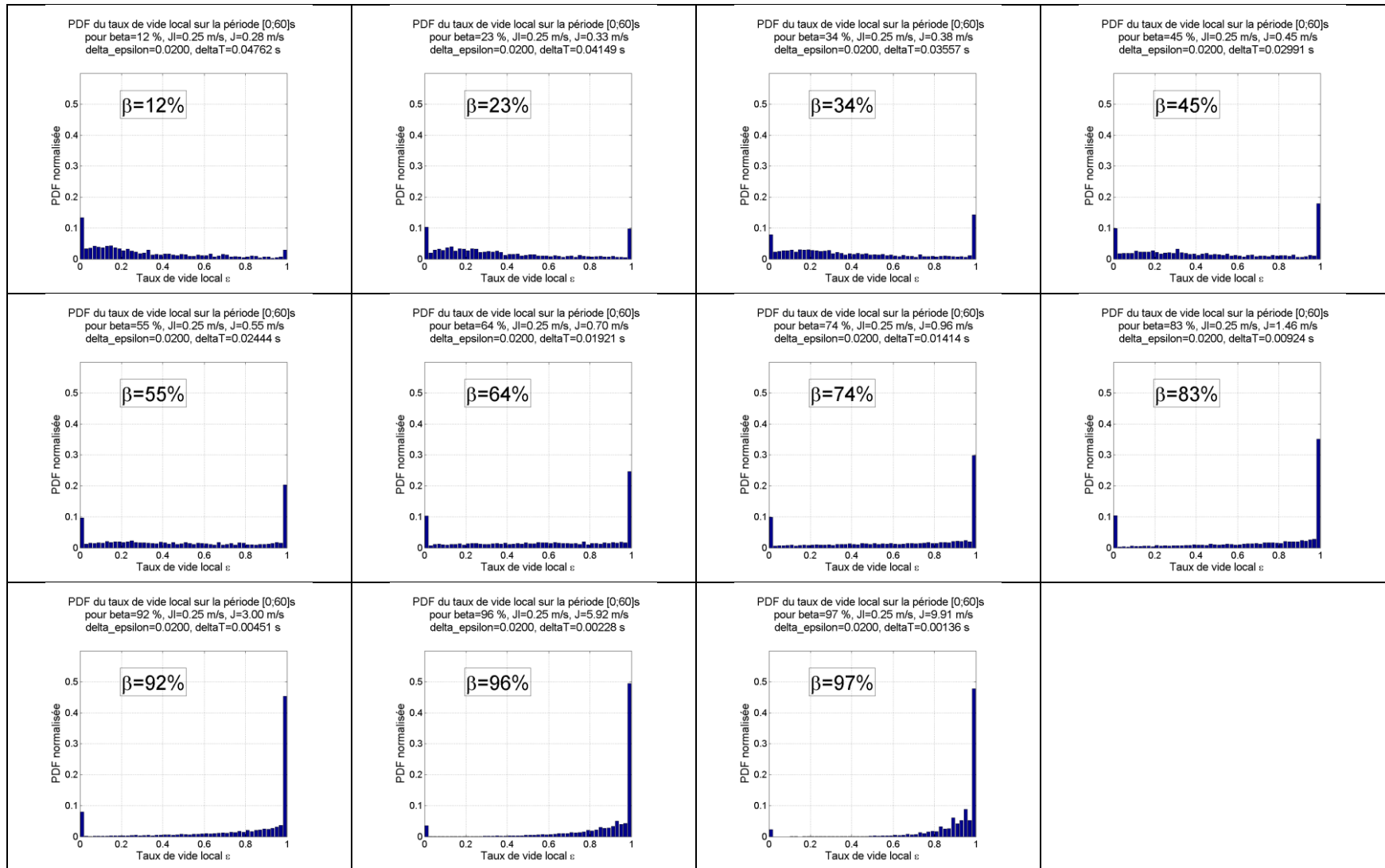
Tableau B. 4 : PDF pour $j_l = 0.25$ m/s

