

Décomposition des distributions granulométriques issues des analyses par diffraction laser à des fins de bancairisation des données

M. Masson, L. Richard

Novembre 2023

Document final

Avec le soutien de

Contexte de programmation et de réalisation

Ce rapport a été réalisé dans le cadre du programme scientifique et technique AQUAREF pour l'année 2023, au titre de l'action D1.1a « Etudes scientifiques et techniques, développement et validation de méthodes pour l'analyse ».

Auteurs :

Matthieu Masson
INRAE
matthieu.masson@inrae.fr

Loïc Richard
INRAE
Loic.richard@inrae.fr

Vérification du document :

Jean-Philippe Ghestem
BRGM
jp.ghestim@brgm.fr

Les correspondants

OFB : Pierre-François Staub, pierre-francois.staub@ofb.gouv.fr

INRAE : Marina Coquery, marina.coquery@inrae.fr

Référence du document : Matthieu Masson, Loïc Richard - Décomposition des distributions granulométriques issues des analyses par diffraction laser à des fins de bancarisation des données - Rapport AQUAREF 2023 - 33 p.

Droits d'usage :	<i>Accès libre</i>
Couverture géographique :	<i>International</i>
Niveau géographique :	<i>National</i>
Niveau de lecture :	<i>Professionnels, experts</i>
Nature de la ressource :	<i>Document</i>

1. INTRODUCTION.....	7
2. METHODES DE DECOMPOSITION DES DISTRIBUTIONS GRANULOMETRIQUES	9
2.1 fonction <i>normalmixEM</i> du logiciel R.....	9
2.2 Fonction <i>peakfit</i> du logiciel Matlab.....	10
3. EVALUATION DES OUTILS DE DECOMPOSITION	11
3.1 A partir de distributions générées numériquement	11
3.2 Tests de mélanges de particules en laboratoire	15
3.2.1 <i>Description de l'expérimentation</i>	15
3.2.2 <i>Modélisation des mélanges</i>	16
3.3 distributions granulométriques d'échantillons de MES réels (station de Jons, Rhône)	18
3.3.1 <i>Description des données</i>	18
3.3.2 <i>Comparaison des décompositions distributions granulométriques des particules prélevées dans le Rhône à Jons</i>	18
4. MISE EN PRATIQUE DE LA DECOMPOSITIONS SUR UN GROS JEU DE DONNEES ET QUALITE DES DONNEES BANCARISEES	22
4.1.1 <i>Description du jeu de données et du modèle utilisé</i>	22
4.1.2 <i>Décomposition granulométrique des MES du Rhône et de ses principaux affluents</i>	23
4.1.3 Bancarisation des données	25
5. CONCLUSIONS ET PERSPECTIVES	27

Liste des annexes :

ANNEXE 1 : décompositions des fractions tamisés A, B et C avec les modèles Matlab et R.

ANNEXE 2 : résultats des décompositions granulométriques (modes, proportions et écart quadratique moyen RMSE) à l'aide des modèles R et Matlab pour les échantillons de MES collectés par centrifugation/pompage (n=77) et par pièges à particules (n=61) à la station de Jons (Rhône) entre août 2012 et juin 2016.

RESUME

La mesure de la distribution granulométrique est un paramètre essentiel à la caractérisation des sédiments et des matières en suspension. La méthode par diffraction laser est de plus en plus utilisée, mais génère des formats de résultats différents (nombre et valeurs des classes de particules) en fonction des granulomètres utilisés. Afin d'uniformiser les résultats, il est proposé de décomposer les distributions granulométriques mesurées en sous-populations gaussiennes dont les caractéristiques (mode, écart-type et proportion) pourront être bancariser de façon homogène avec un nombre restreint de valeurs. Dans cette étude, deux fonctions, *normalmixEM* du package *R Mixtools* et *peakfit* du logiciel *Matlab*, ont été testées sur des distributions granulométriques issues de simulations numériques, d'expérimentations de mélanges de particules en laboratoire et de MES de rivières collectées dans le cadre de l'Observatoire des Sédiments du Rhône (OSR). L'évaluation des résultats a été réalisée en se basant (i) sur la valeur de l'écart quadratique moyen (RMSE) entre la distribution modélisée (somme des distributions des sous-populations gaussiennes modélisées) et la distribution mesurée, et (ii) sur l'adéquation entre les paramètres (modes, écarte-types et/ou proportions) extraits des distributions modélisées et les paramètres théoriques. Dans la plupart des cas, les deux outils donnent des résultats satisfaisants. L'outil du logiciel *Matlab* semble le mieux adapté pour réaliser ces décompositions avec une utilisation plus simple, plus fiable et un temps de calcul beaucoup plus restreint que l'outil *R*. Cette étude montre qu'il est possible de bancariser de façon fiable les distributions granulométriques d'échantillons de MES (voire de sédiments) en un jeu de 13 paramètres (triplets de modes, écart-types et proportions pour 4 sous-population, et un critère de qualité comme le RMSE), et cela indépendamment du nombre de classes mesurées par les granulomètres laser. Pour cela, des améliorations doivent être apportées pour automatiser et fiabiliser l'utilisation de la fonction *Matlab*, aider à la validation des résultats et rendre l'outil accessible à des personnes non expertes. Enfin, ce rapport présente quelques exemples d'utilisation des résultats de la décomposition pour répondre à des préoccupations plus scientifiques.

Mots clés (thématique et géographique) :

Distribution granulométrique, algorithme de décomposition, sous-population gaussienne, matières en suspension, bancarisation de données

ABSTRACT

Measuring particle size distribution (PSD) is an essential parameter for characterizing sediments and suspended solids. The laser diffraction method is increasingly used, but generates different result formats (number and values of particle classes) depending on the equipment used. In order to standardize the results, we propose to decompose the measured PSD into Gaussian subpopulations whose characteristics (mode, standard deviation and proportion) could be uniformly bancarized with a limited number of values. In this study, two functions, *normalmixEM* from the *R* package *Mixtools* and *peakfit* from *Matlab* software, were tested on PSD derived from numerical simulations, laboratory experiments with particle mixtures and suspended solids collected as part of the Rhône Sediment Observatory (OSR). Results were evaluated on (i) the value of the root mean squared error (RMSE) between the simulated (sum of the simulated Gaussian subpopulations) and measured distributions, and (ii) the accuracy between the parameters (modes, standard deviations and/or proportions) extracted from the simulated distributions and the theoretical parameters. In most cases, both tools give satisfactory results. The *Matlab* software tool appears to be the most effective for running the PSD decomposition, as it is easier to use, more reliable and needs considerably less calculation time than the *R* tool. This study shows that the particle size distributions of TSS samples (or even sediments) can be reliably bancarized into a set of 13 parameters (triplets of modes, standard deviations and proportions for 4 subpopulations, and a quality criteria such as RMSE), independently of the number of classes measured by laser particle size analyzers. To achieve this, improvements need to be implemented to automate the *Matlab* function, improve its reliability, assist in the validation of results and provide non-expert access to the tool. Finally, a few examples of using the decomposition results to address more scientific concerns are presented in this report.

Key words (thematic and geographical area):

grain-size distribution, decomposition algorithm, Gaussian subpopulation, suspended matter, data bancarization

1. INTRODUCTION

La nature des sédiments et des matières en suspension (MES) dans les cours d'eau, et en particulier leur granulométrie et leur composition en matière organique, peut influencer les concentrations en contaminants. En effet, il est communément admis que les contaminants ont une forte affinité pour les particules les plus fines en raison de leurs surfaces spécifiques plus élevées (Duinker, 1986; Olsen et al., 1982; Pierard et al., 1996) ou pour les particules riches en matière organique (Olsen et al., 1982). C'est pourquoi la distribution granulométrique des particules est un des paramètres à surveiller lors des opérations d'échantillonnage et d'analyse des sédiments (Coquery et al., 2017). Ainsi, Yari et al. (2019) suggèrent d'interpréter les données de surveillance des sédiments en prenant en compte une correction granulométrique (par exemple la proportion massique de la fraction des particules de diamètres inférieurs à 63 μm). Des relations entre la distribution granulométrique et les concentrations en certains métaux (e.g. Al, Fe) peuvent aussi être utilisées pour normaliser les concentrations (e.g. Dabrin et al., 2021).

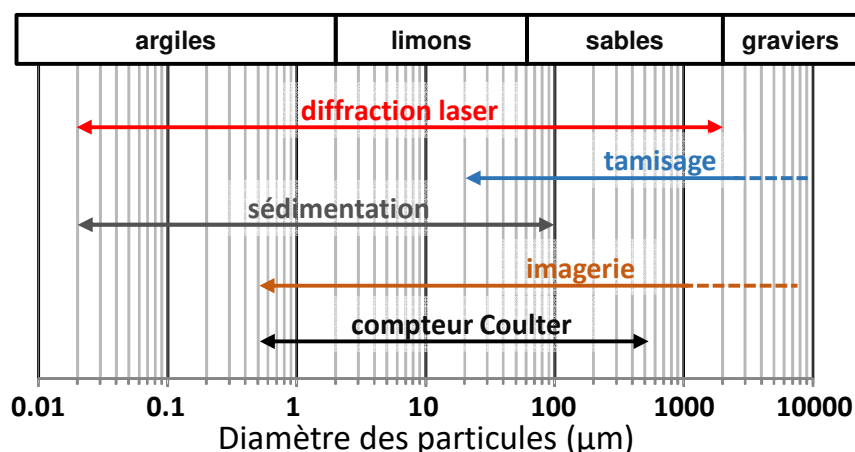


Figure 1 : Gammes de tailles de particules mesurables avec différentes techniques granulométriques (adapté de Delanghe et al., 2016).

Il existe de nombreuses techniques d'analyse de distribution granulométrique (tamisage sec ou humide, sédimentation, compteur Coulter, analyse d'image ou diffraction laser) qui couvrent une gamme de tailles de particules qui leur sont propres (Figure 1). La granulométrie par diffraction est bien adaptée pour mesurer la taille des particules des argiles, des limons et des sables fins (d'une dizaine de nanomètres à quelques centaines de micromètres) mais des précautions sont nécessaires (Delanghe et al., 2018 ; Lepage et al., 2019 ; Masson et al., 2021). Il est important de noter qu'il semble plus difficile d'analyser des échantillons sableux comme peuvent être les sédiments (avec des particules de diamètres supérieurs à 300 μm) : problèmes de surreprésentation des grosses particules et d'homogénéisation des échantillons avant analyse (Masson et al., 2021). Malgré cela, cette technique est aujourd'hui largement utilisée pour caractériser les distributions granulométriques des échantillons de sédiment et de MES car elle permet de fournir des distributions granulométriques à haute définition (généralement plus de 50 classes de tailles de particules), avec un temps d'analyse plus réduit que d'autres techniques et une apparente facilité de mise en place.

Si l'analyse par granulométrie laser est maintenant bien maîtrisée, la question de la bancarisation des données granulométriques se pose. En effet, il semble compliqué de bancariser l'ensemble des données granulométriques car les données issues d'appareils différents n'ont pas le même format (en particulier un nombre de classes de diamètres différents ; par exemple 116 classes de 0.04 à 2000 µm pour l'appareil Beckman Coulter LS 13 320, 100 classes de 0.04 à 2500 µm pour le Cilas 1190 et 100 classes de 0.02 à 2000 µm pour le Malvern Mastersizer 2000). Ensuite, les paramètres classiques de description des distributions granulométriques, comme le diamètre médian (D50) ou la fraction de particules de l'échantillon correspondant à une classe de taille donnée (par ex. <20 µm : code Sandre 6228), ne permettent pas toujours de décrire de façon optimale les distributions. Ceci est particulièrement vrai lorsqu'une distribution multimodale est observée comme c'est généralement le cas pour les MES des cours d'eau (Masson et al., 2018 ; Lepage et al., 2023). Ainsi, une méthode de décomposition des distributions granulométriques a été proposée par Launay (2014) et appliquée à des distributions granulométriques d'échantillons de MES collectées dans le Rhône (Masson et al., 2018). Cette méthode permet de décomposer les distributions granulométriques en plusieurs sous-populations (généralement 2 à 4 pour des MES de cours d'eau) homogènes de particules qui suivent chacune une distribution lognormale. Chaque sous-population est caractérisée par son mode (μ), sa dispersion (σ) et sa proportion (λ), et admet pour densité de probabilité :

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \times \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Cette méthode de décomposition peut s'appliquer sur des jeux de distributions granulométriques comportant des nombres différents de classes, réparties sur des gammes de tailles différentes. Finalement, en partant du principe que les distributions granulométriques peuvent être modélisées par un nombre n de sous-populations gaussiennes (typiquement 2 à 5 sous-populations), la bancarisation des données granulométriques nécessiterait la bancarisation de seulement $3n$ paramètres, indépendamment du type de granulomètre laser utilisé.

Il existe très peu de littérature sur l'utilisation d'outils pour la décomposition des distributions granulométriques des échantillons de sédiment ou de MES (e.g. Clark, 1976 ; Sheridan et al., 1987, Bah et al., 2009, Chapuis et al., 2014, Launay, 2014 ; Delanghe et al., 2017). L'objectif de ce rapport est d'estimer le potentiel de ces traitements de données granulométriques pour une bancarisation homogène des données sans perte de l'information globale en évaluant leur qualité et leur faisabilité. Dans cette étude, deux outils mathématiques de décomposition proposés par les logiciels R et Matlab pour décomposer les distributions granulométriques d'échantillons de MES ont été identifiés et testés. Ces deux outils mathématiques ont été testés sur des distributions granulométriques issues de i) simulations numériques, ii) d'expérimentations de mélanges de particules en laboratoire et iii) de MES de rivières collectées dans le cadre de l'Observatoire des Sédiments du Rhône¹ (OSR).

¹ <https://observatoire-sediments-rhone.fr>

2. METHODES DE DECOMPOSITION DES DISTRIBUTIONS GRANULOMETRIQUES

2.1 FONCTION *NORMALMIXEM* DU LOGICIEL R

La première méthode de décomposition utilise la fonction *normalmixEM* du package *R Mixtools* (Benaglia et al., 2009) permettant d'identifier les paramètres des distributions normales dans un échantillon composé d'un mélange de populations normalement distribuées. L'utilisation de cet outil mathématique sur les résultats bruts (obtenus via les granulomètres laser) pour la décomposition des distributions granulométriques d'échantillons de sédiments ou MES est décrite dans la thèse de Marina Launay (2014) et dans Masson et al., 2018. Le calcul se fait de façon itérative par la méthode du maximum de vraisemblance avec l'algorithme *EM* (pour Espérance-Maximisation) proposé par Dempster et al. (1977). Cette outil prend comme données d'entrée la distribution granulométrique mesurée (classes de diamètres des particules et proportions volumiques pour chaque classe) ainsi que le nombre envisagé de sous-populations. Pour tenir compte de l'hypothèse que les sous-populations suivent des lois de distribution log-normale, la fonction logarithme a été appliquée aux diamètres des particules avant de les utiliser. La fonction crée des points supplémentaires par interpolation ; le nombre d'interpolation a été choisi à 10000. Pour chaque sous-population, la fonction *normalmixEM* retourne son mode (μ ; en μm), son écart-type (σ ; en μm) et sa proportion (λ ; en %). Plusieurs étapes supplémentaires non prévues dans la fonction sont à appliquer pour « nettoyer » les résultats (Launay, 2014) :

- élimination des sous-populations aberrantes : $\sigma > 0.5$ (loi normale qui s'étale sur deux ordres de grandeur) ou $\lambda < 5\%$ (sous-populations négligeable) ;
- regroupement des sous-populations au sein des 4 classes de diamètres médians généralement identifiées dans les MES des échantillons du Rhône et de ses affluents (1-10 μm , 10-27 μm , 27-80 μm et >80 μm ; Launay, 2014). Le mode résultant de ce regroupement correspond à la moyenne des modes. L'écart-type est calculée comme la moyenne des écarts-types pondérée par leurs proportions respectives ;
- normalisation des proportions de chaque sous-population pour que leur somme atteigne 100%.