Tableau B. 5 : PDF pour $j_l = 0.30$ m/s

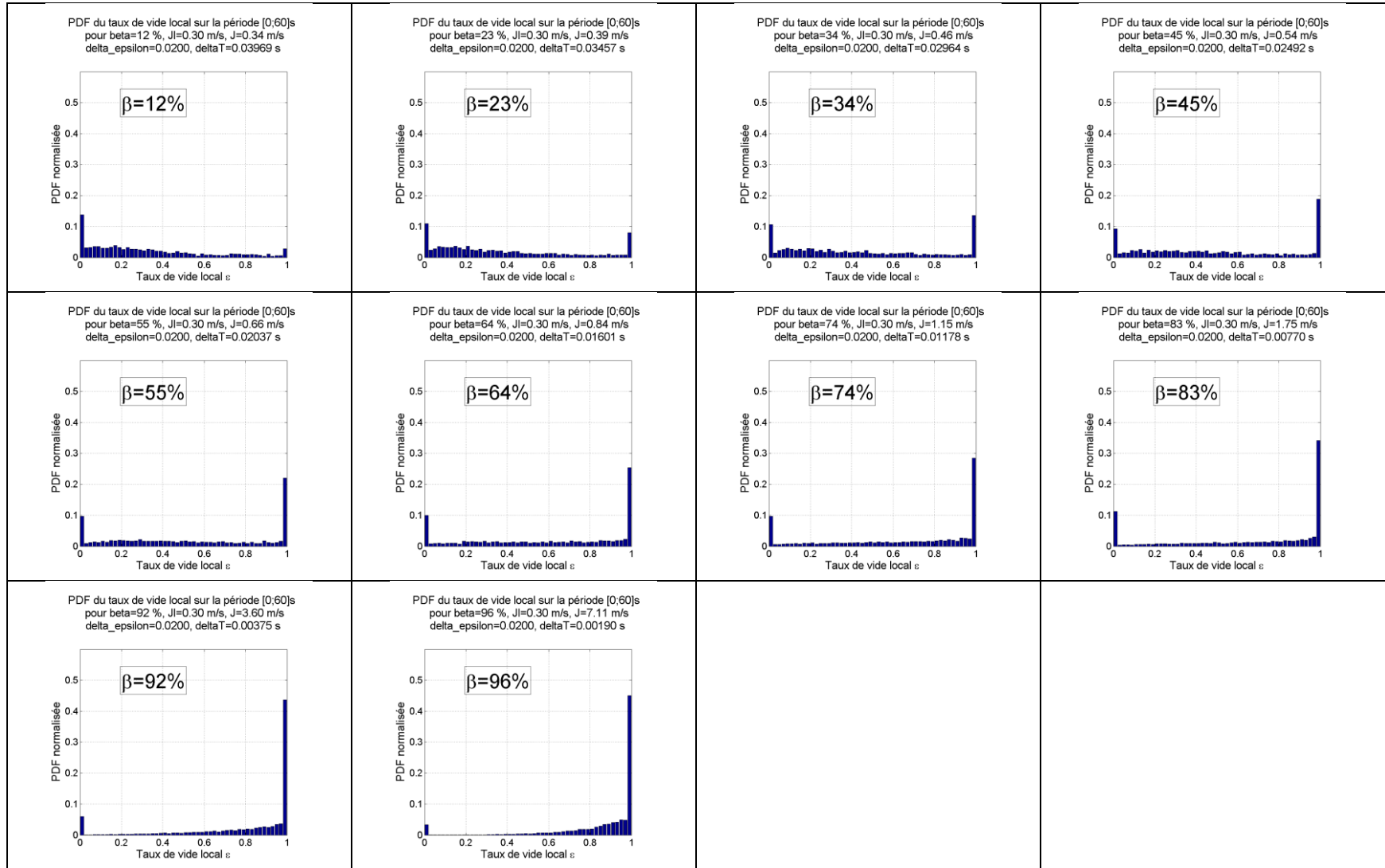


Tableau B. 6 : PDF pour $j_l = 0.35$ m/s

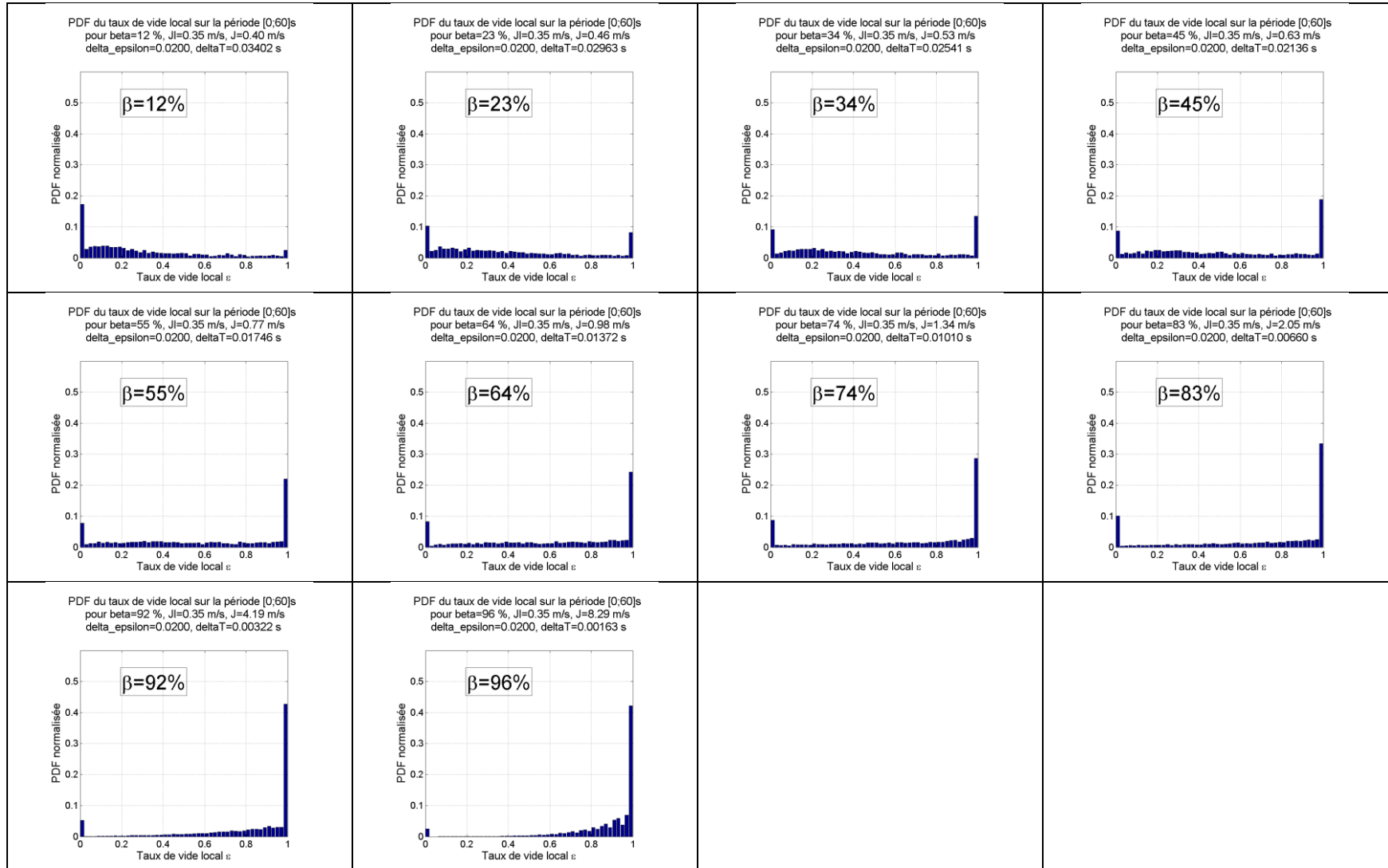


Tableau B. 7 : PDF pour $j_l = 0.40$ m/s

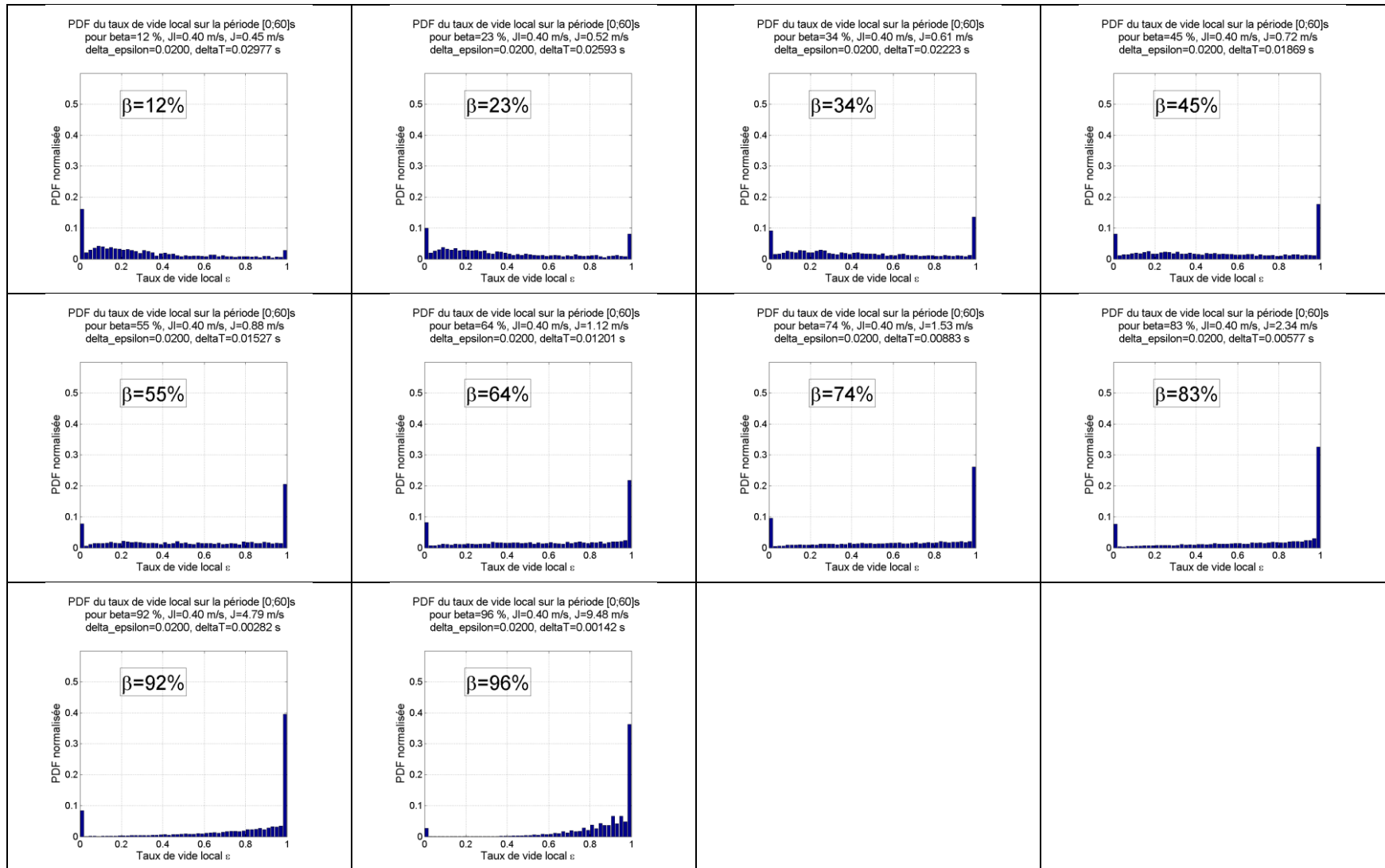


Tableau B. 8 : PDF pour $j_l = 0.50$ m/s

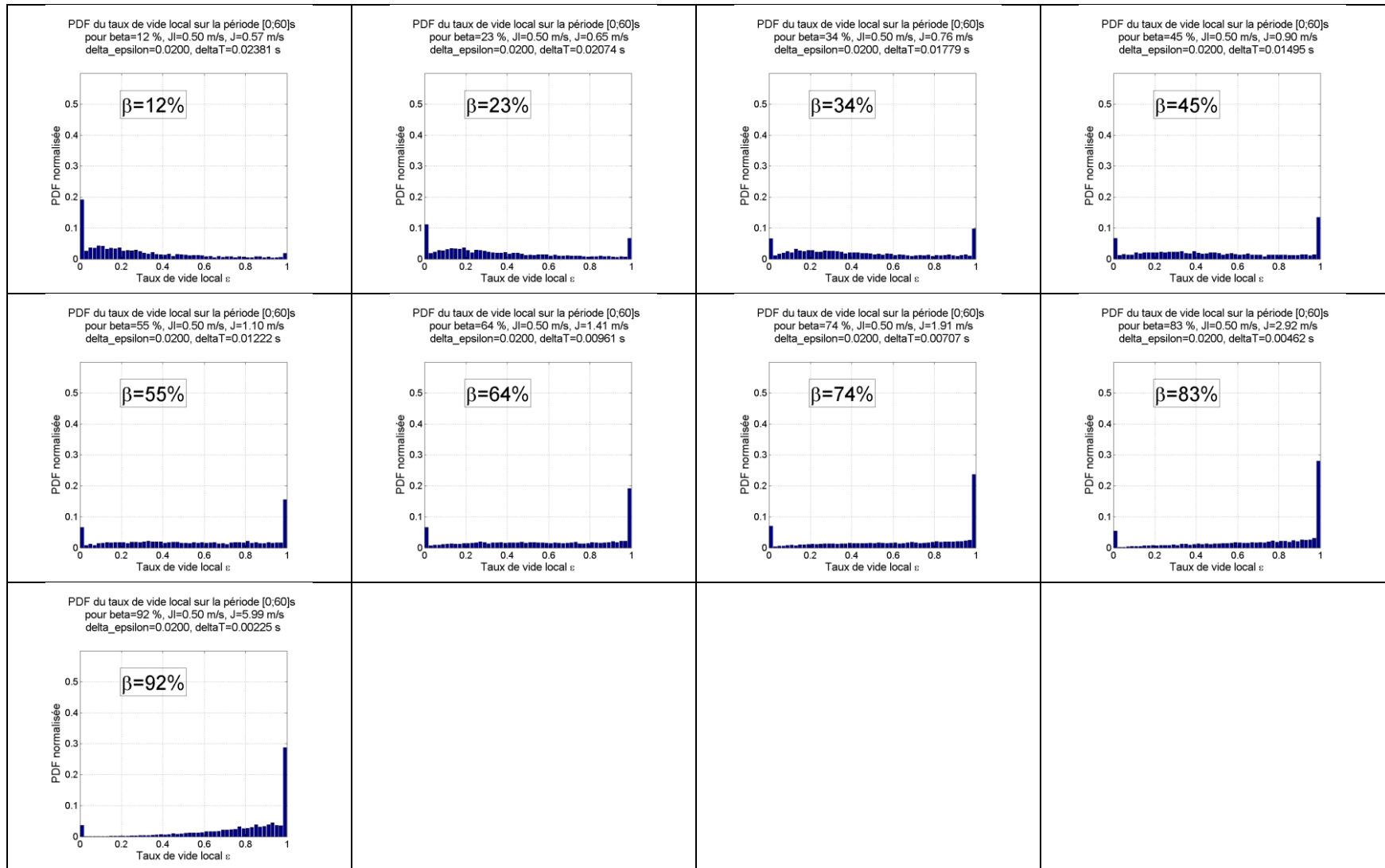


Tableau B. 9 : PDF pour $j_l = 0.60$ m/s

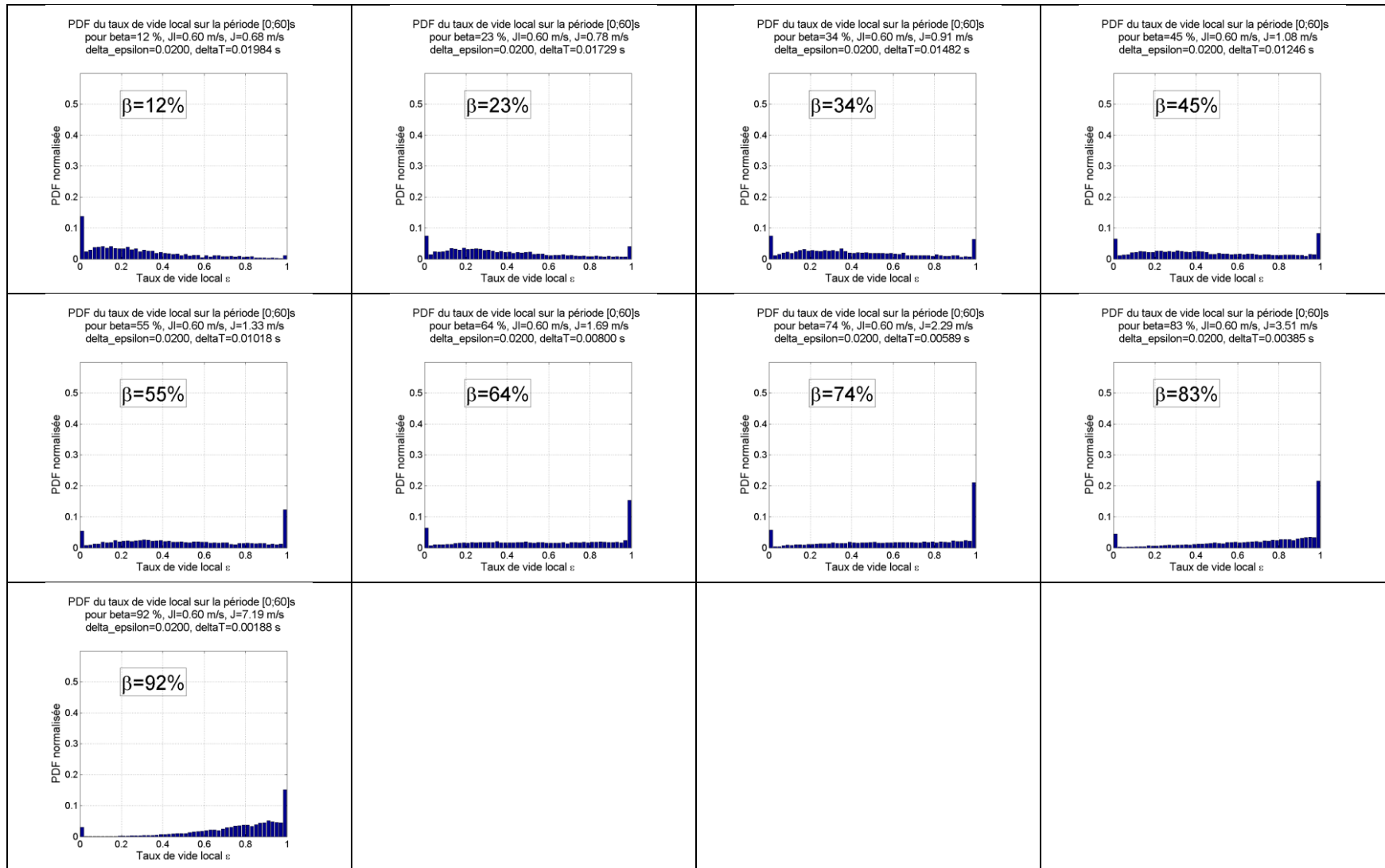


Tableau B. 10 : PDF pour $j_l = 0.68$ m/s

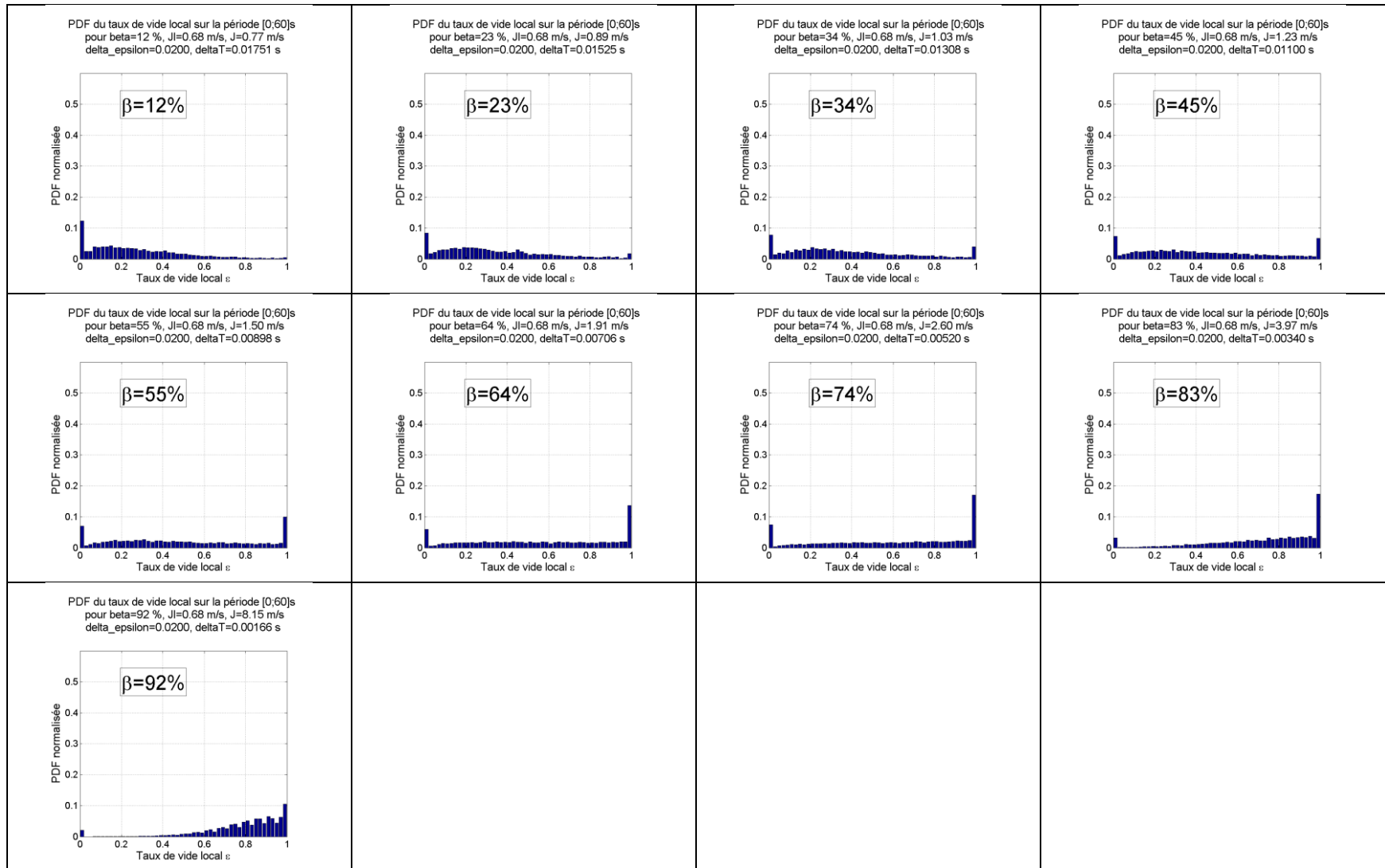
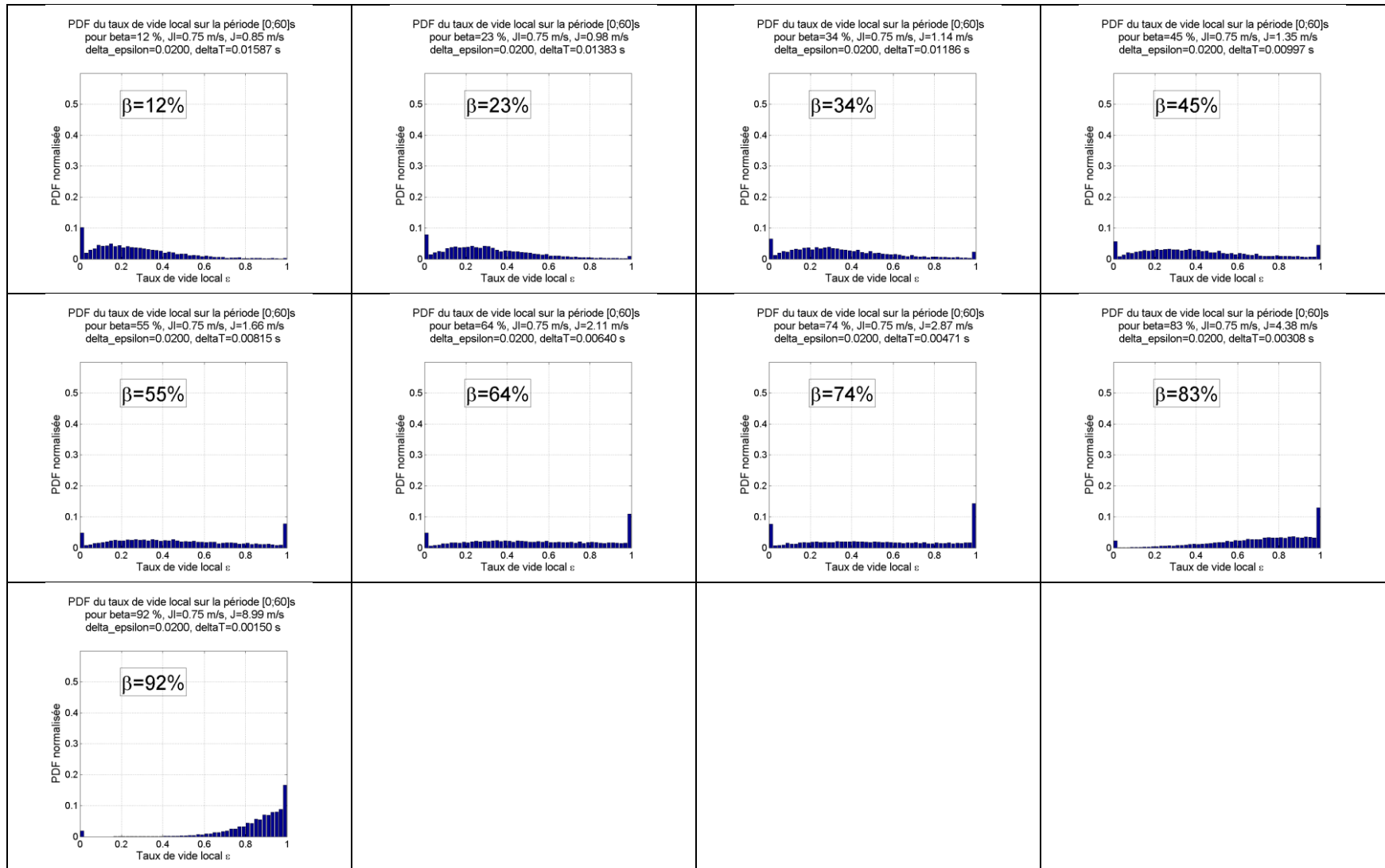


Tableau B. 11 : PDF pour $j_l = 0.75$ m/s



ANNEXE C – MOMENTS STATISTIQUES DU SIGNAL BRUT DE LA SONDE

Tableau C. 1 : Valeurs numériques des moments statistiques du signal brut pour l'ensemble des points de mesure

| β (%) | j_l (m/s) | M | Std | Skw | K |
|-------------|-------------|------|-------|-------|------|
| 12 | 0.1 | 0.29 | 0.32 | 0.76 | 1.65 |
| 12 | 0.15 | 0.30 | 0.32 | 0.84 | 1.75 |
| 12 | 0.2 | 0.28 | 0.33 | 0.81 | 1.69 |
| 12 | 0.25 | 0.29 | 0.32 | 0.75 | 1.59 |
| 12 | 0.3 | 0.28 | 0.32 | 0.73 | 1.55 |
| 12 | 0.35 | 0.28 | 0.34 | 0.81 | 1.68 |
| 12 | 0.4 | 0.29 | 0.33 | 0.78 | 1.64 |
| 12 | 0.5 | 0.25 | 0.28 | 0.95 | 1.97 |
| 12 | 0.6 | 0.29 | 0.34 | 0.90 | 1.85 |
| 12 | 0.68 | 0.26 | 0.34 | 0.97 | 1.97 |
| 12 | 0.75 | 0.26 | 0.32 | 1.04 | 2.12 |
| 23 | 0.1 | 0.28 | 0.30 | 0.65 | 1.48 |
| 23 | 0.15 | 0.29 | 0.30 | 0.45 | 1.24 |
| 23 | 0.2 | 0.33 | 0.32 | 0.45 | 1.25 |
| 23 | 0.25 | 0.33 | 0.34 | 0.44 | 1.21 |
| 23 | 0.3 | 0.31 | 0.33 | 0.46 | 1.23 |
| 23 | 0.35 | 0.31 | 0.32 | 0.39 | 1.18 |
| 23 | 0.4 | 0.36 | 0.34 | 0.39 | 1.19 |
| 23 | 0.5 | 0.30 | 0.29 | 0.46 | 1.24 |
| 23 | 0.6 | 0.34 | 0.37 | 0.46 | 1.23 |
| 23 | 0.68 | 0.30 | 0.33 | 0.66 | 1.48 |
| 23 | 0.75 | 0.29 | 0.33 | 0.76 | 1.62 |
| 34 | 0.1 | 0.33 | 0.33 | 0.45 | 1.27 |
| 34 | 0.15 | 0.35 | 0.34 | 0.29 | 1.13 |
| 34 | 0.2 | 0.38 | 0.36 | 0.33 | 1.16 |
| 34 | 0.25 | 0.37 | 0.34 | 0.16 | 1.04 |

| β (%) | j_l (m/s) | M | Std | Skw | K |
|-------------|-------------|------|-------|-------|------|
| 34 | 0.3 | 0.35 | 0.32 | 0.19 | 1.07 |
| 34 | 0.35 | 0.37 | 0.34 | 0.14 | 1.05 |
| 34 | 0.4 | 0.47 | 0.41 | 0.11 | 1.07 |
| 34 | 0.5 | 0.39 | 0.33 | 0.16 | 1.10 |
| 34 | 0.6 | 0.41 | 0.35 | 0.25 | 1.11 |
| 34 | 0.68 | 0.35 | 0.36 | 0.46 | 1.24 |
| 34 | 0.75 | 0.34 | 0.36 | 0.58 | 1.37 |
| 45 | 0.1 | 0.35 | 0.33 | 0.30 | 1.15 |
| 45 | 0.15 | 0.35 | 0.32 | 0.19 | 1.09 |
| 45 | 0.2 | 0.38 | 0.31 | 0.02 | 1.05 |
| 45 | 0.25 | 0.43 | 0.37 | -0.03 | 1.02 |
| 45 | 0.3 | 0.40 | 0.34 | -0.09 | 1.05 |
| 45 | 0.35 | 0.44 | 0.34 | -0.13 | 1.07 |
| 45 | 0.4 | 0.47 | 0.37 | -0.15 | 1.05 |
| 45 | 0.5 | 0.44 | 0.32 | -0.04 | 1.09 |
| 45 | 0.6 | 0.41 | 0.34 | 0.11 | 1.10 |
| 45 | 0.68 | 0.39 | 0.37 | 0.26 | 1.16 |
| 45 | 0.75 | 0.37 | 0.37 | 0.32 | 1.13 |
| 55 | 0.1 | 0.38 | 0.33 | 0.10 | 1.06 |
| 55 | 0.15 | 0.39 | 0.32 | -0.08 | 1.07 |
| 55 | 0.2 | 0.40 | 0.30 | -0.14 | 1.07 |
| 55 | 0.25 | 0.44 | 0.35 | -0.23 | 1.08 |