L'étape d'interpolation engendre de la variabilité dans les résultats produits. Aussi, pour chaque distribution granulométrique à traiter, 100 décompositions sont réalisées. Les valeurs médianes des paramètres de chaque sous-population sont alors calculées et bancarisées. Un exemple de décomposition est donnée dans la Figure 2. Pour ces deux échantillons, 3 sous-populations ont été identifiées (modes à 4, 15 et 39 μm pour le premier échantillon et à 4, 13, et 35 μm pour le second échantillon).

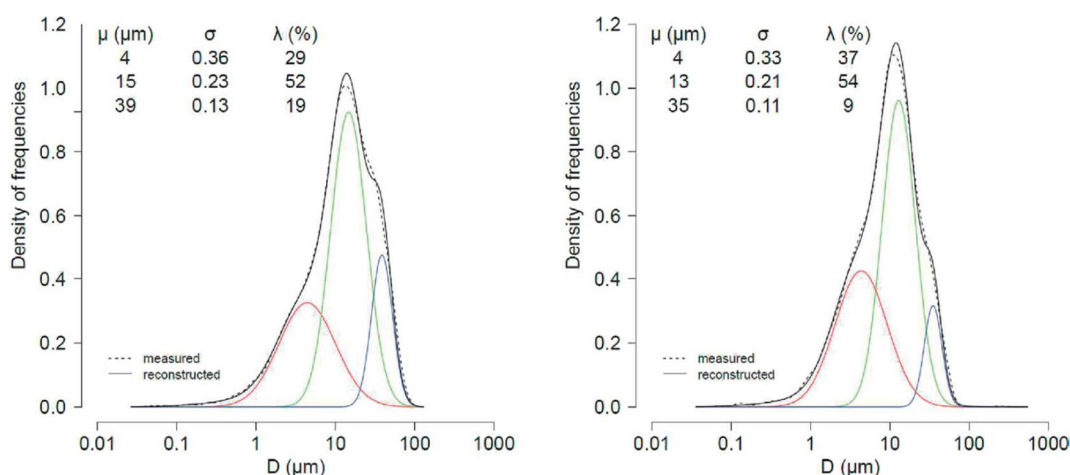


Figure 2 : Exemple de décomposition (fonction *normalmix* du package *R Mixtools*) en 3 sous-populations gaussiennes (courbes rouge, verte et bleue) de distribution granulométrique mesurées (ligne pointillée) pour deux échantillons de MES échantillonnés dans le Rhône à la station de Jons (d'après Masson et al., 2018).

2.2 FONCTION PEAKFIT DU LOGICIEL MATLAB

La deuxième méthode de décomposition utilise la fonction *peakfit* du logiciel *Matlab R2020a* (O'Haver, 2023). Cet outil mathématique a été utilisé par Boukra et al. (2023) pour la décomposition des distributions en taille de molécules organiques analysées par chromatographie d'exclusion stérique. La même procédure a été utilisée pour les distributions granulométriques dans cette étude. Cette méthode effectue un ajustement itératif sans contrainte par la méthode des moindres carrés pour décomposer un signal complexe à pics superposés en ses éléments constitutifs. Comme pour la méthode précédente, cette fonction prend comme entrée la distribution granulométrique (avec prise en compte du logarithme des diamètres des particules) et le nombre envisagé de sous-populations. De plus, cette fonction permet de paramétrer certaines étapes du calcul. Ainsi, le type de sous-population a été défini comme Gaussien, aucune correction de ligne de base n'est réalisée, et aucune valeur de départ (mode, dispersion ou proportion) n'est utilisée. Contrairement à la méthode précédente, aucune variabilité dans le résultat n'est observée. Les valeurs des modes (μ ; en μm), écart-types (σ ; en μm) et proportions (λ ; en %) sont identiques si on répète la décomposition sur une même distribution granulométrique. Cependant, il est possible de renouveler plusieurs fois la décomposition avec des valeurs de départ légèrement différentes, ce qui peut améliorer les résultats. Dans ce cas, la meilleure décomposition est retenue (celle dont la distribution reconstituée sera la plus proche de la distribution mesurée). Cette dernière option a été utilisée pour certains échantillons dont la décomposition ne donnait pas de résultats satisfaisants.

Un exemple de décomposition est donné dans la Figure 3 pour un échantillon de MES prélevé dans la Saône en 2021 par piège à particules (Figure 2). Pour cet échantillon, le meilleur score de confiance (52%) a été trouvé en prédéfinissant 4 sous-populations gaussiennes. Ce score est considéré comme convenable et la décomposition est considérée comme validée malgré l'absence de modélisation du petit pic granulométrique observé vers $0.4 \mu\text{m}$. Au final, 3 sous populations ont été retenues pour cet échantillon avec des modes à 3, 15 et $51 \mu\text{m}$ pour des proportions de 14.4 %, 65.4 % et 20.2 % respectivement.

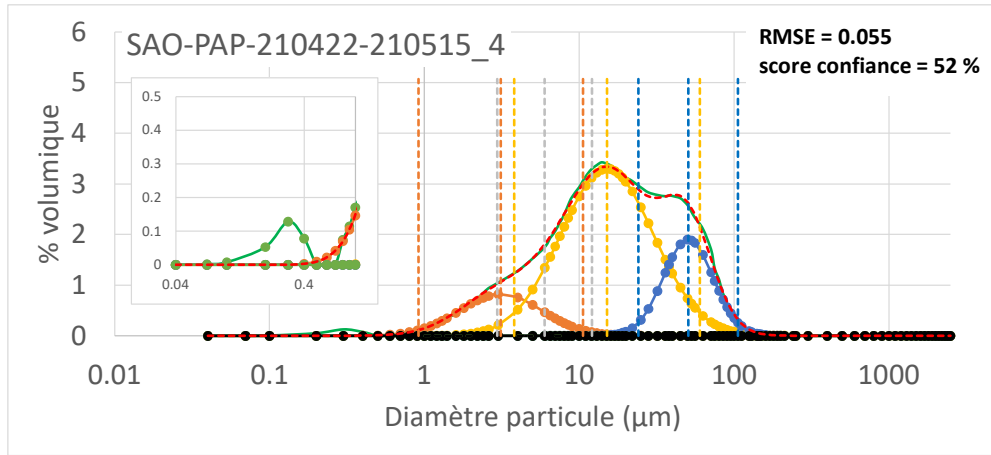


Figure 3 : Exemple de décomposition (fonction peakfit du logiciel Matlab R2020a) en 4 sous-populations gaussiennes (courbes rouge, verte, bleue et noire) d'une distribution granulométrique mesurée (ligne verte) à partir d'un échantillon de MES collecté dans la Saône à la station de Lyon. La distribution recomposée à partir des 4 sous-populations (courbe pointillée rouge) retrace bien l'allure générale de la distribution mesurée excepté pour le pic observé vers 0.4 μm .

3. EVALUATION DES OUTILS DE DECOMPOSITION

3.1 A PARTIR DE DISTRIBUTIONS GENEREES NUMERIQUEMENT

Une première étape de validation des 2 outils de décomposition granulométrique a consisté à appliquer ces outils sur des distributions générées par le mélange de sous-populations gaussiennes bien définies. Pour cela, une série de 4 distributions a été générée en additionnant 3 à 4 sous-populations. Le Tableau 1 reprend les valeurs définies pour chaque sous-population créée dans les 3 mélanges. Dans cette étude, deux approches sont mises en œuvre pour évaluer la qualité des modélisations.

La première approche se base sur la valeur de l'écart quadratique moyen (RMSE) entre la distribution modélisée (somme des distributions des sous-populations gaussiennes modélisées) et la distribution mesurée. La valeur de RMSE, qui représente l'erreur de prévision, est calculée comme suit :

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

où $\hat{y}_i$ et y_i représentent les valeurs prédites mesurées pour chaque classe de particule, et n le nombre de mesures (pour le granulomètre Cilas, $n=100$). Plus la valeur de RMSE se rapproche de 0, et meilleure est la qualité de la modélisation.

La deuxième approche se base sur l'adéquation entre les paramètres (modes, écarte-types et/ou proportions) extraits des distributions modélisées et les paramètres théoriques (i.e. pour ce test, les paramètres qui ont servis à générer les différentes distributions).

Tableau 1 : Modes (μ ; en μm), écart-types (σ ; en μm) et proportions (λ ; en %) de chaque sous-population pour les échantillons de référence (Référence) dont les distributions granulométriques ont été générées numériquement (à partir de la somme de 3 ou 4 sous-populations gaussiennes) et les décomposition à partir du modèle *R* et du modèle *MatLab*. Pour chaque valeur estimée de mode, écart-type et proportion, l'erreur est calculée à partir de la valeur de référence (les erreurs comprises entre -10% et 10% sont indiquées en bleu et celles qui dépassent ces bornes sont indiquées en rouge).

		Référence			modèle R						modèle MatLab					
		μ (μm)	σ (μm)	λ (%)	μ (μm)	erreur μ (%)	σ (μm)	erreur σ (%)	λ (%)	erreur λ (%)	μ (μm)	erreur μ (%)	σ (μm)	erreur σ (%)	λ (%)	erreur λ (%)
mélange 1	pop 1	3.0	0.20	0.30	2.8	-6.7%	0.20	0.0%	0.30	-0.3%	3.0	0.0%	0.20	2.3%	0.30	0.0%
	pop 2	30.0	0.10	0.40	28.7	-4.3%	0.10	0.0%	0.41	3.5%	30.0	0.0%	0.11	5.5%	0.41	3.2%
	pop 3	200	0.10	0.30	189	-5.8%	0.09	-10.0%	0.29	-4.7%	200	0.0%	0.11	5.5%	0.31	3.2%
mélange 2	pop 1	3.0	0.20	0.20	3	-3.3%	0.21	5.0%	0.20	2.0%	2.9	-2.2%	0.20	0.2%	0.19	-3.8%
	pop 2	15.0	0.25	0.60	14.2	-5.3%	0.24	-4.0%	0.58	-2.8%	15.1	0.9%	0.27	6.1%	0.62	3.6%
	pop 3	40.0	0.15	0.20	37.8	-5.5%	0.15	0.0%	0.21	6.5%	40.5	1.1%	0.15	0.7%	0.19	-5.1%
mélange 3	pop 1	4.2	0.21	0.14	4.4	4.8%	0.22	4.8%	0.19	37.9%	4.2	0.6%	0.21	1.5%	0.13	-4.9%
	pop 2	17.2	0.30	0.65	15.5	-9.9%	0.25	-16.7%	0.50	-22.6%	17.3	0.3%	0.31	2.3%	0.65	0.4%
	pop 3	50.6	0.19	0.21	44.7	-11.7%	0.21	10.5%	0.30	44.8%	51.2	1.1%	0.19	-0.3%	0.20	-5.5%
mélange 4	pop 1	2.5	0.22	0.10	2.5	0.0%	0.23	4.5%	0.11	5.0%	2.1	-14.2%	0.16	-27.3%	0.06	-35.4%
	pop 2	12.6	0.27	0.55	10.8	-14.3%	0.23	-14.8%	0.30	-45.5%	13.4	6.6%	0.32	19.7%	0.66	19.1%
	pop 3	42.0	0.22	0.25	21.6	-48.6%	0.27	22.7%	0.32	28.4%	46.3	10.3%	0.21	-6.1%	0.17	-30.8%
	pop 4	112	0.24	0.10	63	-44.2%	0.31	29.2%	0.27	173.0%	113	0.9%	0.24	1.7%	0.10	0.2%

Le premier mélange est réalisé à partir de 3 sous-populations gaussiennes bien définies, i.e. sous-populations qui ne se recoupent pas. Les décompositions en 3 sous-populations gaussiennes par les modèles *R* et *MatLab* sont conformes à la distribution de référence (Figure 4) avec des erreurs sur les modes, écart-types et proportions comprises entre -10% et 10% (Tableau 1). Le calcul de l'écart quadratique moyen (RMSE) permet de dire que la décomposition réalisée avec le modèle *MatLab* (RMSE = 0.017) est meilleure que celle réalisée avec le modèle *R* (RMSE = 0.074).

Deux autres distributions ont été créés à partir de la somme de 3 sous-populations gaussiennes qui se recouvrent. Ces distributions sont similaires à des distributions granulométriques de MES de rivières comme celles généralement mesurées pour les échantillons du Rhône et de ses principaux affluents. Il est à noter que le deuxième mélange présente des modes plus distincts que le troisième mélange (Figure 4). Les résultats de décomposition du deuxième mélange sont très bons pour les deux modèles avec des erreurs sur les paramètres modélisés comprises entre -10% et 10% (Tableau 1), et des RMSE de 0.016 pour le modèle *R* et de 0.008 pour le modèle *MatLab*. Pour le troisième mélange, les erreurs sur les paramètres modélisés avec *R* sont relativement élevées, principalement pour ce qui concerne les proportions des 3 sous-populations. Cependant, le résultat final, i.e. la distribution estimée par la somme des 3 sous-populations sommées, est très proche de la référence comme l'indique la valeur de RMSE de 0.016. Le modèle *MatLab* permet une meilleure approximation des 3 sous-populations (erreurs comprises entre -10% et 10% pour tous les paramètres) avec une valeur de RMSE deux fois plus faible (RMSE = 0.008) qu'avec le modèle *R* (Figure 4).

Enfin, le dernier mélange composé de la somme de 4 sous-populations pourrait représenter un échantillon de MES légèrement sableux. Les distributions modélisées à partir des 4 sous-populations décomposées sont proches des distributions de référence avec une valeur de RMSE de 0.024 pour le modèle *R* et une valeur de RMSE de 0.008 pour le modèle *MatLab* (Figure 4). Cependant, les caractéristiques (mode, écart-type et proportion) des sous-populations générées

par les deux modèles s'éloignent de la réalité, en particulier pour le modèle *R* (Tableau 1). Nous avons ici l'exemple d'une modélisation identifiée comme réussie d'après le critère RMSE (bonne adéquation entre la distribution granulométrique mesurée et celle modélisée) mais pour laquelle les sous-populations générées sont complètement différentes (en terme de mode, écart-type et proportion) des sous-population théoriques ayant servies à générer la distribution granulométrique théorique (i.e. bonne réponse pour de mauvaises raisons).

Ces premiers tests utilisant des distributions générées numériquement à partir de sous-populations gaussiennes qui reflètent des distributions granulométriques typiques de MES de rivières, montrent que les décompositions avec les modèles *R* et *MatLab* permettent d'approcher convenablement les distributions « réelles ». Le modèle *MatLab* semble donner de meilleurs résultats que le modèle *R* (meilleures valeurs de RMSE). Dans un objectif de modéliser les distributions granulométriques à partir des paramètres caractéristiques des-populations gaussiennes modélisées, ceci est très prometteur. Il est à noter cependant que certaines sous-populations sont mal modélisées mais la résultante finale (distribution recréée à partir des sous-populations modélisées) est très proche de la distribution de référence. Dans le futur, ces cas devront être détectés pour savoir si un développement/optimisation des modèles peut améliorer les résultats, ou si ces distributions doivent être bancarisées avec un code qualité particulier (e.g. code qualité « douteux »). L'utilisation des paramètres des sous-populations gaussiennes pour d'autres objectifs (par exemple pour estimer les proportions de particules grossières et réaliser des corrections granulométriques) pourrait engendrer des erreurs d'interprétation si les données qualité d'adéquation comme le RMSE sont bons mais que les gaussiennes des sous-populations ne sont pas représentatives de la réalité.

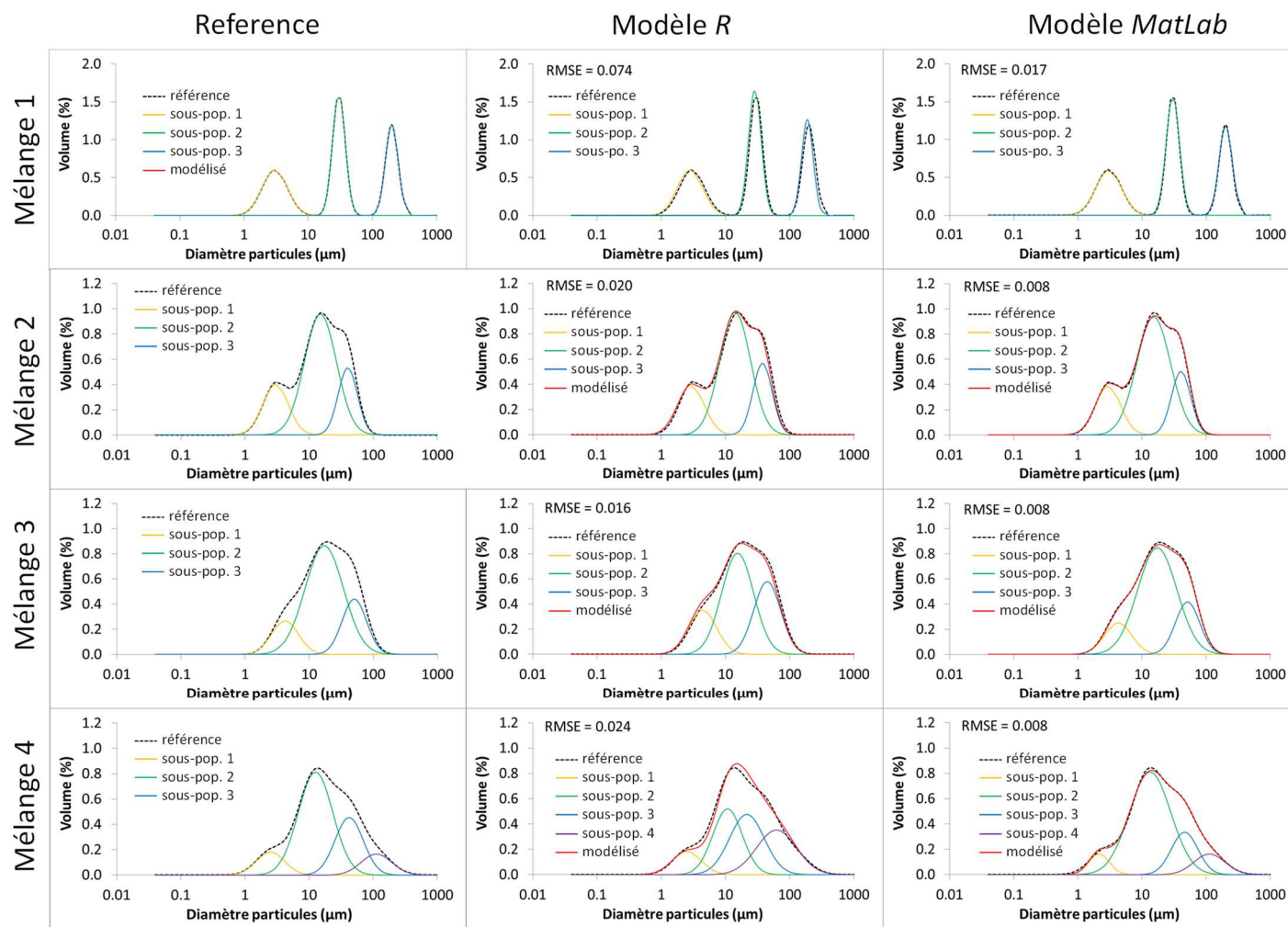


Figure 4 : Distribution volumique (en %) des échantillons de référence (courbes pointillées noires ; générés à partir de la somme de 3 ou 4 sous-populations gaussiennes) et des distributions reconstituées (courbes rouges) à partir des sous-populations gaussiennes (courbes orange, vertes e, bleues et mauves) issues des décompositions des échantillons de référence avec les modèles R et MatLab.

3.2 TESTS DE MELANGES DE PARTICULES EN LABORATOIRE

3.2.1 DESCRIPTION DE L'EXPERIMENTATION

La deuxième étape de validation des 2 outils consiste à décomposer des distributions d'échantillons composites de MES dont les distributions des sous-populations sont contrôlées. Pour cela, un échantillon de MES collecté dans l'Isère été séché et tamisé à travers 2 tamis de 50 et 100 μm . Ainsi, 3 fractions ont été créées :

- La fraction A : fraction < 50 μm , représentative des argiles jusqu'aux limons moyen;
- La fraction B : fraction comprise entre 50 et 100 μm , représentative des limons grossiers à sables très fins ;
- La fraction C : fraction > 100 μm , représentative des sables fins.

Les distributions des trois fractions A, B et C ont été mesurées par diffraction laser (3 réplicats de mesure pour les fractions A et C, et 4 réplicats pour la fraction B). Ensuite, les fractions ont été mélangées deux à deux dans des proportions connues (0:100, 25:75, 50:50, 75:25 et 100:0) pour créer de nouveaux échantillons. Une fois les distributions granulométriques des échantillons de mélange décomposées, les caractéristiques des sous-populations gaussiennes sont comparées aux valeurs théoriques de mélange. Les caractéristiques théoriques des sous-populations d'un mélange sont calculées (mode μ , écart-type σ et proportion λ) à partir des caractéristiques des sous-populations des 2 fractions (μ_1 , σ_1 , λ_1 et μ_2 , σ_2 , λ_2) servant au mélange et des proportions des deux fractions mélangées (p_1 et p_2). Quand deux sous-populations se superposent, i.e. quand les modes sont dans les mêmes gammes 1-10 μm , 10-27 μm , 27-80 μm ou >80 μm (Launay, 2014), elles sont regroupées sous une seule sous-population ayant comme caractéristiques :

- $\mu = p_1.\mu_1 + p_2.\mu_2$
- $\sigma = (p_1.\sigma_1.\lambda_1 + p_2.\sigma_2.\lambda_2) / (p_1.\lambda_1 + p_2.\lambda_2)$
- $\lambda = p_1.\lambda_1 + p_2.\lambda_2$

L'analyse granulométrique des fractions servant à réaliser les différents mélanges de particules (Figure 5) montre que les fractions B et C contiennent majoritairement la fraction granulométrique ciblée, mais qu'elles contiennent aussi des particules de diamètres plus petits (en particulier des limons fins de diamètres proches de 15 μm). Ainsi, il n'a pas été possible de créer trois fractions unimodales, ce qui rend l'analyse des données plus complexe. La première étape a donc été de décomposer les distributions granulométriques des 3 échantillons tamisés.

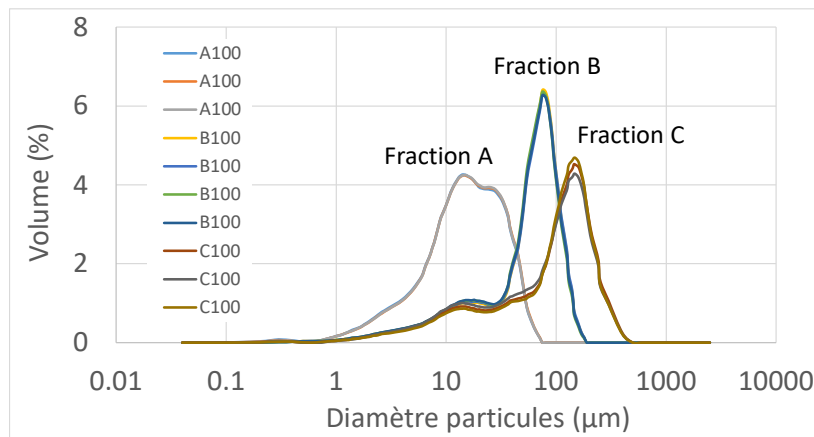


Figure 5 : Distribution volumique (en %) des fractions de l'échantillon tamisés à > 100 µm (fraction C) entre 50 et 100 µm (fraction B) et < 50 µm (fraction A). Trois sous-échantillons des fractions A et B ont été analysés, et 4 réplicats ont été réalisés pour la fraction B.

Le Tableau 2 présente une synthèse des caractéristiques des sous-populations gaussiennes (sous-forme de moyenne et d'écart type sur les 3 ou réplicats mesurés) des fractions A, B et C obtenues avec les modèle *R* et *MatLab*. Les décompositions de la fraction A sont meilleures que celles des fractions B et C (Annexe 1). En effet, de meilleures valeurs de RMSE sont obtenues pour la fraction A que pour les fractions B et C. De plus, la qualité des décompositions est meilleure avec le modèle *Matlab* (RMSE variant de 0.040 à 0.084) qu'avec le modèle *R* (RMSE variant de 0.107 à 0.197).

Tableau 2 : Synthèse des caractéristiques des sous-populations gaussiennes (moyenne et écarts-types des modes μ , écart-type σ , et proportions λ) pour les 3 échantillons tamisés (les fractions A, B et C) obtenues avec les modèles *Matlab* et *R*. Les valeurs en rouge indiquent des écarts relatifs supérieurs à 10% et donc variables

Modèle	Fractions	$\mu 1$ (µm)	$\mu 2$ (µm)	$\mu 3$ (µm)	$\mu 4$ (µm)	$\sigma 1$ (µm)	$\sigma 2$ (µm)	$\sigma 3$ (µm)	$\sigma 4$ (µm)	$\lambda 1$ (%)	$\lambda 2$ (%)	$\lambda 3$ (%)	$\lambda 4$ (%)
MatLab	A	4.3 ± 0.3	14.6 ± 0.1	34.0 ± 0.1	-	0.339 ± 0.016	0.230 ± 0.003	0.139 ± 0.001	-	0.20 ± 0.02	0.58 ± 0.02	0.22 ± 0.01	-
MatLab	B	4.0 ± 1.4	20.4 ± 1.3	75.5 ± 0.7	-	0.377 ± 0.098	0.355 ± 0.018	0.149 ± 0.001	-	0.07 ± 0.03	0.27 ± 0.02	0.66 ± 0.01	-
MatLab	C	3.0 ± 0.9	12.5 ± 2.2	53.9 ± 7.9	147 ± 2	0.143 ± 0.051	0.263 ± 0.095	0.291 ± 0.100	0.17 ± 0.01	0.03 ± 0.03	0.17 ± 0.09	0.24 ± 0.10	0.564 ± 0.03
R	A	4.8 ± 0.1	14.5 ± 0.1	33.1 ± 0.2	-	0.303 ± 0.012	0.210 ± 0.000	0.120 ± 0.000	-	0.22 ± 0.00	0.57 ± 0.00	0.21 ± 0.00	-
R	B	7.9 ± 1.3	29.0 ± 7.1	75.1 ± 2.4	-	0.590 ± 0.018	0.298 ± 0.035	0.133 ± 0.005	-	0.16 ± 0.04	0.26 ± 0.02	0.58 ± 0.05	-
R	C	14.7 ± 1.1	28.4 ± 9.1	89.9 ± 12.1	159 ± 9	0.550 ± 0.052	0.570 ± 0.089	0.183 ± 0.023	0.13 ± 0.01	0.20 ± 0.01	0.20 ± 0.02	0.242 ± 0.03	0.564 ± 0.03

3.2.2 MODELISATION DES MELANGES

Une fois les distributions granulométriques des mélanges analysés par diffraction laser puis décomposés à l'aide des deux modèles, les caractéristiques (modes et proportions) des sous-populations identifiées dans les mélanges sont comparées aux valeurs théoriques. Les valeurs théoriques ont été estimées à partir des décompositions des fractions A, B et C et des équations de mélanges présentées dans le paragraphe précédent.

Pour le mélange entre les fractions A et B, nous pouvons remarquer que le modèle *Matlab* modélise bien les modes des sous-populations 2 (mode compris entre 10 et 27 µm) et 3 (mode entre 27 et 80 µm), alors qu'il sous-estime celui des sous-populations 1 (mode entre 1 et 10 µm) dans tous les mélanges (Figure 6). Les proportions sont systématiquement sous-estimées pour les sous-populations 1 et 3 au détriment d'une surestimation des proportions de la sous-population 2.

Le modèle *R* permet une meilleure modélisation que le modèle *Matlab* des modes des 3 sous-populations (Figure 7). Concernant la modélisation des proportions, le modèle *R* ne fait pas mieux que le modèle *Matlab* en surestimant les proportions des deux premières sous-populations et en sous-estimant celles de la sous-population 3.

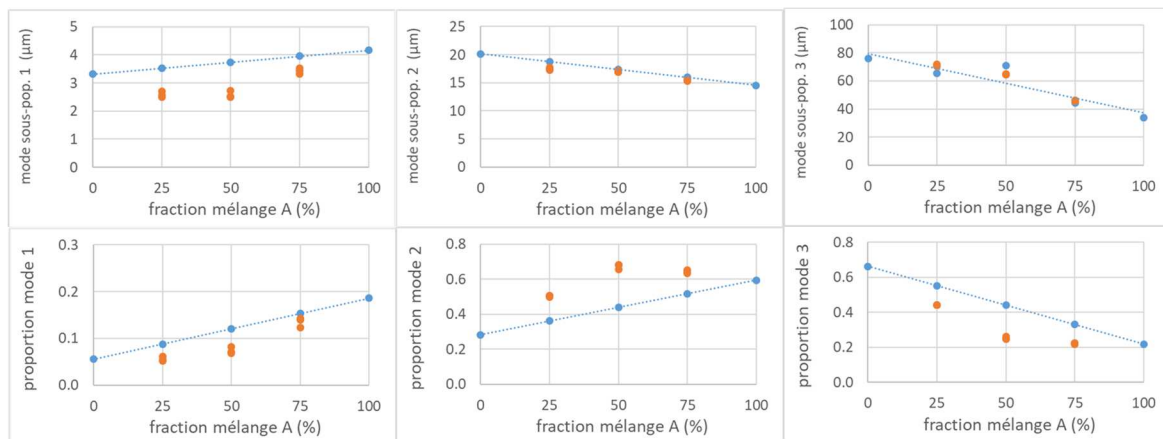


Figure 6 : Résultats des décompositions (pour les modes et les proportions associées) réalisées avec le modèle Matlab sur les fractions A et B et les mélanges théoriques de ces deux fractions (points bleus) et sur les mélanges entre les fractions A et B avec les proportions 25:75,50:50et75:25(points orange).