| 55 | 0.3 | 0.44 | 0.34 | -0.28 | 1.12 |
| 55 | 0.35 | 0.49 | 0.35 | -0.35 | 1.19 |
| 55 | 0.4 | 0.50 | 0.36 | -0.36 | 1.18 |
| 55 | 0.5 | 0.49 | 0.38 | -0.28 | 1.10 |

Tableau C. 1 (suite) : Valeurs numériques des moments statistiques du signal brut pour l'ensemble des points de mesure

| β (%) | j_l (m/s) | M | Std | Skw | K |
|-------------|-------------|------|-------|-------|------|
| 55 | 0.6 | 0.40 | 0.34 | -0.09 | 1.03 |
| 55 | 0.68 | 0.44 | 0.37 | -0.02 | 1.02 |
| 55 | 0.75 | 0.44 | 0.38 | 0.06 | 1.05 |
| 64 | 0.1 | 0.39 | 0.32 | -0.11 | 1.05 |
| 64 | 0.15 | 0.41 | 0.33 | -0.18 | 1.09 |
| 64 | 0.2 | 0.43 | 0.31 | -0.32 | 1.17 |
| 64 | 0.25 | 0.49 | 0.34 | -0.50 | 1.28 |
| 64 | 0.3 | 0.48 | 0.33 | -0.55 | 1.33 |
| 64 | 0.35 | 0.46 | 0.30 | -0.59 | 1.41 |
| 64 | 0.4 | 0.57 | 0.33 | -0.53 | 1.32 |
| 64 | 0.5 | 0.39 | 0.25 | -0.38 | 1.26 |
| 64 | 0.6 | 0.47 | 0.32 | -0.35 | 1.18 |
| 64 | 0.68 | 0.48 | 0.36 | -0.31 | 1.12 |
| 64 | 0.75 | 0.48 | 0.36 | -0.21 | 1.10 |
| 74 | 0.1 | 0.44 | 0.32 | -0.31 | 1.14 |
| 74 | 0.15 | 0.48 | 0.31 | -0.50 | 1.30 |
| 74 | 0.2 | 0.48 | 0.30 | -0.59 | 1.44 |
| 74 | 0.25 | 0.50 | 0.30 | -0.74 | 1.61 |
| 74 | 0.3 | 0.51 | 0.32 | -0.78 | 1.65 |
| 74 | 0.35 | 0.55 | 0.33 | -0.81 | 1.69 |
| 74 | 0.4 | 0.51 | 0.33 | -0.66 | 1.49 |
| 74 | 0.5 | 0.53 | 0.34 | -0.68 | 1.50 |
| 74 | 0.6 | 0.52 | 0.34 | -0.63 | 1.43 |
| 74 | 0.68 | 0.54 | 0.36 | -0.49 | 1.29 |
| 74 | 0.75 | 0.48 | 0.36 | -0.24 | 1.10 |
| 83 | 0.1 | 0.46 | 0.30 | -0.63 | 1.44 |
| 83 | 0.15 | 0.51 | 0.29 | -0.83 | 1.75 |
| 83 | 0.2 | 0.56 | 0.31 | -0.93 | 1.98 |
| 83 | 0.25 | 0.55 | 0.30 | -1.01 | 2.08 |

| β (%) | j_l (m/s) | M | Std | Skw | K |
|-------------|-------------|------|-------|-------|-------|
| 83 | 0.3 | 0.52 | 0.30 | -0.92 | 1.90 |
| 83 | 0.35 | 0.56 | 0.31 | -0.94 | 1.95 |
| 83 | 0.4 | 0.57 | 0.31 | -0.99 | 2.04 |
| 83 | 0.5 | 0.61 | 0.25 | -0.99 | 2.13 |
| 83 | 0.6 | 0.55 | 0.28 | -1.04 | 2.14 |
| 83 | 0.68 | 0.61 | 0.32 | -1.08 | 2.22 |
| 83 | 0.75 | 0.57 | 0.30 | -1.01 | 2.10 |
| 92 | 0.1 | 0.54 | 0.28 | -1.14 | 2.38 |
| 92 | 0.15 | 0.57 | 0.27 | -1.30 | 2.80 |
| 92 | 0.2 | 0.60 | 0.26 | -1.44 | 3.25 |
| 92 | 0.25 | 0.63 | 0.26 | -1.55 | 3.69 |
| 92 | 0.3 | 0.61 | 0.24 | -1.67 | 4.05 |
| 92 | 0.35 | 0.64 | 0.25 | -1.62 | 3.96 |
| 92 | 0.4 | 0.63 | 0.27 | -1.39 | 3.15 |
| 92 | 0.5 | 0.57 | 0.24 | -1.58 | 3.66 |
| 92 | 0.6 | 0.52 | 0.23 | -1.42 | 3.16 |
| 92 | 0.68 | 0.65 | 0.28 | -1.46 | 3.31 |
| 92 | 0.75 | 0.85 | 0.32 | -2.02 | 5.20 |
| 96 | 0.1 | 0.70 | 0.27 | -1.76 | 4.36 |
| 96 | 0.15 | 0.71 | 0.23 | -2.13 | 6.16 |
| 96 | 0.2 | 0.72 | 0.21 | -2.23 | 6.81 |
| 96 | 0.25 | 0.64 | 0.20 | -2.36 | 7.04 |
| 96 | 0.3 | 0.63 | 0.20 | -2.28 | 6.74 |
| 96 | 0.35 | 0.72 | 0.22 | -2.47 | 7.55 |
| 96 | 0.4 | 0.69 | 0.22 | -2.27 | 6.72 |
| 97 | 0.1 | 0.74 | 0.20 | -2.65 | 8.89 |
| 97 | 0.15 | 0.64 | 0.17 | -2.94 | 10.37 |
| 97 | 0.2 | 0.61 | 0.16 | -2.90 | 10.48 |
| 97 | 0.25 | 0.71 | 0.17 | -3.29 | 12.68 |