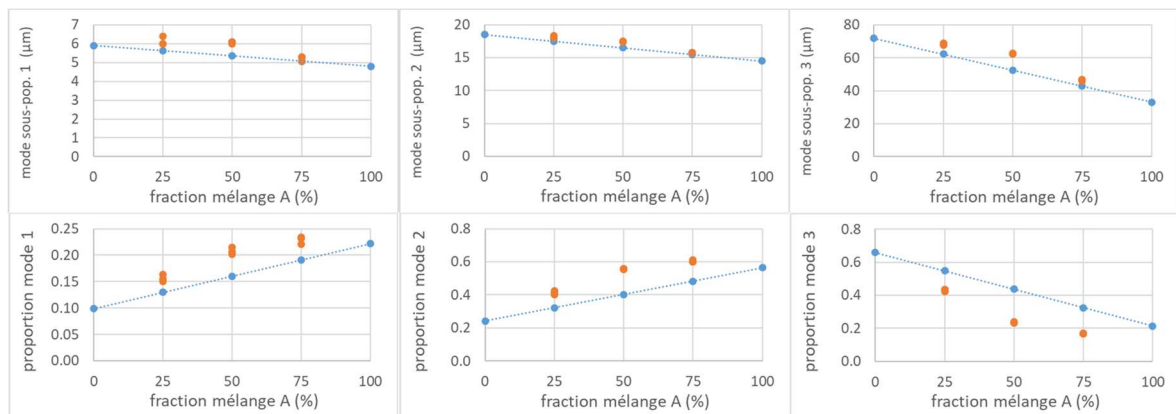


Figure 7 : Résultats des décompositions (pour les modes et les proportions associées) réalisées avec le modèle R sur les fractions A et B et les mélanges théoriques de ces deux fractions (points bleus) et sur les mélanges entre les fractions A et B avec les proportions 25:75,50:50et75:25(points orange).

Ces premiers résultats montrent qu'il est difficile de modéliser le mélange de particules. Des imprécisions sur la position des modes peuvent créer des biais sur les proportions obtenues. Mais surtout, il semble compliqué de modéliser correctement la somme de deux sous-populations très proches en terme de mode. Les équations utilisées restent théoriques et les limites de cette expérimentation sont atteintes du fait que les fractions mélangées ne sont pas unimodales.

Les mauvaises qualités de décomposition des fractions B et C (en particulier avec le modèle R), et surtout des mélanges entre les fractions A et C et les fractions B et C, rendent les résultats de ces tests de mélange difficiles à interpréter. Les résultats de ces mélanges ne sont donc pas montrés ni discutés dans ce rapport. Ceci suggère que des distributions granulométriques mélangeant de nombreuses sous-populations sont difficiles à décomposer que ce soit avec le modèle Matlab ou le modèle R. Des modes trop larges ou mal positionnés sont identifiés par les modèles. Le modèle Matlab semble toutefois légèrement plus robuste que le modèle R (RMSE généralement plus faible).

Aussi, ces expérimentations de mélange montrent qu'il est difficile de tamiser des particules pour obtenir des seuls de coupures nets. Le mélange de fractions de particules en se basant sur les

masses de particules ne permet pas non plus d'avoir de bonnes précisions sur les proportions initialement souhaitées. Des densités différentes entre les classes de particules pourraient expliquer ces écarts entre proportions théoriques et proportions mesurées et modélisées.

3.3 DISTRIBUTIONS GRANULOMETRIQUES D'ECHANTILLONS DE MES REELS (STATION DE JONS, RHONE)

3.3.1 DESCRIPTION DES DONNEES

Les données de granulométrie utilisées dans ce test sont issues de l'Observatoire des Sédiments du Rhône (Thollet et al., 2021). La station de prélèvement de MES est celle de Jons qui se situe sur le Rhône en amont de l'agglomération lyonnaise. Les échantillons de MES sont prélevés par deux techniques différentes : par pompage puis centrifugation (CFC, prélèvement ponctuel) et par piège à particules (PAP, prélèvement intégratif). Les deux types de prélèvement de MES sont intéressants pour cette étude car il a été montré que les MES prélevées par PAP sont généralement plus grossières que celle prélevées par CFC (Masson et al., 2018). Ainsi, les distributions granulométriques de l'ensemble de ces échantillons couvrent des classes de tailles de particules dépassant la centaine de micromètre. Le même jeu de données traité par Masson et al. (2018) est utilisé dans cette étude ; il comprend 138 échantillons de MES prélevés à Jons entre août 2012 et juin 2016 (77 échantillons issus de pompage/centrifugation et 61 échantillons issus de pièges à particules). Les décompositions granulométriques de ces échantillons avaient déjà été réalisées par le modèle *R*. Une nouvelle décomposition granulométrique a donc été réalisée avec le modèle *Matlab* dans cette étude. Les résultats issus de ces deux modèles sont comparés en terme de modes et de proportions pour les 138 échantillons.

3.3.2 COMPARAISON DES DECOMPOSITIONS DISTRIBUTIONS GRANULOMETRIQUES DES PARTICULES PRELEVEES DANS LE RHONE A JONS

Les résultats des décompositions granulométriques (modes et proportions) sont donnés en annexe 2. Des exemples de décomposition à l'aide des deux modèles sont reportées dans la Figure 9. Le premier échantillon (JON-CMO-130108) fait apparaître 3 sous-populations avec les deux modèles. Le deuxième échantillon (JON-CMO-130205) fait apparaître seulement 2 sous-populations avec les deux modèles, la sous-population la plus fine correspondant à une classe de particules allant de 2 à 6 μm n'a pas été détectée.

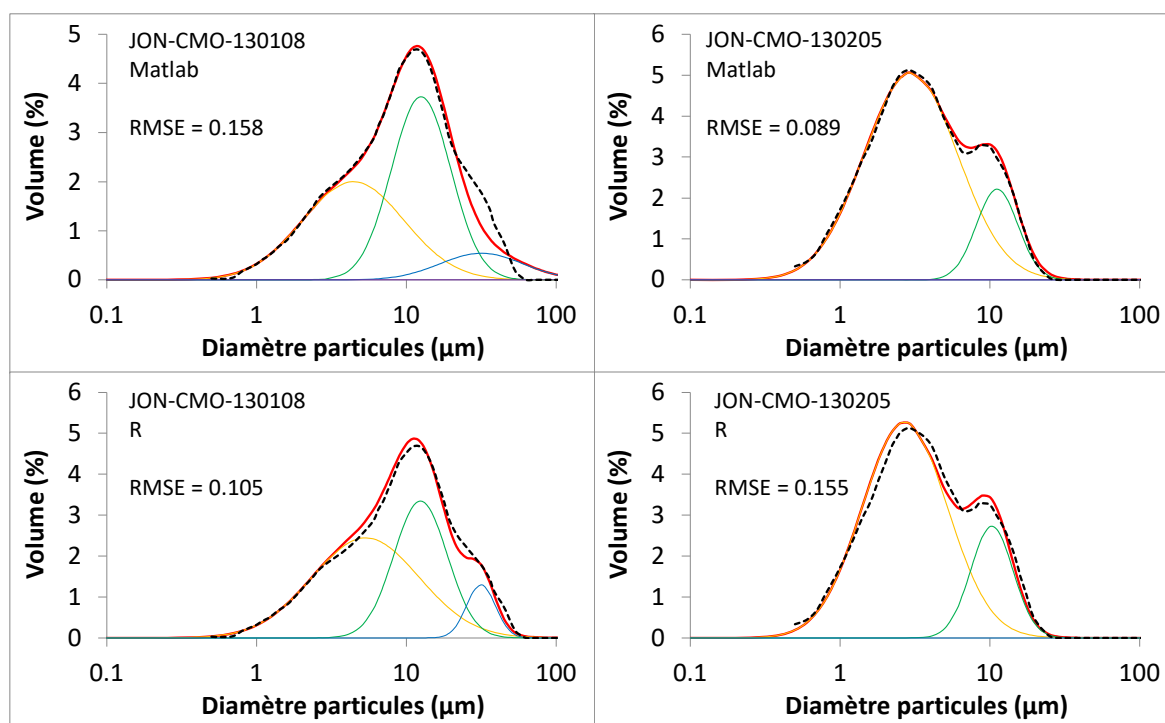


Figure 8 : Résultats des décompositions granulométriques (modes et proportions) obtenus à partir des modèles *R* (en abscisse) et *Matlab* (en ordonnées) pour les échantillons de MES prélevés à la station de Jons (Rhône) pour la période août 2012 à juin 2016. Généralement, 3 sous-populations ont été retrouvées pour chaque distribution granulométrique.

Sur 77 échantillons de CFC, 26 échantillons avaient été décomposés par le modèle *R* à l'aide de deux sous-populations de particules ayant des modes compris entre 2.0 et 4.5 μm pour la plus fine et compris entre 9.4 et 19.0 μm pour la deuxième. Avec le modèle *Matlab*, seulement 10 de ces 26 échantillons sont décomposés avec seulement 2 sous-populations ayant des modes similaires aux sous-populations modélisées avec *R*, compris entre 2.9 et 3.7 μm pour la plus fine et compris entre 10.9 et 13.2 μm pour la seconde. Une 3^{ème} sous-population apparaît pour les autres échantillons dont le mode est compris entre 25.9 et 36.8 μm . Les proportions de cette 3^{ème} sous-population sont faibles et comprises entre 5 et 12%. Seulement 3 échantillons de CFC sont décomposés avec 2 sous-populations par le modèle *Matlab* alors que le modèle *R* en fait apparaître 3. Pour les échantillons prélevés par PAP, les deux modèles décomposent les distributions granulométriques avec 3 sous-populations dont les modes sont compris :

- pour le modèle *R* : (1) entre 2.6 et 6.0 μm , (2) entre 12.5 et 20.7 μm et (3) entre 33.7 et 74.1 μm ;
- pour le modèle *Matlab* : (1) entre 1.6 et 6.4 μm , (2) entre 13.4 et 20.9 μm et (3) entre 33.8 et 77.4 μm .

Ainsi les modes des 3 sous-populations sont comparables entre les deux modèles.

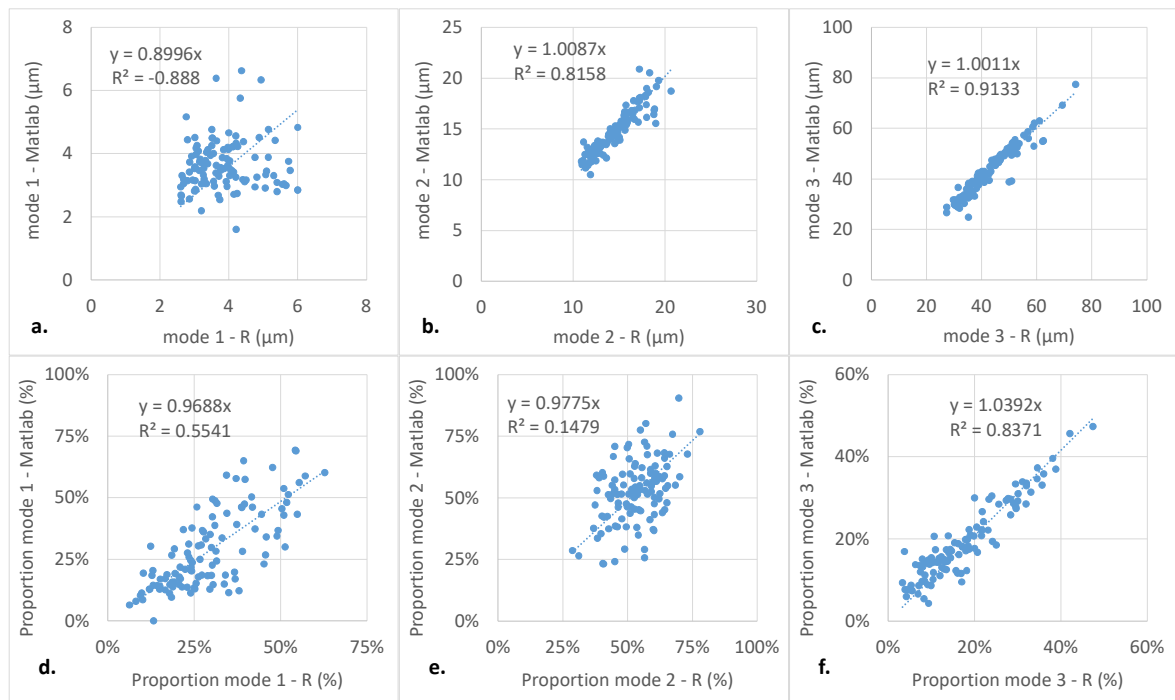


Figure 9 : Résultats des décompositions granulométriques (modes et proportions) obtenus à partir des modèles R (en abscisse) et Matlab (en ordonnées) pour les échantillons de MES prélevées à la station de Jons (Rhône) pour la période août 2012 à juin 2016. Généralement, 3 sous-populations ont été retrouvées pour chaque distribution granulométrique.

En comparant les décompositions issues des deux modèles pour chaque échantillon, il est possible de mettre en évidence certaines différences et similitudes entre les modèles R et Matlab.

Les caractéristiques de la première sous-population sont généralement différentes en terme de modes et de proportions (Figure 9a et d). Toutefois, comme indiqué précédemment, les modes de cette première sous-population sont tous compris entre 1.6 et 6.6 μm et cette sous-population de particules proche des argiles est systématiquement observée. Les proportions sont très variables entre les deux modèles mais une corrélation entre les deux modèles est observée (Figure 9d).

Les modes obtenus pour la deuxième sous-population sont similaires entre les modèles pour l'ensemble des échantillons comme le montre la bonne corrélation observée entre les données avec une pente proche de 1 (Figure 9b). Par contre, les proportions ne sont pas si comparables entre les deux modèles. L'explication la plus probable est que les différences observées sur le mode de la première sous-population influent sur les proportions des 2 sous-populations. Si le mode de la première sous-population est sous-évaluée, cela aura tendance à diminuer la proportion de la première sous-population et, au contraire, à augmenter la proportion de la seconde sous-population. Ainsi une corrélation significative est observée entre les écarts relatifs entre les deux modèles sur les modes de la première sous-population (obtenus avec les deux modèles) et les proportions de la deuxième sous-population (Figure 10).

Enfin, les modes et les proportions obtenus pour la troisième sous-population sont comparables entre les 2 modèles. Cette sous-population de particules correspondant aux limons grossiers semble bien être identifiée par les deux modèles. Pour cette sous-population de particules, les confusions avec la sous-population de particules de taille inférieure semblent être limitées.

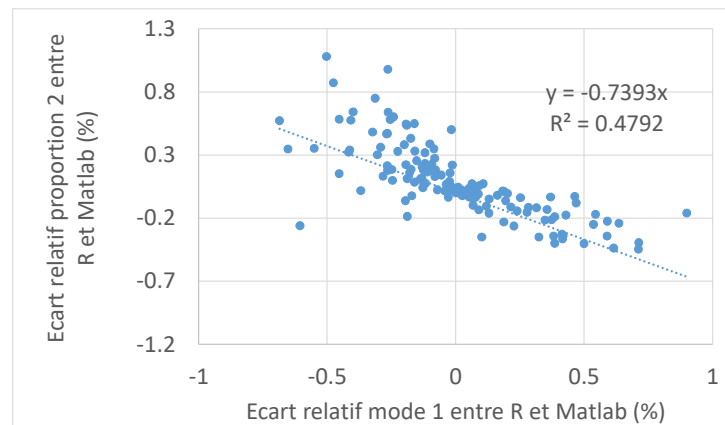


Figure 10 : Ecart relatif entre les deux modèles sur les modes de la première sous-population (obtenus avec les deux modèles) en fonction des écarts relatifs entre les deux modèles sur les proportions de la deuxième sous-population.

La qualité de la modélisation peut être approchée par le calcul du RMSE. Malheureusement, seules les valeurs de RMSE des échantillons de CFC modélisés par le modèle *R* avaient été calculées dans l'étude de Masson et al. (2018). Ainsi la comparaison de la qualité des modélisations entre les deux modèles n'a été réalisée dans cette étude que sur les échantillons collectés par CFC. Cette comparaison se base tout de même sur 76 décompositions granulométriques aux résultats variés étant donné que 2 ou 3 sous-populations granulométriques sont identifiées au total.

Les valeurs de RMSE sont comprises entre 0.068 et 0.985 avec le modèle *R* (valeur médiane de 0.402) et sont comprises entre 0.048 et 0.889 avec le modèle *Matlab* (valeur médiane de 0.172). La dispersion des valeurs de RMSE (Figure 11) ainsi que la valeur médiane 2 fois plus faible pour le modèle *Matlab* que le modèle *R* montrent clairement que le modèle *Matlab* donne des résultats plus fiables que le modèle *R*. Cette évaluation sur des distributions granulométriques de MES de rivière confirme les précédents résultats obtenus sur des mélanges de particules en laboratoire.

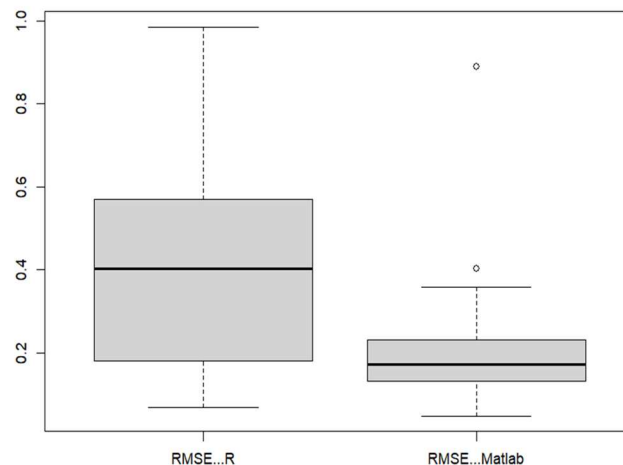


Figure 11 : Boîtes à moustache des valeurs des écarts quadratiques moyens (RMSE) obtenues par les modèles *R* (à gauche) et *Matlab* (à droite) sur les décompositions des 76 distributions granulométriques des échantillons de MES collectés par centrifugation/pompage (CFC) à la station de Jons sur le Rhône.

En plus de qualité des données, d'autres critères sont à prendre en compte dans la comparaison de l'utilisation des deux modèles : la simplicité ou complexité de l'utilisation des scripts de chaque modèle et le temps de calcul.

Pour chaque modèle, l'automatisation de l'utilisation des fonctions de décomposition, nécessite l'édition d'un script permettant:

- d'appliquer la fonction de décomposition sur les données de granulométrie (conversion de certains paramètre comme les diamètre des particules, paramétrisation du nombre de sous-population à rechercher, paramétrisation du nombre de répétition) ;
- d'automatiser la décomposition pour plusieurs échantillons ;
- d'écarter les résultats aberrants ;
- de mettre en forme les résultats dans un format adapté.

En l'état actuel, les scripts utilisés dans cette étude n'ont pas été optimisés. Néanmoins, il apparaît que le script édité pour le modèle *Matlab* est bien plus simple à faire fonctionner que celui du modèle *R*. En effet, le script *R* nécessite la mise en place d'une étape préliminaire pour prédéfinir les données initiales de décomposition (modes des différentes sous-population) et aussi la répétition systématique des décompositions (actuellement 100 répétitions) avec élimination des valeurs extrêmes et aberrantes. La fonction *peakfit* de *Matlab* est plus simple à utiliser car elle comprend des options permettant de réaliser des répétitions et de choisir le résultat le plus performant. Cette valeur a été fixée aussi à 100 dans cette étude pour être comparable au modèle *R*.

La plus grande complexité du modèle *R* engendre un temps de calcul beaucoup plus important que celui du modèle *Matlab*. Pour une décomposition avec 3 sous-populations, le modèle *R* travaille pendant près de 2 minutes alors que le modèle *Matlab* travaille pendant 5 à 6 secondes. Ce temps de travail est près de 2 fois plus long dans le cas d'une décomposition avec 4 sous-populations. Ainsi, le modèle *R* devient difficile à appliquer pour de grandes séries d'échantillons supérieures à quelques centaines d'échantillons et dans lesquels 4 sous-populations sont recherchées.

4. MISE EN PRATIQUE DE LA DECOMPOSITIONS SUR UN GROS JEU DE DONNEES ET QUALITE DES DONNEES BANCARISEES

4.1.1 DESCRIPTION DU JEU DE DONNEES ET DU MODELE UTILISE

Les données de granulométries utilisées dans cette mise en pratique sont issues de la base de données OSR (Thollet et al., 2021). Les données granulométriques collectées sont celles de MES du Rhône (aux stations de Jons, Andancette et Arles) et de ses principaux affluents (Ain, Ardèche, Arve, Bourbre, Drôme, Durance, Fier, Gardon, Gier, Isère, Saône). Au total, 797 échantillons de MES ont été analysés par granulométrie laser.

Etant donné que le temps de travail du modèle *R* dépasse la minute par échantillon, et qu'il faut multiplier les décompositions (en changeant le nombre de sous-population, typiquement entre 2 et 4) pour estimer le bon nombre de sous-population, il n'a pas été possible (ou plutôt raisonnable) de décomposer les 797 échantillons avec le modèle *R*. C'est donc le modèle *Matlab* qui a été utilisé dans cet exercice pratique. Le modèle *Matlab* a été appliqué sur chaque échantillon en faisant varier le nombre de sous-population de 2 à 5 puis en sélectionnant la modélisation avec la valeur de RMSE la plus petite.

4.1.2 DECOMPOSITION GRANULOMETRIQUE DES MES DU RHONE ET DE SES PRINCIPAUX AFFLUENTS

Les décompositions par le modèle *Matlab* ont révélé l'existence de 2 à 4 sous-populations granulométriques dans l'ensemble des échantillons. Ainsi, 4 groupes de sous-populations ont été définis en fonction de la répartition des modes (Figure 12) :

- Sous-population 1 identifiée dans 740 échantillons : modes compris entre 0.9 et 11.9 μm avec une médiane à 3.4 μm . Cette sous-population correspond principalement à des argiles et des limons très fins à fins d'après la classification de Bloot et Pye (2012).
- Sous-population 2 identifiée dans 795 échantillons: modes compris entre 9.4 et 34.5 μm avec une médiane à 15.9 μm . Cette sous-population correspond principalement à des limons moyens et grossiers.
- Sous-population 3 identifiée dans 729 échantillons: modes compris entre 31.0 et 83.0 μm avec une médiane à 50.5 μm . Cette sous-population correspond principalement à des limons grossier et sables très fins.
- Sous-population 4 identifiée dans 20 échantillons: modes compris entre 70.9 et 297 μm avec une médiane à 101 μm . Cette sous-population correspond principalement à des sables très fins à fins.

Les intervalles des modes se recoupent entre sous-populations car pour quelques échantillons, deux sous-populations avec des modes relativement proches ont été identifiées. Ceci conduit à élargir vers le bas ou vers le haut la gamme des modes d'une sous-population.

Un seul échantillon est composé d'une seule sous-population granulométrique. La grande majorité des échantillons sont composés de 3 sous-populations, et 105 échantillons n'en présentent que 2.

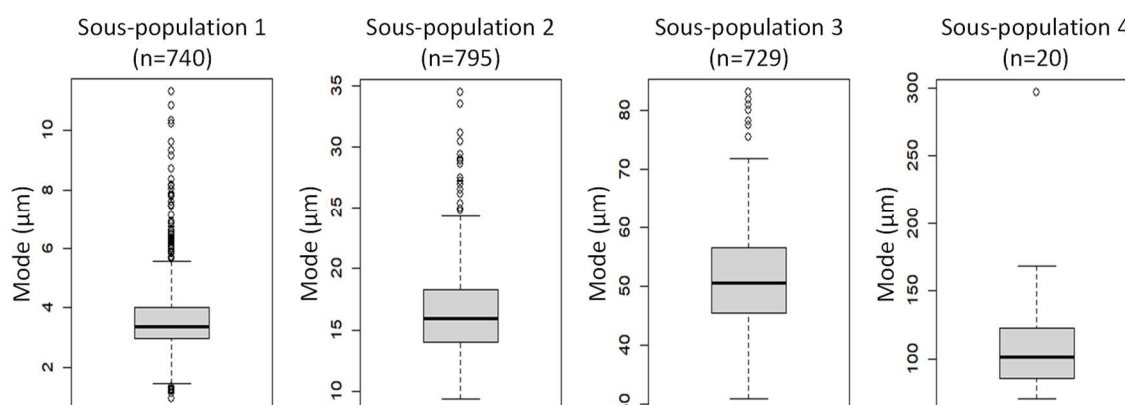


Figure 12 : Boîtes à moustaches des modes des sous-populations granulométriques identifiées sur les 797 échantillons de MES du Rhône et de ses principaux affluents.

De façon globale, la sous-population 2 est celle qui possède les proportions les plus élevées (médiane à 58%) devant la sous-population 3 (médiane à 25%) et la sous-population 1 (médiane à 15%). Quand elle est présente (dans seulement 20 échantillons), les proportions de la sous-population 4 sont élevées avec une médiane à 47%, ce qui traduit un échantillon sableux (Figure 13).

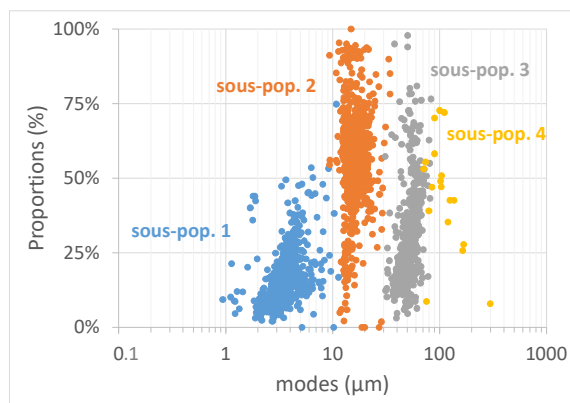


Figure 13 : Proportions en fonction des modes des 4 sous-populations identifiées dans les échantillons de MES du Rhône et de ses principaux affluents.

Plus de la moitié des décompositions granulométriques (503 sur 797 ; soit 63% des échantillons) ont été réalisées avec une valeur de RMSE inférieure à 0.06. Pour ces décompositions, la qualité de la modélisation peut être qualifiée de très bonne comme le montrent les deux exemples de décompositions caractérisées par des RMSE de 0.4 et 0.6 (Figure 14a, b). Les valeurs de RMSE de la grande majorité des échantillons (686 sur 797 ; soit 86% des échantillons) sont inférieures à 0.1 ce qui traduit une bonne qualité des décompositions réalisées comme le montre la Figure 14c. Finalement, seulement 18 échantillons (soit 2% des échantillons) ne sont pas correctement décomposés avec des valeurs de RMSE supérieures à 0.2 (Figure 14d). Pour ces échantillons, il semblerait qu'une des sous-populations ne soit pas identifiée, ce qui entraîne un élargissement trop important des distributions des sous-populations modélisées. Ceci pourrait être sûrement amélioré en multipliant les décompositions ou en fixant des valeurs initiales (principalement sur les modes) lors de la décomposition par le modèle *Matlab*. De telles optimisations n'ont pas été explorées dans cette étude. Une étude d'optimisation devra être conduite avant d'utiliser ce modèle pour décomposer les distributions granulométriques afin de maximiser la qualité des modélisations.

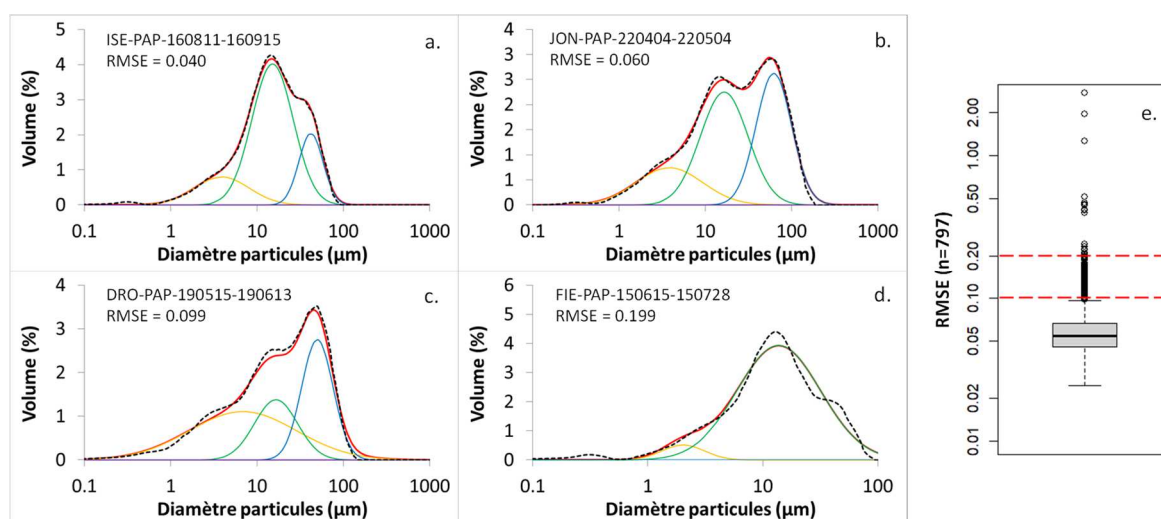


Figure 14 : Exemples de décompositions granulométriques caractérisées par différents niveaux de valeurs de RMSE (a. RMSE = 0.040 ; b. RMSE = 0.060 ; c. RMSE = 0.099 ; d. RMSE = 0.199) et distribution des valeurs de RMSE (sous forme de boîte à moustache avec échelle logarithmique) pour les décompositions granulométriques des 797 échantillons de MES collectées dans le Rhône et ses principaux affluents.