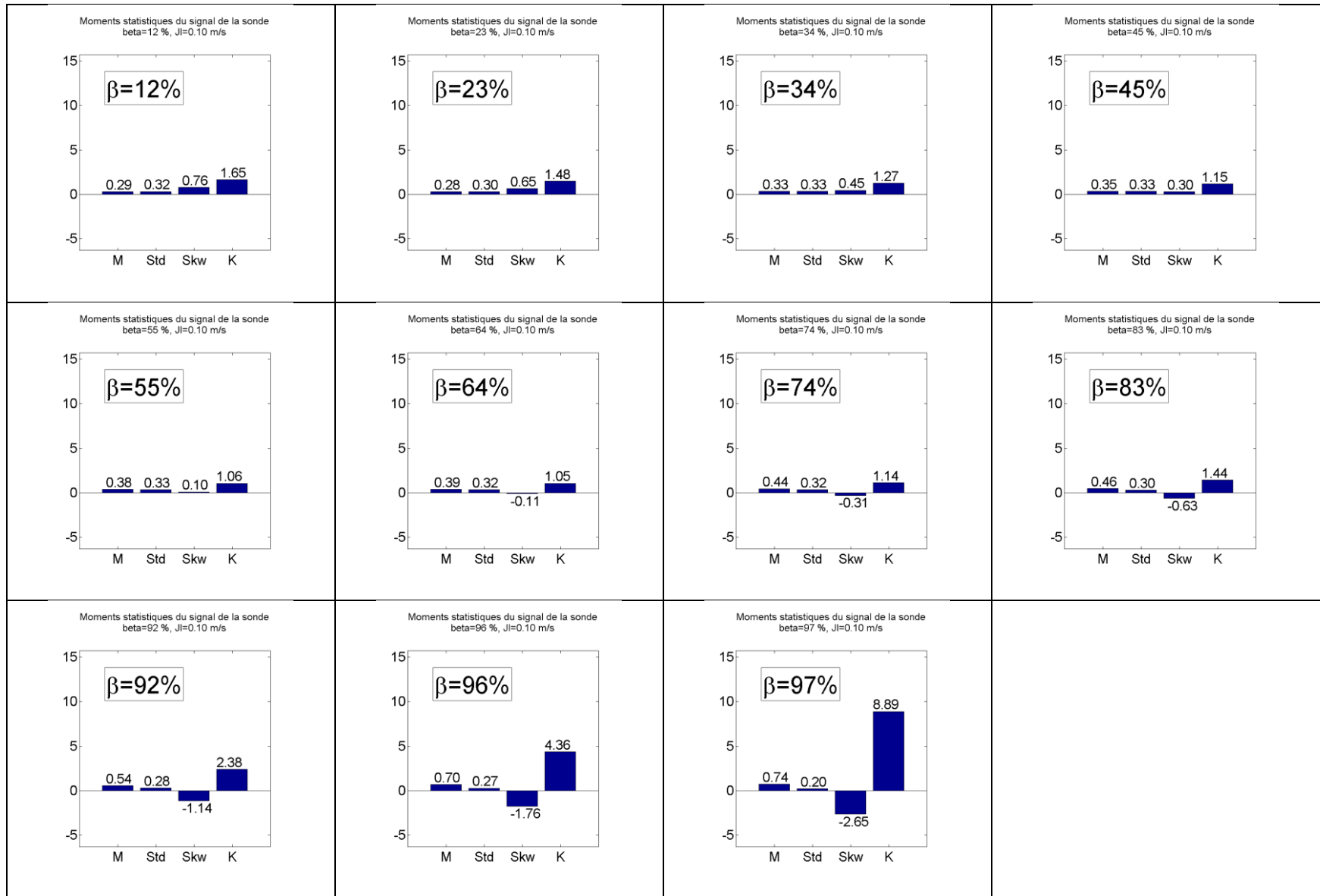
Tableau C. 2 : Graphiques des moments statistiques : $j_l=0.1$ m/s