Finalement, sans réelle optimisation du modèle *Matlab*, la grande majorité des décompositions peuvent être validées directement (celles pour un RMSE < 0.1) ; environ 12% des échantillons (ceux avec un RMSE compris entre 0.1 et 0.2) pourraient être décomposés de nouveau en augmentant le nombre d'itération afin de voir si la qualité s'améliore. Ces valeurs de RMSE (à 0.1 et 0.2) pourraient constituer des valeurs seuils pour valider les décompositions. Bien entendu, d'autres critères de validation pourraient être mis en place : le nombre de pics définis dans une même plage granulométrique pourrait être redéfini pour limiter les pics trop proches en terme de mode ; un critère sur l'étalement des pics (par exemple un rapport largeur sur hauteur supérieur à 1) qui permettrait d'invalider des distributions trop étalées et peu réalistes ; ou un critère sur l'écart maximal autorisé entre valeur modélisée et mesurée pour chaque classe qui permettrait de détecter les distributions avec une bonne allure globale (valeur de RMSE correcte) mais des sous-populations aberrantes (e.g. mauvais modes).

4.1.3 BANCARISATION DES DONNEES

En se basant sur les données granulométriques de cours d'eau comme le Rhône et ses affluents, il est raisonnable de penser que 4 sous-populations de particules permettent de décrire les distributions granulométriques des MES. Ceci devrait être vérifié pour les petits cours d'eau pour lesquels le transport de fraction sableuse peut être plus important. Ainsi, 12 paramètres sont nécessaires à la bancarisation des modes, écart-types et proportions modélisés. Il serait aussi intéressant de rajouter un critère de qualité comme la valeur de RMSE issue de la décomposition afin d'apprécier la qualité de la modélisation granulométrique. Ces données sont bancarisées indépendamment du type de granulomètre laser utilisé pour réaliser les mesures, i.e. indépendamment du nombre de classes de taille de particules rendus par le granulomètre.

A partir de ces 13 paramètres, il est possible de modéliser une courbe granulométrique et de calculer certaines caractéristiques granulométriques de l'échantillon comme les différentes fractions de particules <20 µm (code Sandre 6228), comprises entre 20 et 63 µm (code Sandre 3054), entre 63 et 150 µm (code Sandre 7042), entre 150 et 200 µm (code Sandre 7043) et > 200 µm (code Sandre 7044) (e.g. Amalric et al., 2016) ; ou les déciles d10, d50 (grain médian) et d90. Cela nécessite d'utiliser la loi log-normale avec les paramètres descriptifs bancarisés (mode, écart-type et proportion) pour chaque sous-population. La création d'outil mathématique (type feuille excel, script R...) permettrait d'automatiser les calculs. Bien entendu, il est possible de bancariser ces informations (5 fractions et 3 déciles) calculés à partir des distributions granulométriques mesurées en plus des paramètres de décomposition. Mais cela impliquerait la bancarisation de 21 paramètres ce qui semble beaucoup. Il serait intéressant de comparer les valeurs des 5 fractions et des 3 déciles obtenus à partir des distributions mesurées et modélisées afin de s'assurer que ces valeurs sont similaires dans les deux cas (travail non réalisé dans cette étude).

Au-delà de la bancarisation, les résultats de décomposition peuvent être exploités pour décrire les distributions granulométriques. Par exemple, Masson et al. (2018) ont comparé les modes et proportions des 3 sous-populations de MES collectées par pièges à particules et pompage centrifugation. Cette comparaison a permis de mettre en évidence un biais granulométrique induit par les pièges à particules (Figure 15a) qui ont tendance à collecter des particules plus grossières (limons grossiers) au détriment des particules les plus fines (limons très fins). Ce biais est d'autant plus important que le débit de la rivière est élevé (Figure 15b).

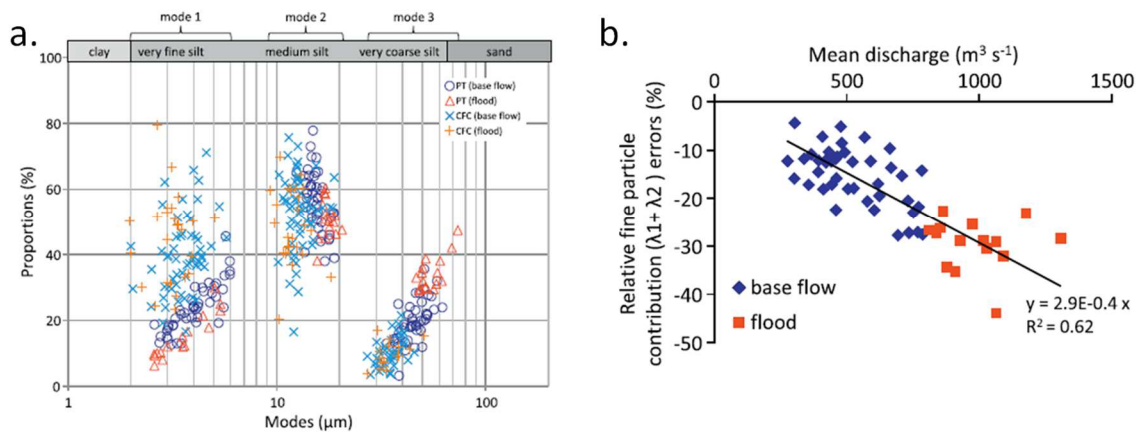


Figure 15 : à gauche (a), distribution des modes et des proportions (sous forme de boîtes à moustaches) associées des 3 sous-populations identifiées dans les distributions granulométriques des MES collectées dans le Rhône par pièges à particules et par centrifugation ; à droite (b) contributions cumulées des deux sous-populations les plus fines en fonction du débit de la rivière. D'après Masson et al., 2018.

En reprenant les données des distributions granulométriques des MES du Rhône et de ses principaux affluents, il est possible d'utiliser les modes et les proportions associées pour comparer la granulométrie des cours d'eau. Par exemple, le grain médian d_{50} est plus élevé et plus variable dans le Rhône à Jons ($d_{50} = 18.8 \pm 7.6$) qu'à l'exutoire de la Saône à Lyon ($d_{50} = 12.1 \pm 2.5$). Une observation grossière de la distribution des modes et proportions des 3 sous-populations identifiées dans les distributions granulométriques montre que les diamètres médians des limons moyens (sous-population 2) et grossiers (sous-population 2) sont légèrement plus élevés dans le Rhône que dans la Saône. De plus, le Rhône transporte des proportions de limons grossiers plus important que la Saône, alors que la tendance inverse est observée pour les limons très fins (sous-population 1). Ces observations n'auraient pas été possibles avec les seules données généralement bancarisées comme le grain médian ou les fractions.

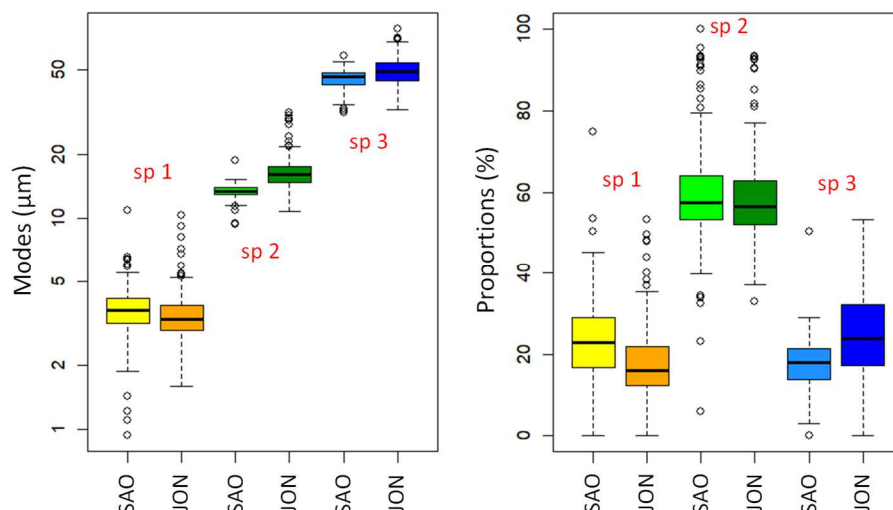


Figure 16 : Distribution des modes et des proportions (sous forme de boîtes à moustaches) associées des 3 sous-populations granulométriques principales identifiées dans les échantillons de MES collectés sur le Rhône (à la station de Jons) et sur la Saône (à la station de Lyon). L'échelle des modes est exprimée en logarithme pour permettre une meilleure représentation graphique des 3 sous-populations.

5. CONCLUSIONS ET PERSPECTIVES

Deux fonctions permettant la décomposition des distributions granulométriques ont été identifiées et testées dans cette étude : *normalmixEM* du package *R Mixtools* et *peakfit* du logiciel *Matlab*.

Les deux modèles ont été testés sur des mélanges numériques puis des mélanges de particules en laboratoire. Les premiers résultats montrent la capacité de ces modèles pour décomposer les distributions granulométriques et identifier les différentes sous-populations gaussiennes. Les critères de qualité comme le RMSE peuvent être utilisés pour valider les décompositions. Il conviendra tout de même de mettre en place d'autres critères afin d'invalider les décompositions modélisant très bien les mesures mais avec des sous-populations aberrantes (par exemple sous-population ayant des modes trop proches, ou un étalement trop important). Les tests de mélanges de particules en laboratoire ont été confrontés à certaines limites comme le fait d'arriver à créer un échantillon ayant une distribution unimodale, ou la transformation de deux sous-populations proches en terme de mode en une seule sous-population en se basant sur les lois empiriques d'addition de loi normales.

Les deux modèles ont été testés sur des distributions granulométriques réelles. Les résultats sont globalement similaires : similitudes dans le nombre de sous-populations identifiées et dans les classes de taille des particules des différentes sous-populations. Quelques différences sont à noter pour les modes des populations les plus fines qui engendrent un peu de variabilité dans les proportions des 2 classes les plus fines (sans doute un partage entre les deux fractions difficiles à faire). Les contrôles qualité RMSE sont clairement meilleurs pour le modèle *Matlab*. En particulier le temps de décomposition avec le modèle *Matlab* est bien plus rapide qu'avec le modèle *R*. En effet, les deux algorithmes sont différents et induisent une gestion différente avec la création d'un script d'utilisation plus complexe dans le cas du modèle *R*.

Ce temps important de décomposition *R* nous a contraints à ne pas envisager de décomposer les distributions granulométriques d'un grand jeu de données (797 échantillons) avec le modèle *R*. Les résultats obtenus avec le modèle *Matlab* sont cohérents avec généralement de bonnes valeurs de RMSE et une bonne répartition des modes dans 4 groupes de diamètres de particules identifiés au préalable (des limons très fins à fins ; limons moyens et grossiers ; limons grossier et sables très fins ; sables très fins à fins).

Un dernier aspect reste à tester avant une bancarisation possible des données générées : l'effet du nombre de classes de mesure sur les caractéristiques des sous-populations identifiées. Pour cela, il faudrait pouvoir décomposer des distributions granulométriques avec plusieurs types de granulomètres laser (par exemple Cilas 1190 avec 100 classes de particules, Beckman Coulter LS 13 320 avec 116 classes ou Sequoia LISST streamside avec 32 classes

Cette étude montre qu'il est possible de bancariser de façon fiable les distributions granulométriques d'échantillons de MES (voire de sédiments) en un jeu de 13 paramètres (triplets de modes, écart-types et proportions pour 4 sous-population, et un critère de qualité comme le RMSE), et cela indépendamment du nombre de classes mesurées par les granulomètres laser. Cependant, la décomposition des distributions granulométriques n'est pas si triviale et nécessite une certaine expertise dans la mise en place des algorithmes mathématiques et dans la validation des décompositions. L'algorithme *peakfit* du logiciel *Matlab* semble clairement le mieux adapté pour réaliser ces décompositions qui sont plus fiables, plus faciles à mettre en œuvre avec un temps de calcul plus court que l'algorithme *normalmixEM* du package *R Mixtools*. Il est tout de même important de noter que l'utilisation du logiciel *R* est libre alors que celle de *Matlab* est payante. Quelques améliorations pourraient être apportées pour aider à la validation des résultats et éviter d'avoir les bonnes réponses (bonne adéquation du spectre global indiquée par bonne valeur de RMSE) pour des mauvaises raisons (allures aberrantes ou non naturelles des sous-populations modélisées) : critères d'étalement des pics ou de recouvrement trop important des sous-populations. Ces critères de qualité pourraient être importés dans un script permettant une

décomposition et surtout une validation automatique des résultats. Mais ces critères doivent être rigoureusement évalués en amont.

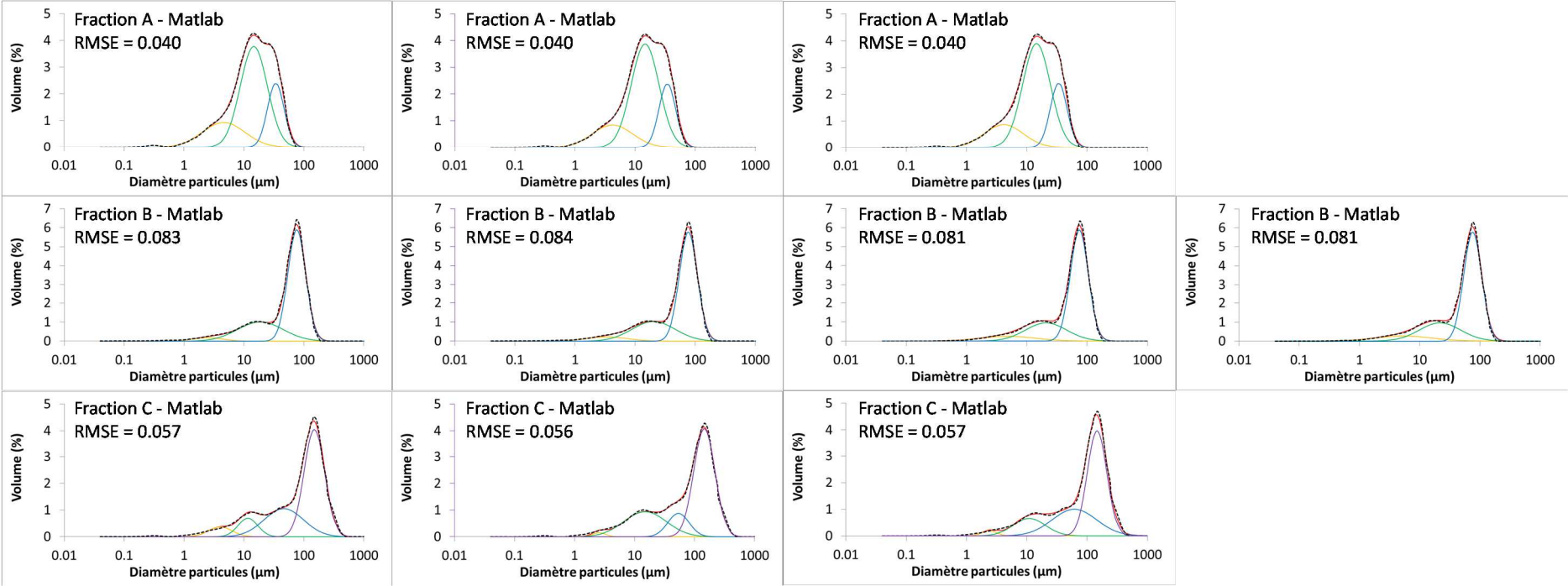
A ce stade, de tels outils semblent plus adaptés pour une utilisation de recherche scientifique que pour la bancarisation de données dans le cadre de la surveillance de la qualité des sédiments. Il semble tout de même intéressant d'optimiser cet outil de décomposition pour permettre une utilisation plus générale des données produites : par exemple pour la caractérisation plus fine des MES et des sédiments, l'utilisation à des fins de normalisation, ou l'étude du lien entre les contaminants et les différentes fractions granulométrique des MES.

Bibliographie

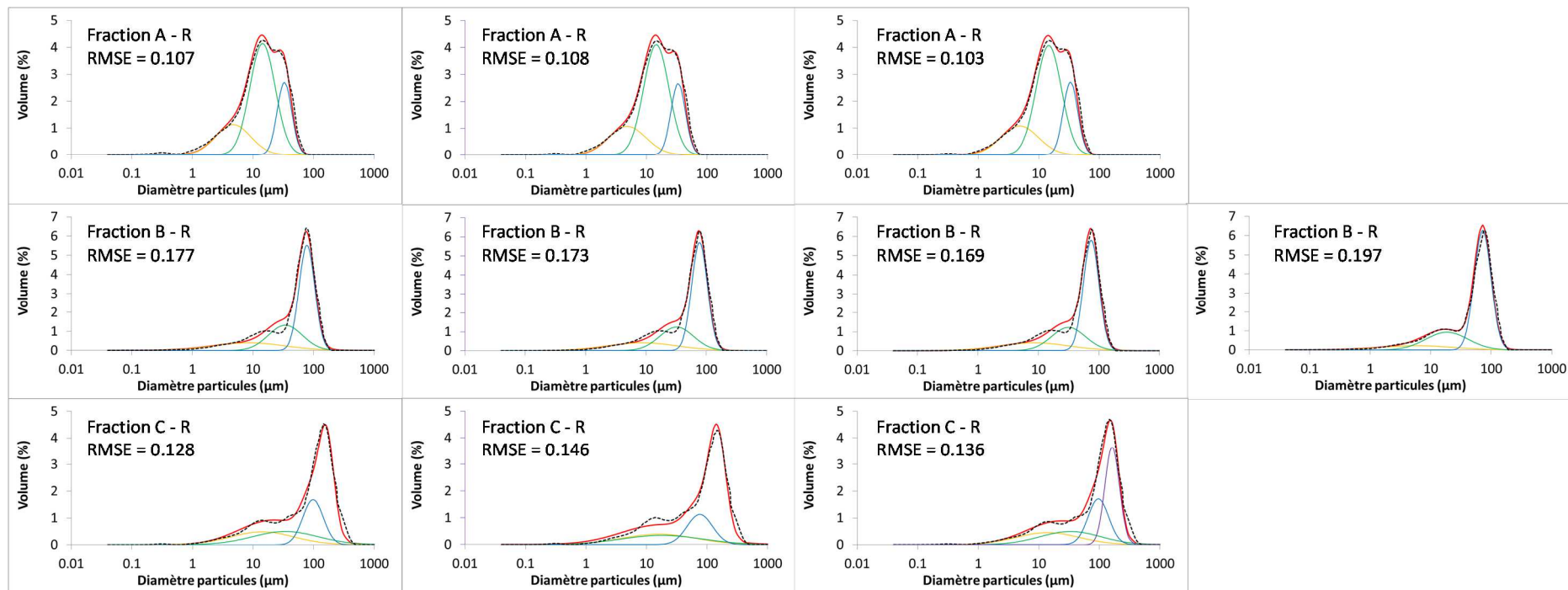
- AFNOR (2009a). ISO 13320 : Analyse granulométrique – Méthode par diffraction laser – Principes généraux. 51 p.
- AFNOR (2009b). NF EN ISO 5667 : Qualité de l'eau - Échantillonnage - Partie 15 : Lignes directrices pour la conservation et le traitement des échantillons de boues et de sédiments. 28 p.
- Amalric L., Lestremeau F., Strub M.-P., Lionard E., Lardy-Fontan S. (2016). Opérations d'analyse physico-chimique des eaux et des sédiments en milieu continental dans le cadre des programmes de surveillance DCE - Recommandations techniques. Rapport AQUAREF 2016, 34.
- Bah A.R., Kravchuk O., Kirchhof G. (2009). Fitting Performance of Particle-size Distribution Models on Data Derived by Conventional and Laser Diffraction Techniques. *Soil Science Society of America Journal* 73(4), 1101-1107.
- Benaglia T., Chauveau D., Hunter D.R., Young D.S. (2009). Mixtools: an R package for analyzing finite mixture models. *J. Stat. Softw.* 32, 1–29.
- Blott S.J., Pye K. (2001). GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surface Processes and Landforms* 26, 1237-1248.
- Chapuis, R.P., Dallaire V., Saucier A. (2014). Getting information from modal decomposition of grain size distribution curves. *Geotechnical Testing Journal* 37, 282-295.
- Cilas (2011) Analyseurs granulométriques – Manuel utilisateur. Troisième édition, version française, 114 p.
- Clark M.W. (1976). Some Methods for Statistical Analysis of Multimodal Distributions and Their Application to Grain-Size Data. *Mathematical Geology* 8, 267-282.
- Coquery M., Lionard E., Yari A. (2017). Opérations d'échantillonnage de sédiments en milieu continental (cours d'eau et plan d'eau) dans le cadre des programmes de surveillance DCE - Recommandations techniques – Edition 2017. Rapport AQUAREF, 22 p.
- Dabrin A., Bégorre C., Bretier M., Dugué V., Masson M., Le Bescond C., Le Coz J., Coquery M. (2021). Reactivity of particulate element concentrations: apportionment assessment of suspended particulate matter sources in the Upper Rhône River, France. *Journal of Soils and Sediments* 21, 1256-1274.
- Delanghe D., Lepage H., Masson M., Le Bescond C. (2018). Synthèse sur les techniques granulométriques - Méthodologie et inter-comparaison des analyses granulométriques. Rapport OSR4 2015-2017, 66 p.
- Dramais G. (2020). Observation et modélisation des flux de sable dans les grands cours d'eau. Mécanique des fluides. Thèse de doctorat de l'université de Lyon. 294 p.
- Duinker J.C. (1986). The role of small, low-density particles on the partition of selected PCB congeners between water and suspended matter (North-Sea area). *Netherlands Journal of Sea Research* 20:229-238.

- Gruat A., Coquery M., Le Coz J., Thollet F., Lagouy M., Dabrin A., Masson M., et al. (2020). Rapport sur le fonctionnement du réseau OSR d'observation des flux de matières en suspension et de contaminants particuliers (OSR5–année2020). Rapport de recherche ZABR, Lyon. 36 p.
- Juvigné E. (1982). L'utilisation rationnelle de l'eau oxygénée pour la destruction de matières organiques en granulométrie. *Bulletin de la Société géographique de Liège*, 18:19-29.
- Lepage H., Masson M., Delanghe D., Le Bescond C. (2019). Grain size analyzers: results of an intercomparison study. *Springer Nature Applied Sciences* 1:1100.
- Lepage H., Masson M., Gruat A., Richard L. (2023). Relation entre les concentrations en contaminants et les variations de granulométrie ou de la nature de MO observée. Rapport final. Observatoire des Sédiments du Rhône – 6ème programme d'action. 27 p.
- Masson M., Angot H., Le Bescond C., Launay M., Dabrin A., Miège C., Le Coz J., Coquery M. (2018). Sampling of suspended particulate matter using particle traps in the Rhône River: Relevance and representativeness for the monitoring of contaminants. *Science of the Total Environment* 637-638:538-549.
- Masson M., Richard L., Gruat L. (2021). Recommandations techniques pour la mesure de la distribution granulométrique des sédiments et matières en suspension de cours d'eau – Rapport AQUAREF 2021 – 26 p.
- Lepage H., Masson M., Gruat A., Richard L. (2023). Effets de la nature des particules sédimentaires (granulométrie et matière organique) sur la concentration des contaminants. OSR6, Axe B, Action B.4.2, Rapport scientifique intermédiaire. Rapport de recherche IRSN, INRAE, 28p.
- Olsen C.R., Cutshall N.H., Larsen I.L. (1982). Pollutant particle associations and dynamics in coastal marine environments - a review. *Marine Chemistry* 11:501-533.
- Phillips J.M., Walling D.E. (2005). Intra-storm and seasonal variations in the effective particle size characteristics and effective particle density of fluvial suspended sediment in the Exe Basin, Devon, United Kingdom. Droppo I.G., Leppard G.G., Liss S.N., Milligan T.M. (eds), *Flocculation in Natural and Engineered Environmental Systems*. CRC Press Boca Raton, pp. 47-70.
- Pierard C., Budzinski H., Garrigues P. (1996). Grain-size distribution of polychlorobiphenyls in coastal sediments. *Environ. Sci. Technol* 30:2776-2783.
- Sheridan M.F., Wohletz K.H., Dehn J. (1987). Discrimination of grain-size subpopulations in pyroclastic deposits. *Geology* 15, 367-370.
- Slattery M.C., Burt T.P. (1997). Particle size characteristics of suspended sediment in hillslope runoff and stream flow. *Earth Surface Processes and Landforms*, 22:705-719.
- Thollet F., Le Bescond C., Lagouy M., Gruat A., Grisot G., Le Coz J., Coquery M., Lepage H., Gairoard S., Gattacceca J.C., Ambrosi J.-P., Radakovitch O., Dur G., Richard L., Giner F., Eyrolle F., Angot H., Mourier D., Bonnefoy A., Dugué V., Launay M., Troudet L., Labille J., Kieffer L. (2021). Observatoire des Sédiments du Rhône, INRAE. <https://dx.doi.org/10.17180/OBS.OSR>
- Yari A., Dabrin A., Coquery A. (2019). Méthodologie d'évaluation des tendances temporelles de contamination dans les sédiments et les matières en suspension des systèmes aquatiques continentaux. *Techniques Sciences Méthodes* 6:71-84.

ANNEXE 1 : décompositions des fractions tamisés A, B et C avec les modèles *Matlab* et *R*.



ANNEXE 1 (suite) : décompositions des fractions A, B et C avec les modèles Matlab et R.



ANNEXE 2 : résultats des décompositions granulométriques (modes, proportions et écart quadratique moyen RMSE) à l'aide des modèles R et Matlab pour les échantillons de MES collectés par centrifugation/pompage (n=77) et par pièges à particules (n=61) à la station de Jons (Rhône) entre août 2012 et juin 2016.

échantillon	modèle R							modèle Matlab						
	μ_1 μm	μ_2 μm	μ_3 μm	λ_1 %	λ_2 %	λ_3 %	RMSE -	μ_1 μm	μ_2 μm	μ_3 μm	λ_1 %	λ_2 %	λ_3 %	RMSE -
JON-TAS-120801-120814	5.05	15.11	38.25	30%	57%	13%		2.91	13.92	39.69	15%	73%	13%	
JON-TAS-120814-120828	4.12	13.84	36.80	23%	62%	14%		4.22	14.26	38.41	26%	59%	15%	
JON-TAS-120828-120911	5.15	14.85	39.81	31%	52%	17%		4.77	14.68	40.18	28%	53%	18%	
JON-TAS-121121-121204	3.51	18.31	51.12	13%	51%	36%		4.50	20.54	54.06	20%	46%	33%	
JON-TAS-121204-121217	2.80	17.20	52.10	10%	51%	39%		4.44	20.90	55.54	19%	44%	37%	
JON-TAS-130122-130205	3.02	18.19	59.14	12%	53%	34%		3.15	18.61	62.06	13%	53%	35%	
JON-TAS-130219-130305	4.20	16.66	44.78	21%	61%	18%		3.26	15.97	45.44	17%	63%	20%	
JON-TAS-130305-130319	3.29	14.18	41.43	21%	66%	13%		3.20	14.23	41.57	18%	68%	14%	
JON-TAS-130409-130423	3.56	16.71	49.24	15%	55%	30%		3.12	16.84	51.30	14%	58%	28%	
JON-TAS-130423-130506	2.61	19.30	74.10	6%	46%	47%		2.48	19.78	77.41	6%	46%	47%	
JON-TAS-130506-130521	3.63	17.20	48.26	12%	59%	29%		6.39	18.04	49.19	30%	41%	29%	
JON-TAS-130604-130618	3.63	14.91	40.99	22%	59%	19%		3.39	15.06	42.89	21%	62%	17%	
JON-TAS-130618-130709	4.02	15.61	38.42	24%	51%	25%		3.53	16.70	41.20	22%	60%	18%	
JON-TAS-130730-130812	2.61	15.00	38.80	19%	78%	3%		2.69	15.39	38.89	14%	77%	9%	
JON-TAS-130910-130924	3.46	14.16	41.00	23%	59%	18%		3.95	14.67	41.54	28%	53%	20%	
JON-TAS-130924-131008	4.08	14.01	41.55	30%	59%	11%		3.47	14.11	41.78	23%	60%	17%	
JON-TAS-131008-131022	4.94	16.56	49.85	28%	46%	26%		6.34	17.77	51.37	33%	38%	28%	
JON-TAS-131022-131105	2.62	17.20	61.00	9%	53%	38%		2.66	17.81	62.94	10%	50%	40%	
JON-TAS-131203-131217	4.50	15.89	46.39	29%	48%	23%		3.17	15.64	48.37	18%	60%	22%	
JON-TAS-131217-140114	2.85	18.82	69.31	8%	50%	42%		2.56	16.88	69.17	8%	46%	46%	
JON-TAS-140114-140128	2.60	17.00	58.50	10%	59%	31%		2.95	17.71	60.56	11%	55%	34%	
JON-TAS-140211-140225	2.69	17.20	56.80	10%	60%	30%		3.06	17.10	58.71	9%	58%	33%	
JON-TAS-140225-140311	5.66	17.08	46.60	34%	38%	28%		3.00	15.67	49.02	15%	59%	26%	
JON-TAS-140311-140326	3.00	15.23	46.74	17%	64%	19%		3.52	15.78	47.49	19%	59%	22%	
JON-TAS-140326-140409	5.56	16.47	44.78	30%	50%	21%		3.03	16.07	47.72	13%	70%	17%	
JON-PAP-140617-140702	4.42	14.71	38.74	30%	54%	16%		4.38	14.35	39.51	30%	51%	19%	
JON-PAP-140924-141021	3.71	14.71	43.94	21%	61%	18%		3.29	14.70	46.60	15%	68%	17%	
JON-PAP-141021-141118	4.24	18.79	52.32	25%	53%	22%		2.74	16.44	52.52	13%	63%	24%	
JON-PAP-141118-141203	5.14	17.87	48.44	27%	44%	28%		3.89	18.20	50.49	19%	52%	30%	
JON-PAP-150106-150120	3.04	15.59	51.33	18%	52%	30%		2.84	16.13	54.41	15%	58%	27%	
JON-PAP-150120-150203	3.00	14.58	62.45	15%	61%	24%		2.82	15.32	55.08	13%	57%	30%	
JON-PAP-150203-150217	3.94	15.77	43.03	21%	57%	22%		3.85	16.03	44.87	22%	56%	22%	
JON-PAP-150217-150303	6.00	19.00	57.00	38%	39%	23%		2.85	15.55	56.01	12%	58%	30%	
JON-PAP-150428-150504	3.58	19.06	50.26	12%	52%	36%		2.98	19.19	52.60	13%	51%	36%	
JON-PAP-150507-150512	4.75	20.65	49.86	18%	48%	35%		2.95	18.73	50.79	11%	52%	37%	
JON-PAP-150512-150526	3.76	17.87	46.95	17%	50%	33%		3.48	18.30	49.71	17%	51%	31%	
JON-PAP-150526-150609	4.13	15.60	44.99	19%	61%	20%		3.42	15.48	45.29	19%	60%	21%	
JON-PAP-150609-150623	4.21	15.79	44.14	24%	58%	18%		1.60	15.43	46.36	20%	68%	12%	
JON-PAP-150623-150707	4.76	13.85	42.72	31%	55%	14%		3.88	14.46	41.93	24%	55%	21%	
JON-PAP-150707-150721	5.73	13.68	36.63	46%	41%	14%		3.76	13.66	39.02	27%	59%	14%	
JON-PAP-150721-150804	3.51	12.50	33.74	22%	65%	13%		4.76	13.42	33.81	37%	48%	15%	
JON-PAP-150804-150818	3.05	14.77	50.00	18%	70%	12%		3.98	13.86	38.81	27%	59%	15%	
JON-PAP-150818-150901	4.02	13.58	38.07	24%	64%	12%		3.31	13.47	39.09	21%	68%	11%	
JON-PAP-150901-150915	3.21	14.33	47.06	19%	64%	17%		3.04	14.54	48.50	16%	66%	18%	
JON-PAP-150915-150929	3.03	15.43	52.17	14%	65%	20%		2.87	15.17	52.06	14%	63%	23%	
JON-PAP-150929-151013	3.20	14.01	43.68	18%	57%	24%		2.19	14.00	47.46	10%	71%	19%	
JON-PAP-151013-151027	2.80	14.62	46.13	15%	73%	12%		3.15	14.86	46.43	17%	68%	15%	
JON-PAP-151027-151110	6.00	18.00	49.44	35%	45%	20%		2.85	16.15	50.74	11%	71%	18%	
JON-PAP-151110-151124	3.14	15.86	52.94	13%	65%	22%		3.47	15.87	52.49	18%	55%	27%	
JON-PAP-151124-151208	3.95	15.63	46.65	26%	56%	18%		2.97	14.88	47.42	18%	62%	20%	
JON-PAP-151208-151222	4.02	14.70	37.89	24%	61%	15%		3.78	14.50	37.70	24%	59%	17%	
JON-PAP-151222-160105	2.76	15.53	49.26	13%	70%	17%		5.16	16.35	51.97	0%	90%	10%	
JON-PAP-160111-160119	5.40	17.31	62.14	23%	45%	32%		3.09	18.11	54.83	14%	53%	33%	
JON-PAP-160119-160202	5.30	18.87	53.61	19%	53%	28%		3.31	16.97	53.92	16%	54%	30%	
JON-PAP-160202-160216	5.07	15.74	50.55	30%	38%	32%		3.33	17.35	52.06	18%	53%	28%	
JON-PAP-160216-160301	5.40	15.86	51.90	25%	45%	30%		2.80	16.51	50.25	13%	57%	29%	
JON-PAP-160301-160315	4.46	17.95	51.81	22%	48%	30%		3.11	17.39	53.69	14%	55%	31%	
JON-PAP-160315-160329	4.81	15.87	52.88	26%	53%	22%		3.26	16.63	49.96	15%	64%	21%	
JON-PAP-160412-160426	3.35	16.47	55.33	13%	60%	27%		3.05	16.47	57.29	14%	56%	29%	
JON-PAP-160426-160510	6.00	18.00	50.93	34%	46%	20%		4.83	19.00	39.15	19%	51%	30%	
JON-PAP-160705-160719	5.60	16.36	58.99	24%	44%	32%		3.05	16.83	52.95	11%	55%	34%	