Tableau C. 3 : Graphiques des moments statistiques : $j_i = 0.15$ m/s

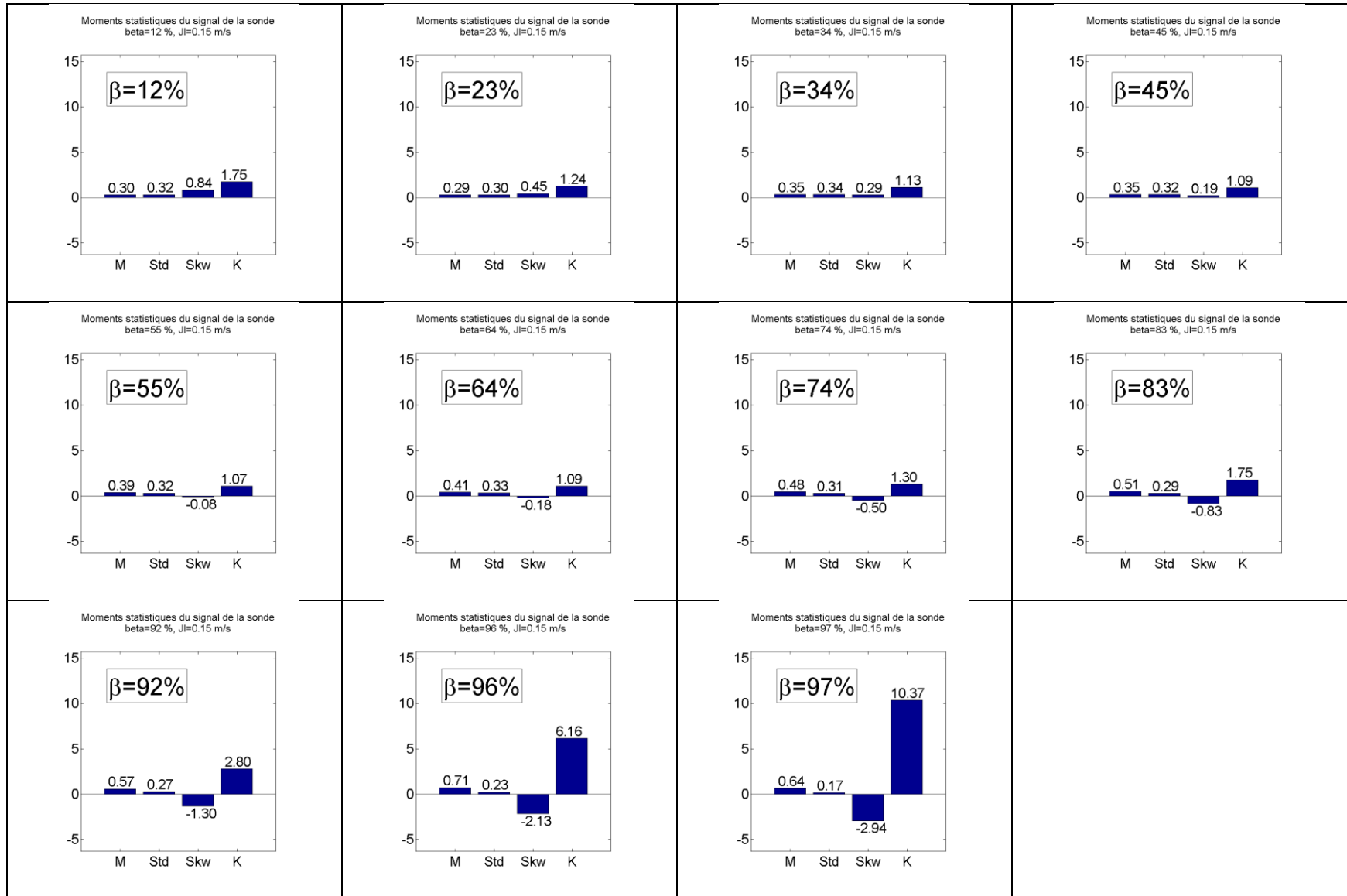


Tableau C. 4 : Graphiques des moments statistiques : $j_l = 0.2$ m/s

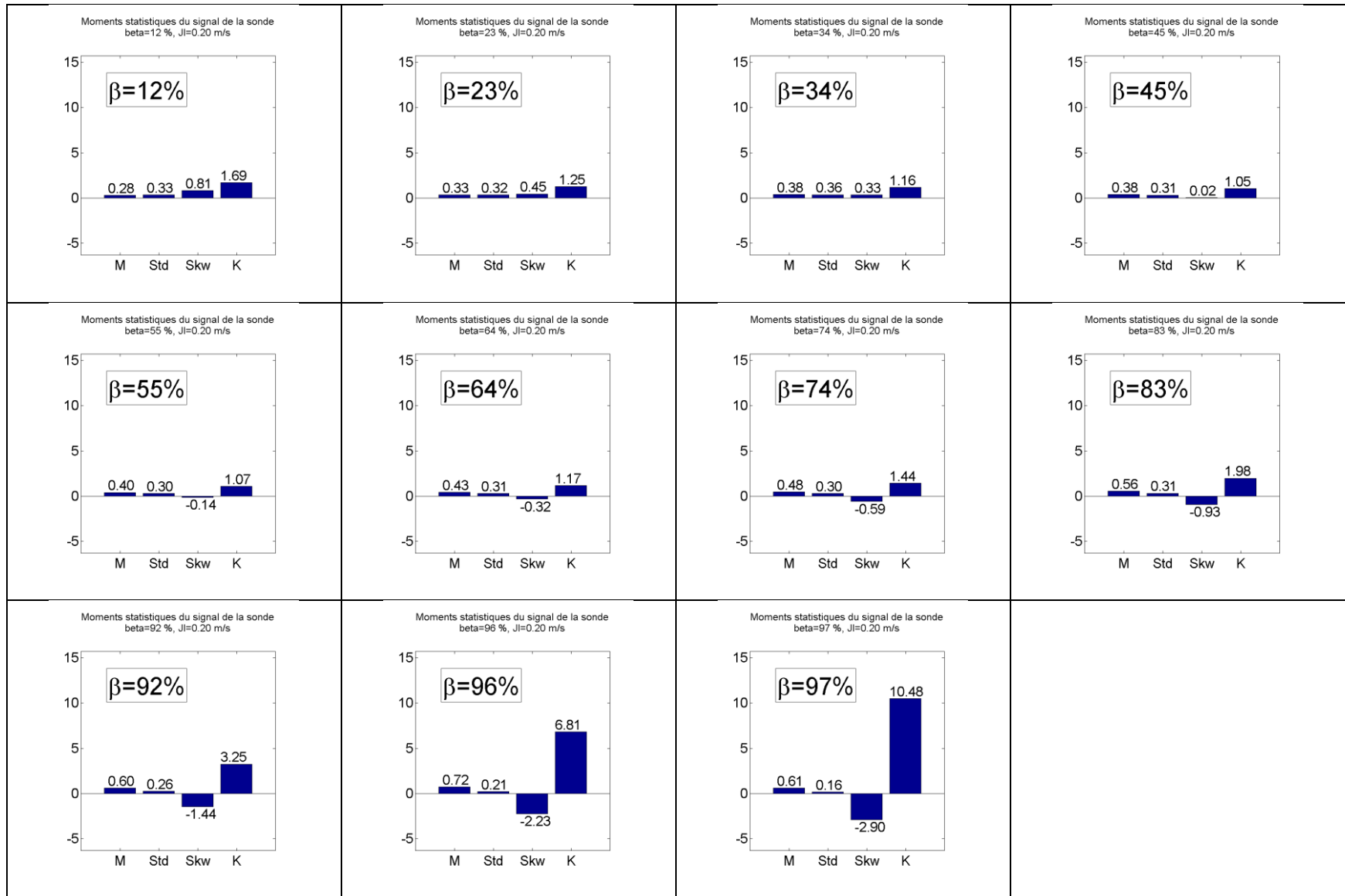


Tableau C. 5 : Graphiques des moments statistiques : $j_l = 0.25$ m/s

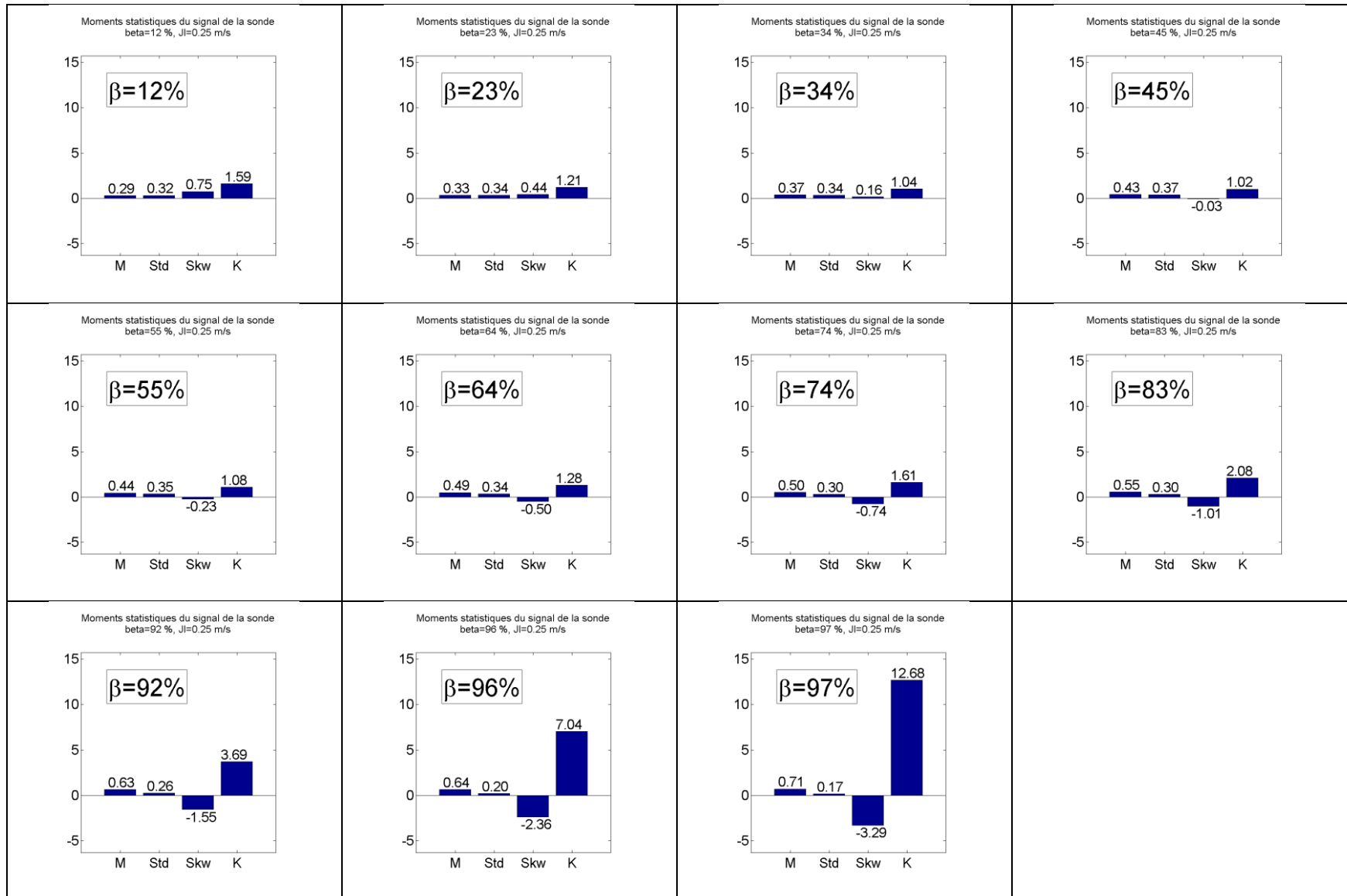


Tableau C. 6 : Graphiques des moments statistiques : $j_l = 0.3 \text{ m/s}$