échantillon	modèle R							modèle Matlab						
	μ_1 μm	μ_2 μm	μ_3 μm	λ_1 %	λ_2 %	λ_3 %	RMSE -	μ_1 μm	μ_2 μm	μ_3 μm	λ_1 %	λ_2 %	λ_3 %	RMSE -
JON-CMO-120801	4.15	12.28	32.58	37%	57%	6%	0.127	2.71	12.01		20%	80%		0.107
JON-CMO-120814	4.53	18.96		36%	64%		0.565	3.56	13.22		26%	74%		0.086
JON-CMO-120828	2.89	12.70		27%	73%		0.172	4.37	12.87	33.71	47%	40%	12%	0.196
JON-CMO-120911	2.92	10.22		62%	38%		0.311	3.55	11.36		78%	22%		0.048
JON-CMO-121121	3.92	11.86	28.29	46%	51%	4%	0.117	3.12	11.75		34%	66%		0.085
JON-CMO-121204	3.30	11.47	34.58	27%	60%	12%	0.128	3.74	12.50	35.96	37%	50%	13%	0.218
JON-CMO-121217	2.69	11.88	43.36	52%	37%	11%	0.120	3.19	12.27	42.99	48%	38%	14%	0.254
JON-CMO-130108	5.34	12.43	31.47	55%	37%	8%	0.105	4.42	12.54	31.74	43%	47%	10%	0.158
JON-CMO-130122	4.26	11.03		38%	62%		0.409	2.93	10.91		36%	64%		0.101
JON-CMO-130205	2.70	10.30		80%	20%		0.155	2.97	11.18		84%	16%		0.089
JON-CMO-130219	3.20	12.51		33%	67%		0.424	4.18	13.07	36.75	47%	42%	11%	0.186
JON-CMO-130305	5.78	15.01	45.88	45%	44%	10%	0.082	3.47	14.21	48.24	23%	67%	10%	0.170
JON-CMO-130319	3.33	18.19		45%	55%		0.548	4.07	12.74	33.99	54%	37%	9%	0.149
JON-CMO-130409	3.69	12.61		33%	67%		0.2	3.74	13.10	34.69	40%	53%	6%	0.155
JON-CMO-130423	3.01	12.76	34.87	32%	54%	14%	0.162	3.59	13.23	33.92	48%	35%	17%	0.282
JON-CMO-130506	3.41	11.36		58%	42%		0.156	3.70	12.05		68%	32%		0.073
JON-CMO-130521	2.99	14.69		53%	47%		0.457	3.34	11.60		60%	40%		0.057
JON-CMO-130604	2.95	13.86	39.70	24%	54%	22%	0.113	3.16	14.38	40.83	25%	53%	22%	0.357
JON-CMO-130618	3.95	13.59		39%	61%		0.17	3.64	12.38	29.19	34%	58%	8%	0.149
JON-CMO-130709	3.97	13.03	30.32	39%	48%	13%	0.109	4.20	13.46	31.76	46%	41%	13%	0.210
JON-CMO-130716	4.21	12.99	27.34	44%	46%	9%	0.069	4.33	13.24	28.85	43%	48%	9%	0.140
JON-CMO-130812	2.85	11.61	40.97	55%	31%	13%	0.603	3.42	11.41	38.58	56%	27%	17%	0.402
JON-CFI-130924	4.26	12.76	30.37	63%	29%	9%	0.515	4.23	12.64	29.46	60%	29%	11%	0.194
JON-CFI-131008	4.20	12.07	31.19	57%	38%	5%	0.513	4.56	11.84	28.89	59%	34%	8%	0.120
JON-CFI-131022	4.33	13.36	40.16	42%	43%	16%	0.323	5.76	13.56	43.02	50%	37%	12%	0.222
JON-CFI-131105	2.64	12.33	50.88	24%	60%	15%	0.485	3.31	13.37	49.34	38%	43%	19%	0.302
JON-CMO-131203	4.11	16.66		46%	54%		0.182	4.64	12.47	35.42	50%	39%	11%	0.174
JON-CMO-131217	3.15	12.30	42.76	30%	62%	8%	0.917	3.75	12.76	39.43	35%	52%	13%	0.209
JON-CMO-140114	3.28	13.25	38.54	23%	65%	12%	0.617	3.54	13.69	38.34	31%	54%	15%	0.259
JON-CMO-140128	5.10	13.27	32.06	51%	40%	8%	0.54	3.45	12.43	32.79	30%	60%	10%	0.170
JON-CMO-140211	3.15	18.31		67%	33%		0.724	3.48	11.74	32.11	71%	22%	7%	0.150
JON-CMO-140225	3.04	10.96	35.22	40%	57%	4%	0.985	4.18	11.47	24.83	57%	26%	17%	0.250
JON-CMO-140311	3.63	14.76	38.54	27%	54%	19%	0.576	3.63	14.96	39.10	25%	54%	21%	0.342
JON-CMO-140326	2.92	12.79	36.63	30%	63%	6%	0.171	3.92	13.68	34.80	42%	44%	14%	0.227
JON-CMO-140409	3.69	15.52	38.80	17%	67%	16%	-	2.69	15.87	42.32	13%	76%	12%	-
JON-CFI-140617	4.36	12.99	35.20	37%	54%	9%	0.168	6.63	13.68	37.77	58%	38%	4%	0.152
JON-CFI-140702	3.51	13.00	38.92	28%	62%	10%	0.228	4.26	13.46	37.58	36%	50%	14%	0.236
JON-CFI-140715	2.27	10.40		30%	70%		0.36	3.79	11.22	26.76	68%	21%	11%	0.170
JON-CFI-140826	3.32	12.90	35.56	31%	55%	14%	0.416	3.78	13.61	36.09	39%	45%	17%	0.270
JON-CFI-140924	4.00	12.68	33.08	40%	49%	11%	0.068	4.66	12.75	31.63	47%	38%	14%	0.211
JON-CFI-141021	3.87	11.90	30.14	51%	43%	6%	0.356	3.53	11.74	29.86	43%	50%	7%	0.124
JON-CFI-141118	3.97	12.47	33.07	50%	42%	7%	0.465	3.63	11.88	31.69	46%	43%	12%	0.203
JON-CFI-141203	3.84	12.89	36.07	37%	53%	10%	0.587	4.12	12.66	34.66	39%	45%	15%	0.238
JON-CFI-150106	2.02	9.39		40%	60%		0.625	3.08	11.19		57%	43%		0.099
JON-CFI-150120	3.17	11.07	31.25	54%	40%	5%	0.537	3.83	11.90	29.83	69%	23%	7%	0.138
JON-CFI-150203	3.26	11.34	27.26	51%	45%	4%	0.835	3.33	11.17	26.60	54%	39%	8%	0.146
JON-CFI-150217	3.07	11.71	36.81	30%	60%	10%	0.161	4.26	12.69	34.91	49%	37%	14%	0.227
JON-CFI-150303	3.12	11.82	38.14	31%	60%	9%	0.404	4.08	12.68	36.27	49%	37%	14%	0.236
JON-CFI-150414	3.98	14.00	40.82	27%	57%	16%	0.306	4.13	14.59	42.59	31%	53%	16%	0.259
JON-CFI-150428	3.65	12.81	38.51	33%	58%	9%	0.46	4.41	13.14	37.61	34%	51%	15%	0.273
JON-CFI-150504	3.40	11.13	31.38	49%	40%	11%	0.56	4.13	13.72	30.56	37%	43%	21%	0.231
JON-CFI-150507	3.82	13.61	30.71	34%	49%	17%	0.574	3.88	12.15	30.59	59%	29%	12%	0.306
JON-CFI-150512	3.35	11.54	31.48	49%	43%	8%	0.208	4.04	13.21	36.57	34%	51%	14%	0.197
JON-CFI-150526	3.38	12.70	37.43	26%	64%	10%	0.459	3.68	12.67	33.14	46%	45%	9%	0.232
JON-CFI-150609	3.31	11.61	32.99	39%	54%	7%	0.464	3.68	12.67	33.14	46%	45%	9%	0.169
JON-CFI-150623	3.76	11.72	29.73	52%	41%	7%	0.467	3.56	11.70	31.70	51%	42%	7%	0.143
JON-CFI-150707	2.83	13.06		41%	59%		0.731	3.65	12.67	30.99	55%	32%	12%	0.213
JON-CFI-150721	4.12	12.62	32.52	26%	63%	10%	0.874	4.23	13.03	32.38	30%	58%	12%	0.191
JON-CFI-150804	3.54	15.91		46%	54%		0.711	3.60	11.18		48%	52%		0.099
JON-CFI-150818	4.02	12.47		66%	34%		0.18	3.63	11.10		51%	49%		0.089
JON-CFI-150901	2.01	10.80		43%	57%		0.396	2.92	11.32		44%	56%		0.098
JON-CFI-150915	2.52	11.49		24%	76%		0.366	3.99	12.22	30.01	48%	42%	11%	0.174
JON-CFI-150929	3.37	11.65		51%	49%		0.629	3.95	11.82	25.90	60%	35%	5%	0.100
JON-CFI-151013	3.74	11.89	35.27	37%	55%	8%	0.382	2.54	10.51	38.23	17%	78%	5%	0.141
JON-CFI-151027	4.88	14.51	35.81	43%	50%	8%	0.558	4.51	13.57	33.08	37%	48%	15%	0.227
JON-CFI-151110	4.41	14.01	31.24	39%	50%	11%	0.396	3.18	15.02		28%	72%		0.124
JON-CFI-151124	2.86	12.69	40.80	19%	68%	12%	0.601	3.74	13.81	39.81	29%	55%	16%	0.249
JON-CFI-151208	3.01	11.27	35.13	39%	57%	4%	0.387	4.51	12.50	32.48	65%	29%	6%	0.117
JON-CFI-151222	3.55	12.08	31.92	34%	60%	5%	0.401	4.01	12.58	31.80	44%	48%	9%	0.146
JON-CFI-160105	2.06	11.38		30%	70%		0.976	4.20	12.34	30.85	51%	39%	10%	0.158
JON-CFI-160111	3.03	10.89	33.70	54%	41%	5%	0.291	3.56	11.81	30.35	69%	23%	8%	0.143
JON-CFI-160119	2.53	10.33		36%	64%		0.184	4.11	11.93	28.73	68%	25%	7%	0.116
JON-CFI-160202	2.00	9.80		50%	50%		0.816	3.93	11.96	30.96	55%	35%	11%	0.889
JON-CFI-160216	3.00	10.04		45%	55%		0.597	3.92	12.09	29.72	76%	19%	5%	0.145
JON-CFI-160301	3.52	12.00	31.96	48%	45%	7%	0.297	4.49	11.99	28.39	62%	24%	13%	0.230
JON-CFI-160315	3.49	18.67		46%	54%	0%	0.641	4.55	11.99	27.41	66%	28%	6%	0.112
JON-CFI-160719	3.22	12.19	35.94	42%	40%	18%	0.269	3.70	13.12	37.98	46%	35%	18%	0.284