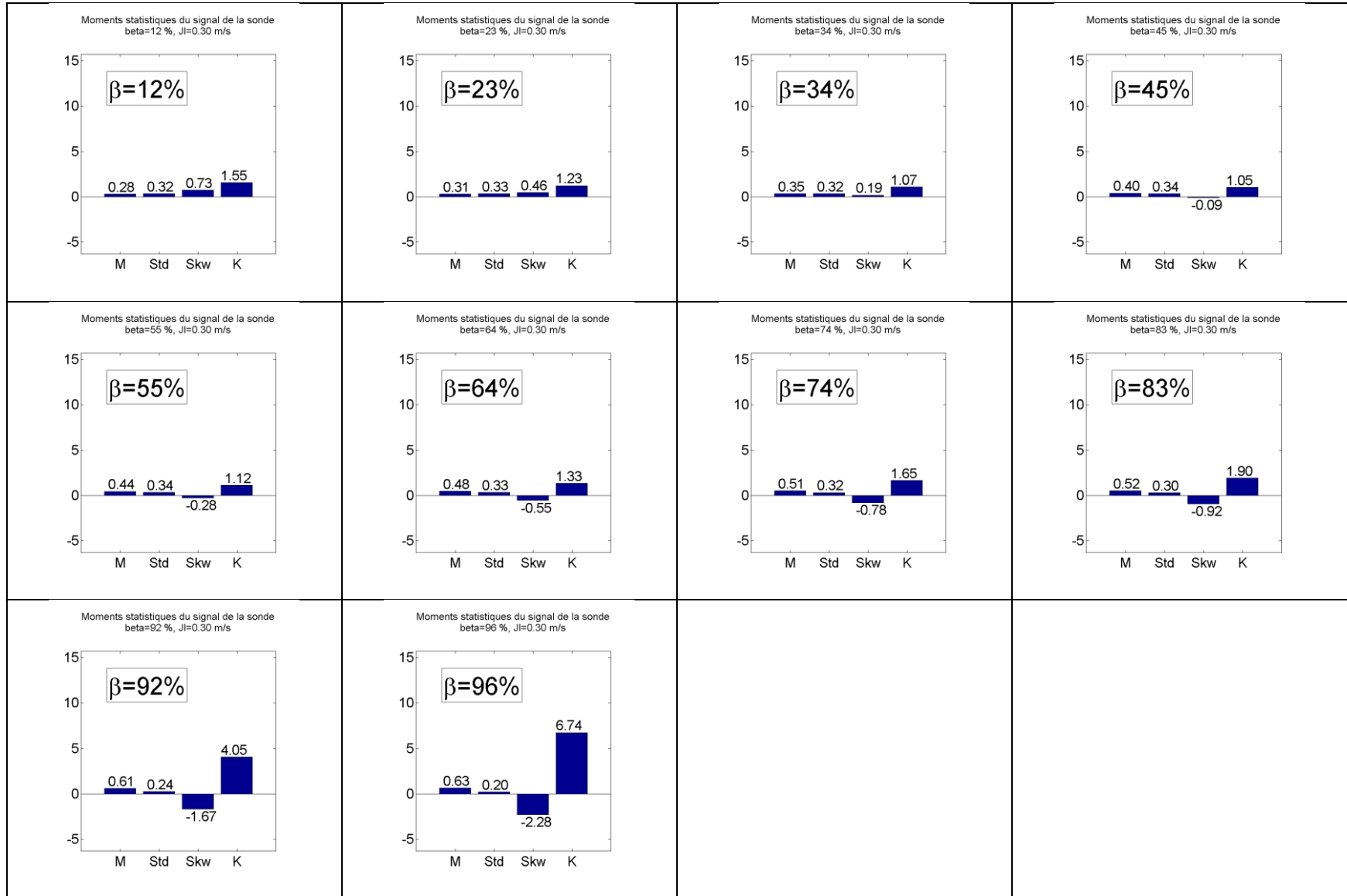


Tableau C. 7 : Graphiques des moments statistiques : $j_l = 0.35$ m/s

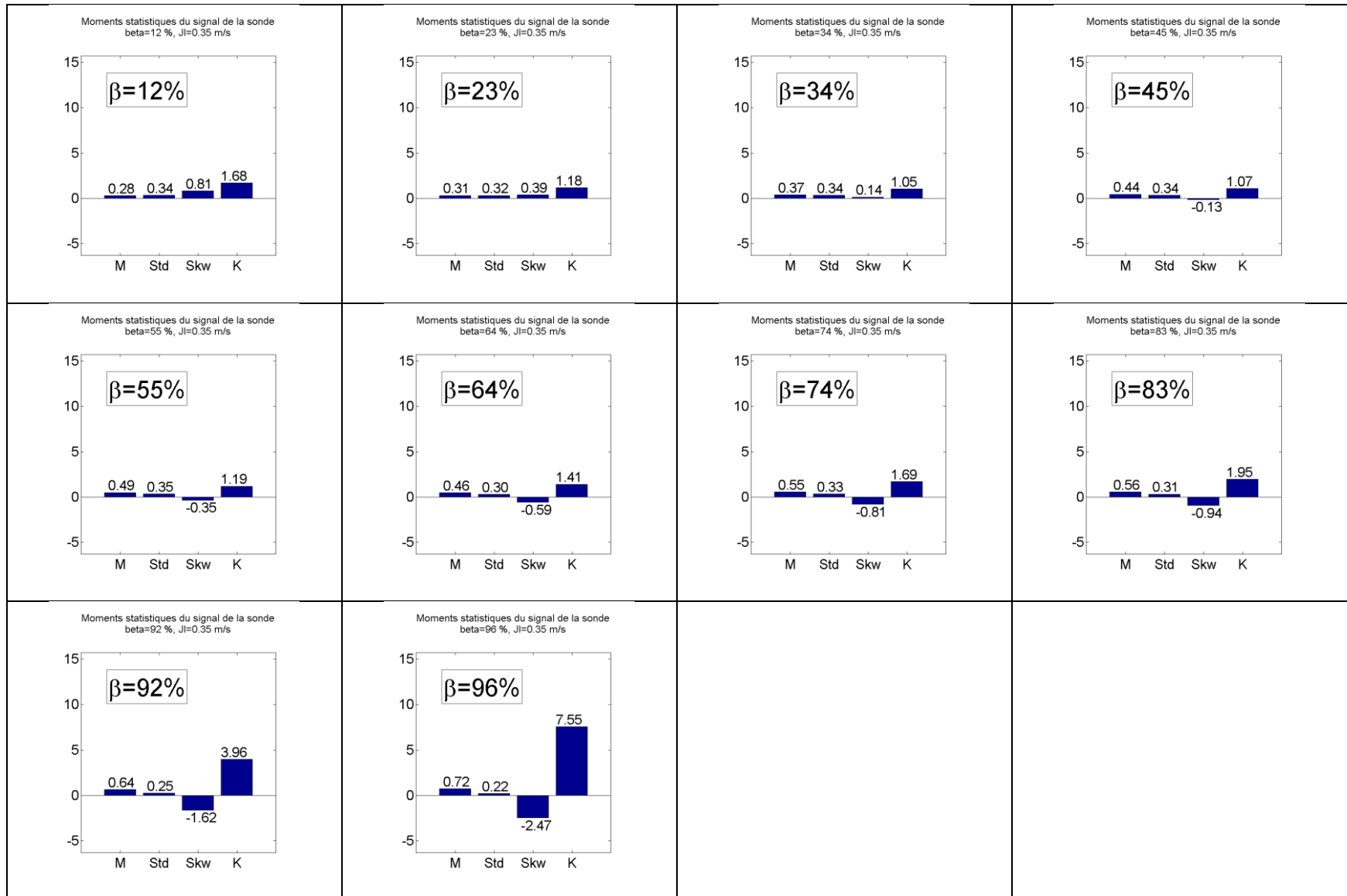


Tableau C. 8 : Graphiques des moments statistiques : $j_l = 0.4 \text{ m/s}$

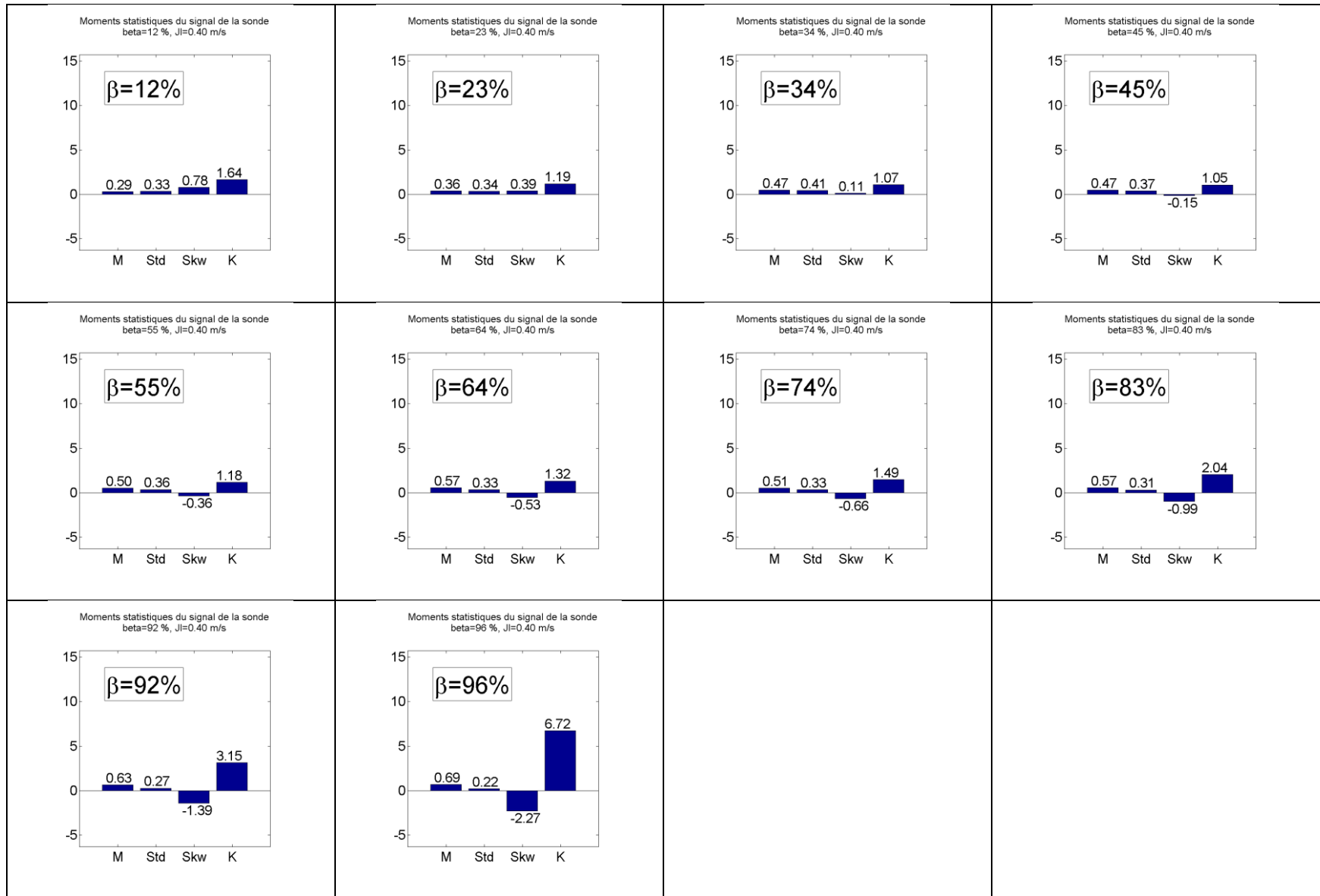


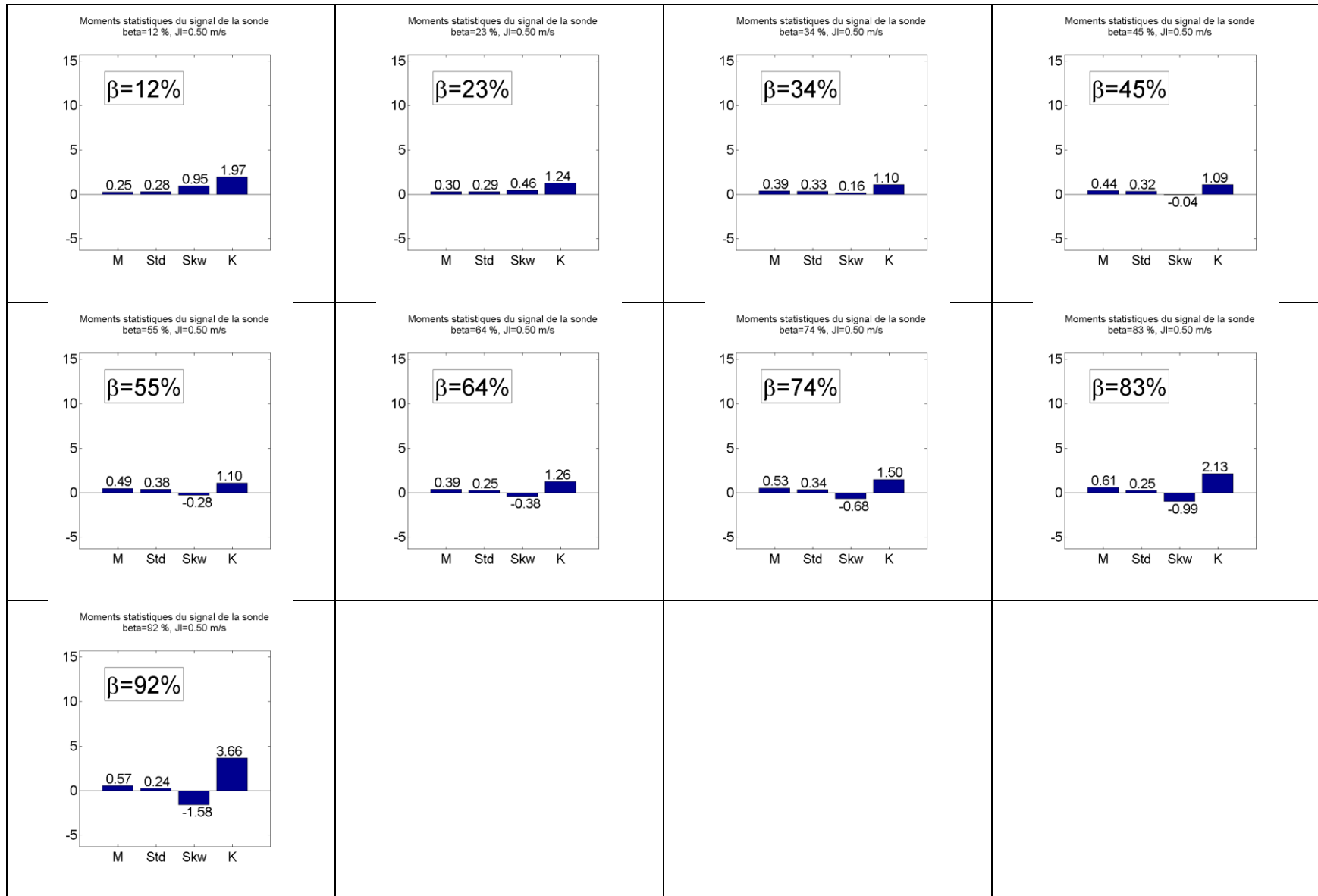
Tableau C. 9 : Graphiques des moments statistiques : $j_l = 0.5$ m/s

Tableau C. 10 : Graphiques des moments statistiques : $j_l = 0.6$ m/s

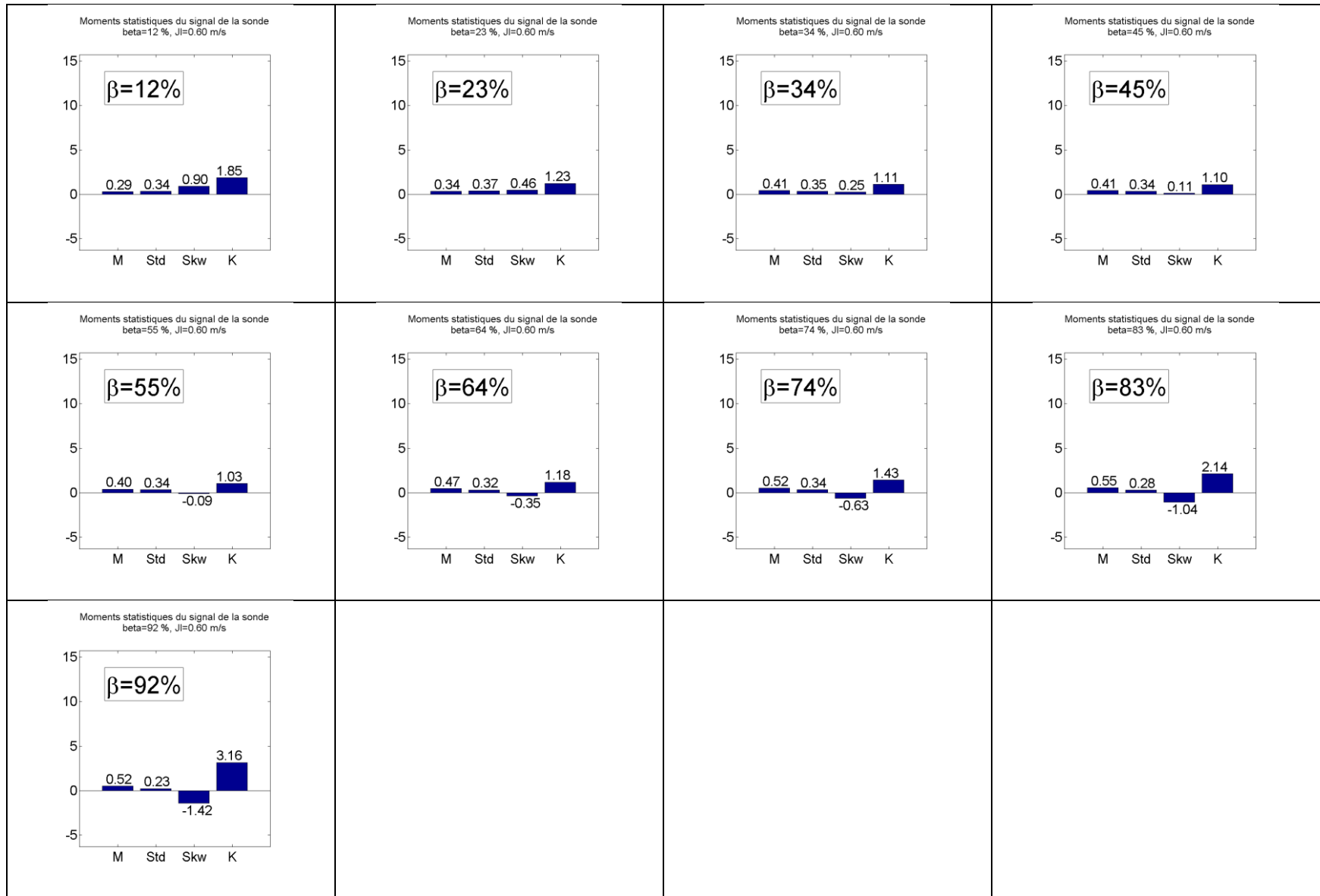


Tableau C. 11 : Graphiques des moments statistiques : $j_l = 0.68$ m/s

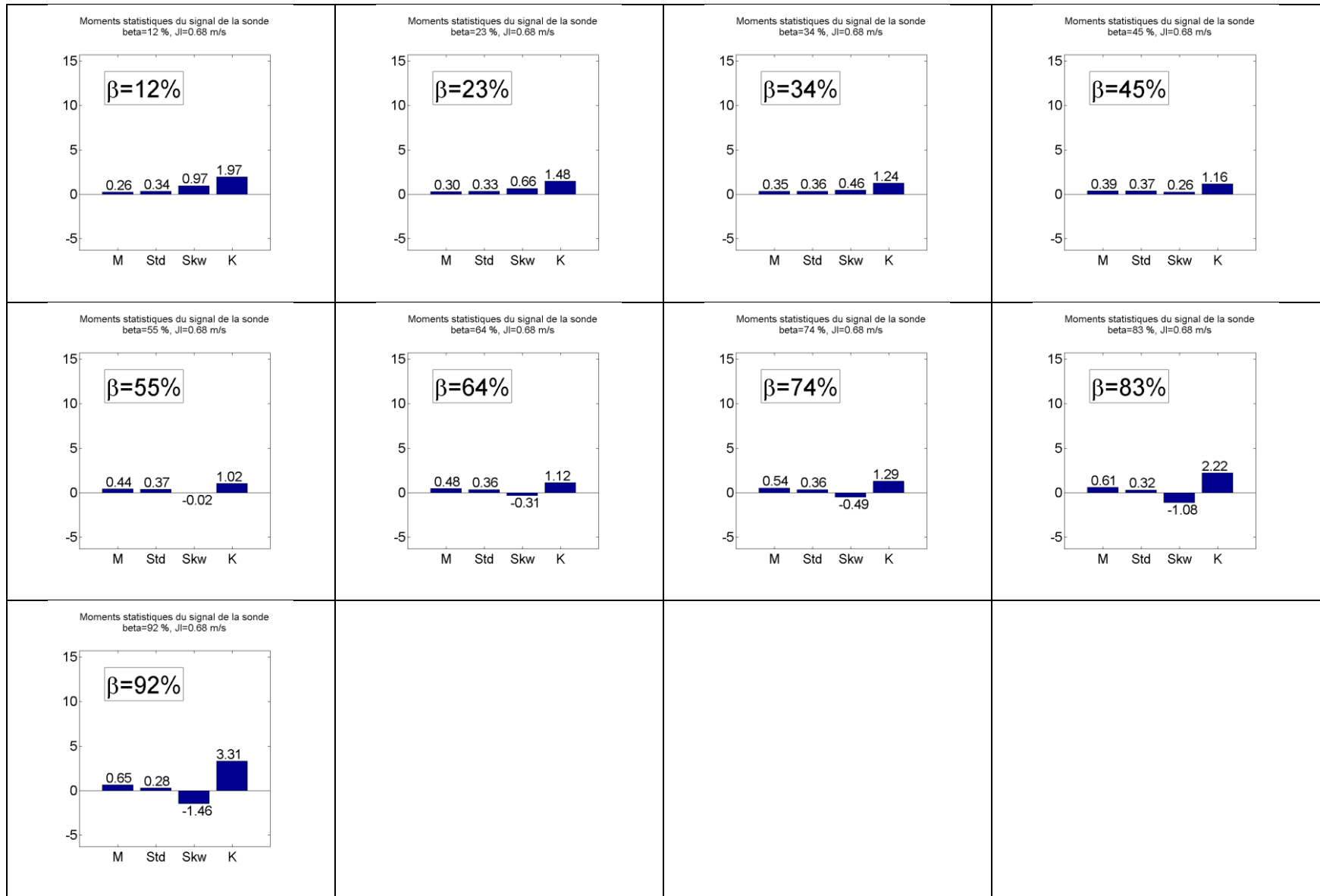
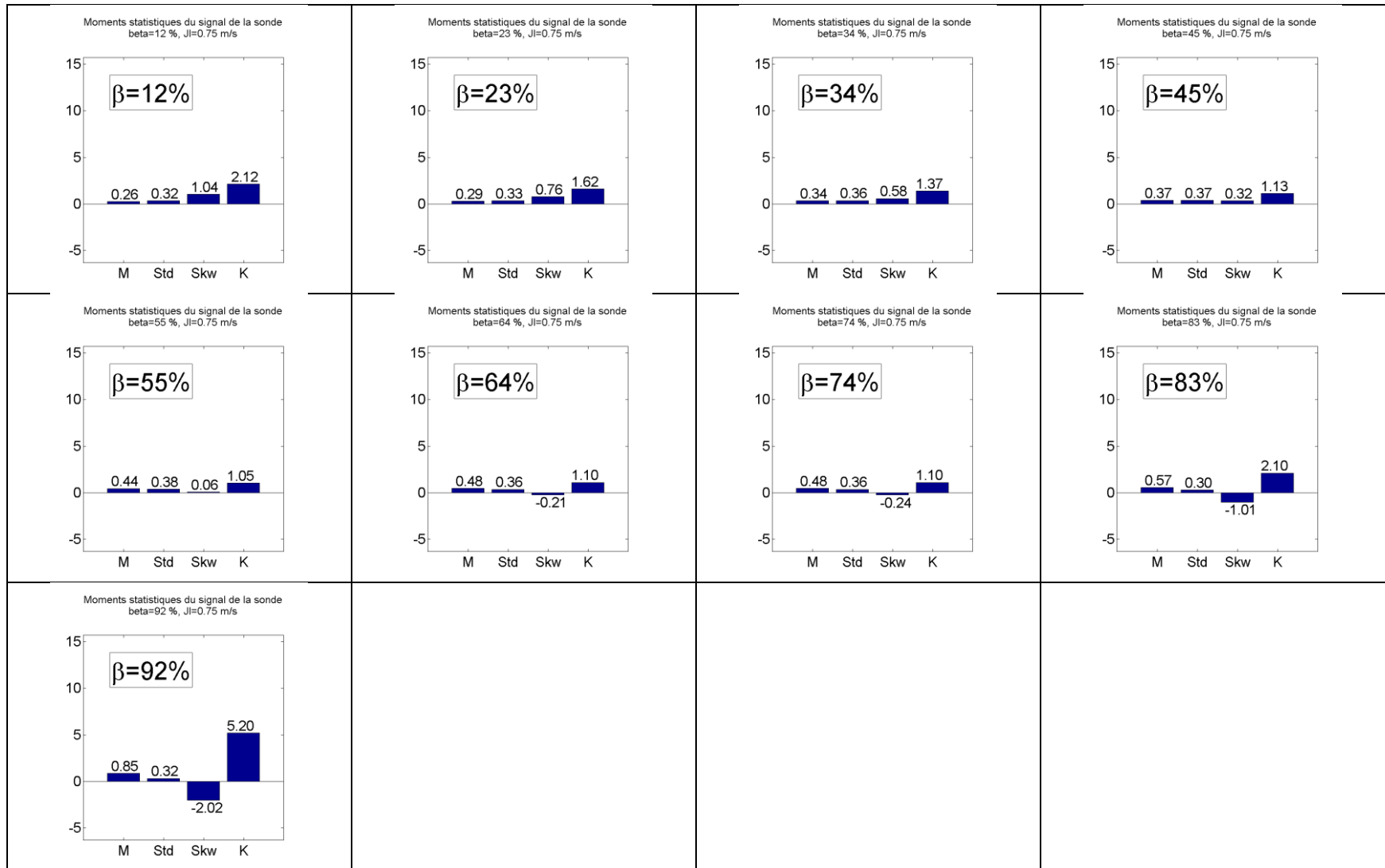


Tableau C. 12 : Graphiques des moments statistiques : $j_l = 0.75$ m/s



ANNEXE D – LIENS VERS LES RESSOURCES UTILISÉES POUR LA SVM

Package LIBSVM

Le package contient les fonctions nécessaires à l'implémentation d'une SVM : entraînement, classification, gestion des multi-classifications, et autres outils, disponibles pour de nombreux langages de programmation. Ce package est maintenu à jour par ses auteurs et a été largement utilisé depuis 2003 (pour donner un ordre d'idées, l'article correspondant – ci-dessous – a été cité 13 333 fois à ce jour d'après Google Scholar).

Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.

Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

Guide d'utilisation pratique proposé sur le site de LIBSVM :

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification.

<https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

Site très complet sur l'utilisation de LIBSVM sous MATLAB

Site personnel de Kittipat Kampa (Integrated Brain Imaging Center, UW Medical Center, Seattle). LIBSVM for Matlab : https://sites.google.com/site/kittipat/libsvm_matlab

Cours en ligne sur le machine learning

accessible sur la plateforme Coursera, donné par Andrew Ng (Stanford University) à l'été 2015. Des cours semblables sont donnés régulièrement sur Coursera.

<https://www.coursera.org/learn/machine-learning/>

ANNEXE E – RÉSULTATS OBTENUS AVEC LES VALEURS PAR DÉFAUT DE C ET γ : MAUVAISE CLASSIFICATION FINALE

On rappelle que pour une C -SVM basée sur une fonction noyau gaussienne, les paramètres C et γ doivent être choisis par l'utilisateur. Il est crucial d'optimiser ce choix, car sinon les résultats, notamment la classification finale, donnés par la SVM peuvent être mauvais.

Afin d'illustrer ce point, on va ici utiliser les valeurs par défaut des programmes de LIBSVM pour le noyau gaussien :

$$C = 1$$

$$\gamma = \frac{1}{n_{feat}} = 0.25$$

où n_{feat} est le nombre de composantes des vecteurs (features), qui vaut ici 4.

Pour une exécution donnée, on réalise donc l'entraînement de la SVM avec les données d'entraînement choisies, dont les valeurs des composantes sont transformées par la fonction-noyau gaussienne avec $\gamma=0.25$, et $C=1$, afin de se placer dans un espace image dans lequel on peut séparer linéairement ces données d'entraînement. On fournit également à la SVM la classe qui a été attribuée à chaque donnée d'entraînement.

En résolvant le problème d'optimisation de la marge (distance entre l'hyperplan séparateur et les vecteurs de support de chaque classe), on obtient un modèle contenant les paramètres qui définissent cet hyperplan séparateur, qui est donc optimal pour *ces* données d'entraînement et *ces* paramètres C et γ .

Comme on est dans un cas de multi-classification, ce procédé d'apprentissage d'une frontière binaire est répété pour chaque classe selon une procédure *One Versus Rest* : on cherche la frontière entre la classe 1 et une classe artificielle composée de {2+3}, puis entre 2 et {1+3} et ainsi de suite. D'un point de vue technique, on entraîne donc en réalité 3 SVM. On construit ensuite un modèle global optimal à partir de ces trois modèles.

On utilise alors ce modèle pour classer les données du *test set*, qui sont des données inconnues pour la SVM. Mais comme ces données ont été choisies parmi les données de référence, on connaît leur « vraie » classe, et on peut comparer la vraie classe avec celle qui est prédite par la SVM. Ces données permettent ainsi de tester la SVM et d'estimer sa performance de façon non biaisée (car le modèle construit ne s'est pas basé sur ces données de test, donc les considère comme des données complètement nouvelles).

Par exemple, pour les valeurs par défaut pour C et γ , et pour l'exécution 1, on obtient la classification donnée Figure E. 1 pour le *test set*. Le losange intérieur représente la vraie classe tandis que le losange extérieur représente la classe prédite par la SVM avec le modèle retenu.

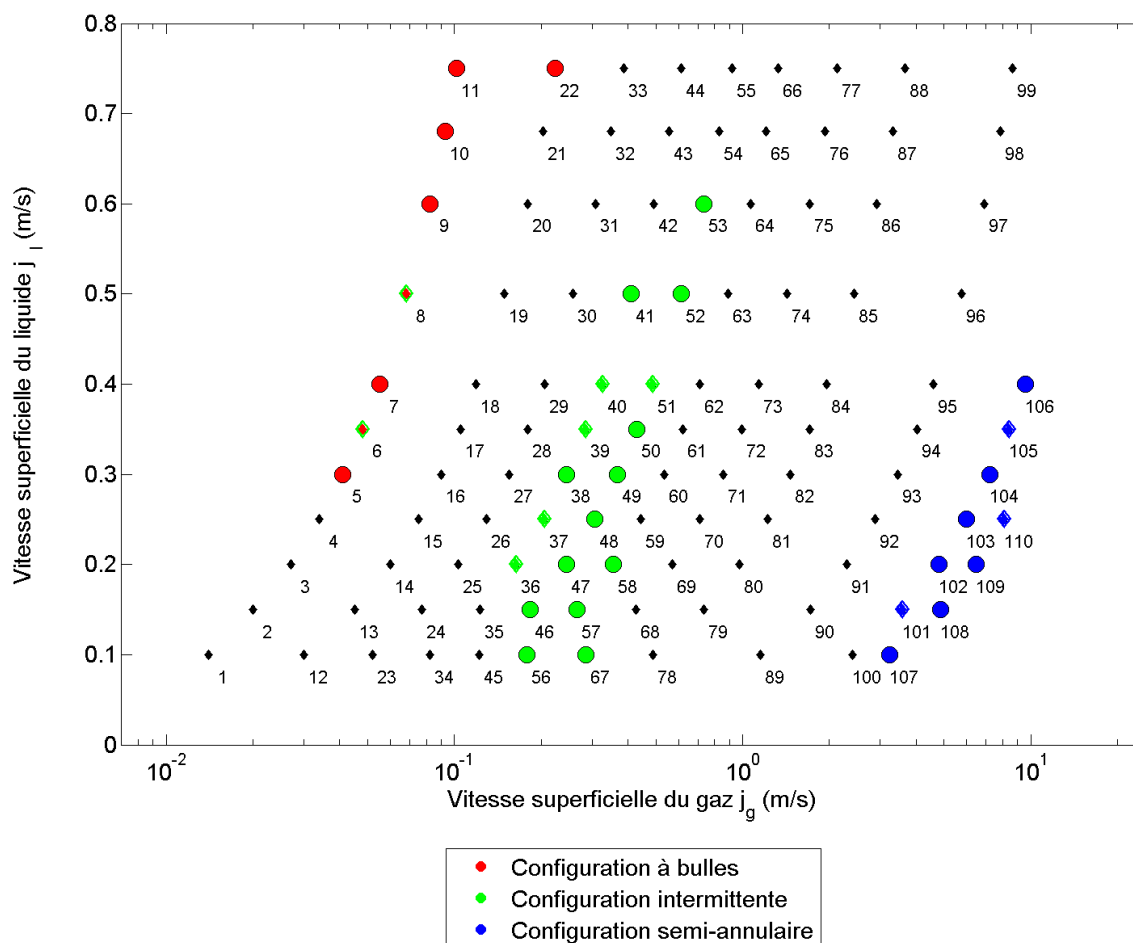


Figure E. 1 : Prédiction des classes du *test set* par la SVM construite pour l'exécution 1 ($C=1$, $\gamma=0.25$)

On constate que deux points du *test set* dont la configuration réelle est à bulles sont prédits en configuration intermittente, et le pourcentage de données de test bien classées à l'issue de la multi-classification (aussi appelé exactitude de la classification par la suite) est 80%, ce qui n'est pas optimal.

Comme les données du *test set* sont considérées par la SVM comme des données inconnues, on peut estimer sur ces points la capacité à généraliser de la SVM, qui serait ici capable de prédire la classe de nouvelles données avec une fiabilité de 80%.

Classification des autres points de mesure par la SVM

On peut alors obtenir grâce à la SVM ainsi construite une prédiction des données de configuration inconnue, dont la fiabilité estimée sera ici de 80%. La classification prédite lors de l'exécution 1 est donnée Figure E. 2, où s'ajoutent aux figures précédentes des losanges extérieurs affichant la classe prédite pour les points dont la configuration était inconnue.

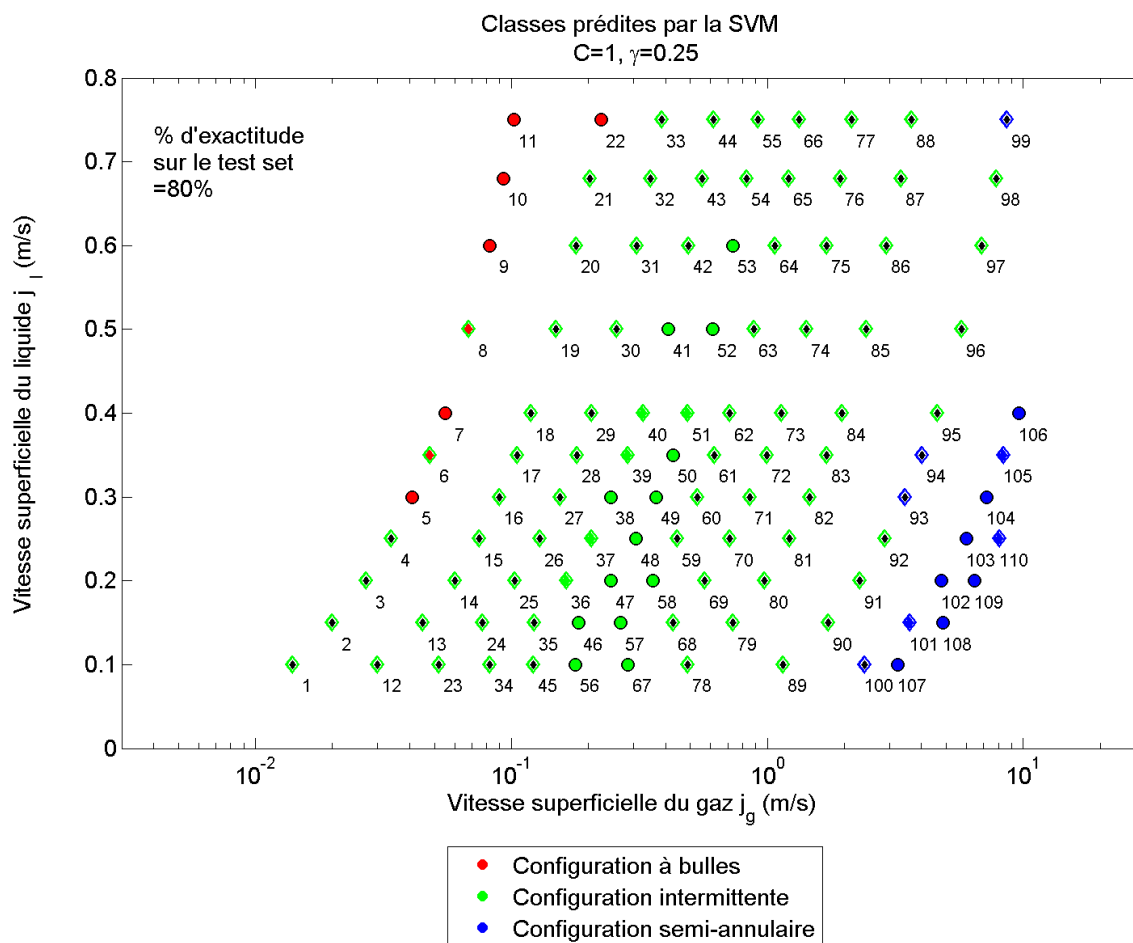
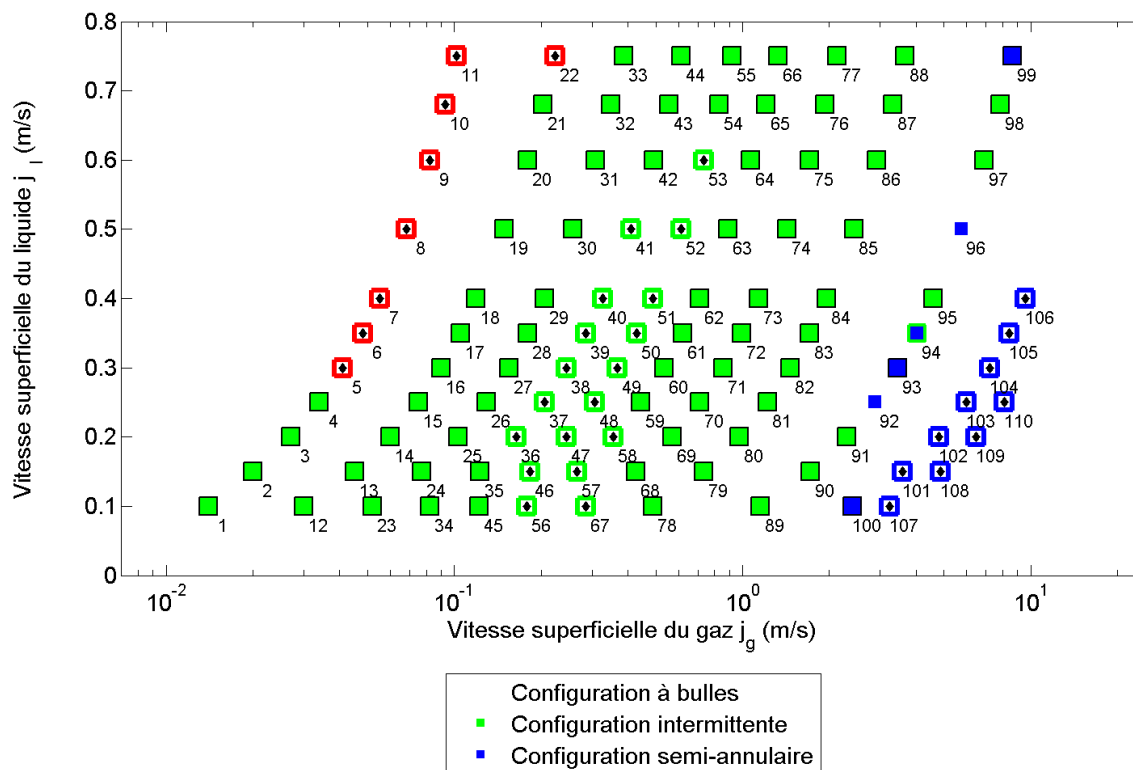


Figure E. 2 : Prédiction des configurations de l'ensemble des points pour l'exécution 1, $C=1$, $\gamma=0.25$

Afin d'obtenir une prédiction plus robuste et indépendante du choix initial des *sets*, on peut effectuer une moyenne sur un ensemble d'exécutions donnant des classifications différentes puisque leurs SVM se sont basées sur des *training sets* différents. Pour cela, on calcule pour chaque point une classe moyenne sur l'ensemble des exécutions, et la « stabilité » de cette classe moyenne prédite, qui correspond à la probabilité qu'un point donné soit prédit dans sa classe moyenne, sur un ensemble d'exécutions. Ce principe est illustré pour les 4 premières exécutions dans la Figure E. 3 où la *vraie* classe des points de référence est représentée par les cadres carrés (les erreurs de classification du *test set* n'apparaissent donc pas ici), et la classe prédite pour les autres points est représentée par un carré plein dont la taille est proportionnelle à la probabilité que la classe

moyenne affichée soit prédite, sur l'ensemble des 4 premières exécutions. Par exemple, la configuration prédite du point 92 est *en moyenne* semi-annulaire, mais il y a eu une ou des exécutions pour laquelle/lesquelles ce point a été prédit en configuration intermittente.



5.15

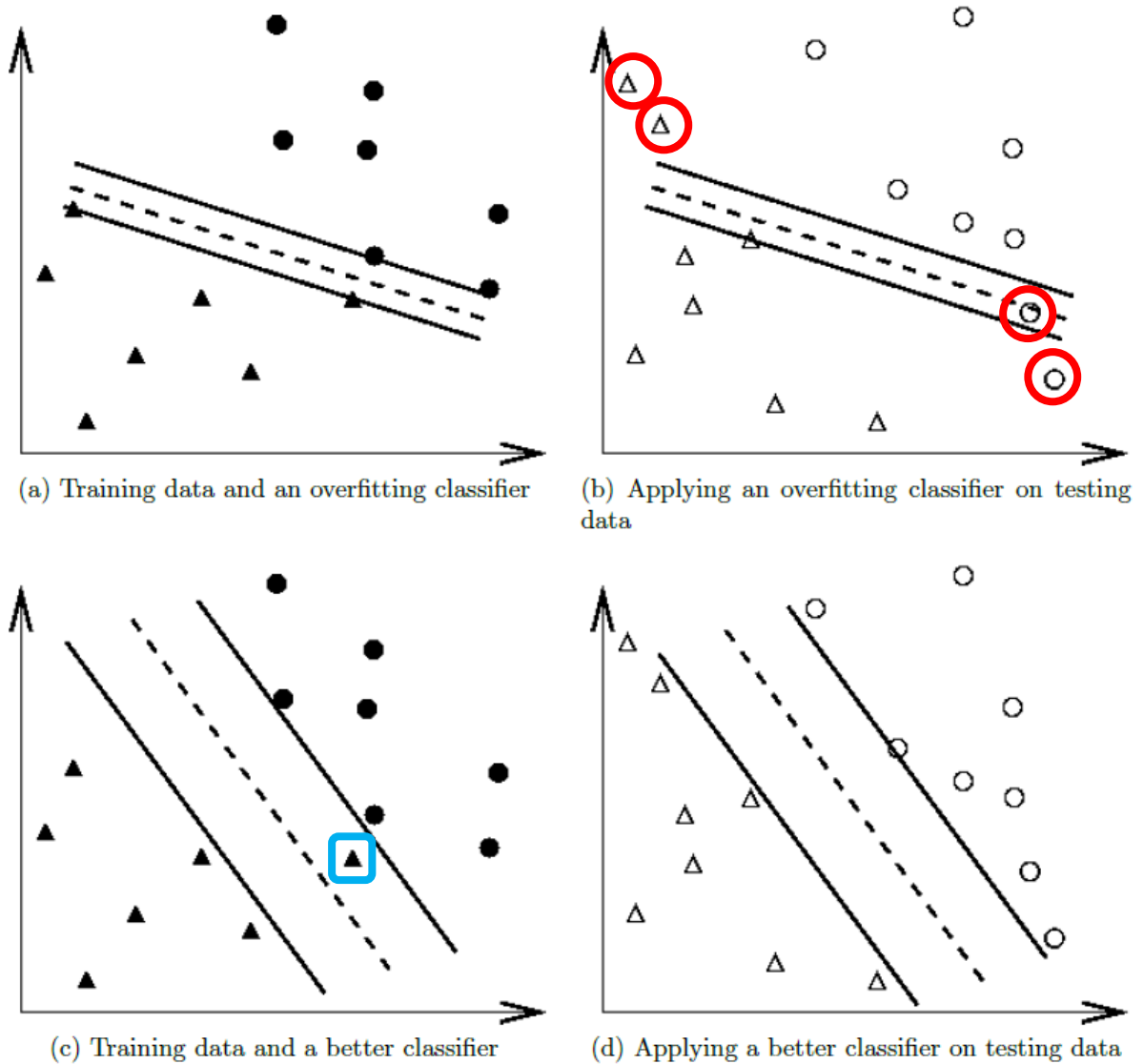
Figure E. 3 : Classification moyenne sur les 4 premières exécutions pour $C=1$ et $\gamma=0.25$

L'exactitude de la classification du *test set* moyennée sur ces 4 exécutions est 77.9%, on peut donc estimer que la SVM peut commettre autour de 22% d'erreur dans la classification proposée pour les points inconnus au départ. On peut aussi émettre un doute sur le fait qu'aucun point n'ait été attribué à la configuration à bulles, surtout en sachant que les erreurs commises dans la classification du *test set* concernaient des points de cette classe.

C'est pour cela qu'il est **absolument nécessaire d'effectuer une procédure de *cross validation*** pour optimiser le choix des paramètres C et γ , comme présenté à la section 6.2.3, qui montre qu'on peut bien obtenir de meilleurs résultats, et prévenir des problèmes d'*underfitting* ou plus souvent d'*overfitting*.

Détails sur le problème d'*overfitting*

Le problème d'*overfitting* est illustré Figure E. 4, où a) et b) représentent dans l'espace image les *training* et *test sets* respectivement, par rapport à une SVM générant de l'*overfitting*, contrairement à la SVM de c) et d) qui se révèle plus flexible et possède donc une meilleure capacité de généralisation car une plus grande marge, ce qui est le plus important. L'hyperplan séparateur est représenté en pointillés.



a) Frontière trop ajustée au training set (\blacktriangle et \bullet)

\Rightarrow b) Mauvaise classification de plusieurs points du test set (\triangle et \circ) : \circ

\neq

c) La SVM « accepte » de mal classer une des données d'entraînement : \square

\Rightarrow d) Tous les points du test set sont bien classés car la marge obtenue est plus grande

Figure E. 4 : Illustration de l'*overfitting* d'après Hsu, Chang, & Lin (2003)