Couplage de l'analyse environnementale et de l'apprentissage automatique pour les procédés de valorisation des déchets lignocellulosiques

Nancy PRIOUX[1], Gilles HETREUX[1] et Jean-Pierre BELAUD[1]

Affiliation 1 : Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France.

Ces dernières années ont vu l'émergence de plans d'action pour le climat, le développement durable ou l'environnement comme le plan européen « *Green Deal* » de janvier 2020. L'un des principaux plans d'action, intitulé "Plan d'action pour l'économie circulaire", propose d'utiliser le modèle d'économie circulaire (EC) pour simuler l'utilisation de modèles durables. Pour parvenir à des modèles durables, la « pensée cycle de vie » peut contribuer à améliorer les performances environnementales tout en maximisant les avantages économiques et sociaux. Plusieurs méthodes globales ont vu le jour pour concevoir des procédés de recyclage des déchets qui s'inscrivent dans l'économie circulaire (Grimaud et al., 2017). L'agriculture est un domaine particulier dans lequel l'EC et la « pensée cycle de vie » se sont développées au cours de la dernière décennie. Une grande partie des déchets agricoles sont des sous-produits lignocellulosiques qui peuvent être transformés en bioénergie, biomolécules ou biomatériaux. Au cours des 30 dernières années, de nombreux procédés de prétraitement ont été étudiés et publiés (Davis et al., 2017). Cependant, il y a un manque de critères pour guider le choix parmi ces différentes voies technologiques. L'utilisation de l'évaluation environnementale, économique et sociale dans un contexte d'EC est un bon moyen de guider ce choix. Ces évaluations nécessitent une grande quantité de données, en particulier des données sur les procédés potentiellement disponibles dans la littérature scientifique ou technique ou alimentées par la simulation (Morales-Mendoza et al., 2012). La quantité d'articles scientifiques décrivant les procédés de valorisation des sous-produits agricoles constitue une source de données importante et est sous-exploitée. Cette quantité peut être difficile à traiter. Cependant, de nombreux outils et méthodes sont désormais disponibles à cette fin grâce aux techniques venant du « *Big Data* » ou de l'intelligence artificielle comme l'apprentissage automatique (ML). Les outils d'évaluation environnementale tels que l'analyse du cycle de vie (ACV) peuvent bénéficier de ces technologies pour analyser les données des procédés et les résultats de ces évaluations.

L'approche présentée fait le pont entre l'apprentissage automatique et l'évaluation de la durabilité pour aider le chercheur ou l'ingénieur dans l'analyse et la comparaison de différents procédés de prétraitement de la biomasse lignocellulosique et de différents types de biomasse dans le contexte de l'économie circulaire. Cette approche est conçue pour faire un premier tri dans les procédés de prétraitement ou les biomasses à partir des données contenues dans la littérature scientifique (Belaud et al., 2019). Elle est divisée en cinq étapes principales : objectif et frontière, architecture de données, évaluation de la durabilité, visualisation et analyse des résultats, et décision. La figure 1 présente les étapes principales de l'approche.

Dans la première étape, l'objectif et les frontières de l'étude doivent être clairement définis. La « pensée cycle de vie » est recommandée. Cette pensée encourage une approche "du berceau à la tombe" ou "du berceau à la porte" si la logistique d'une chaîne de valeur est difficile à obtenir voire du « berceau au berceau » lors de l'ingénierie circulaire. Les frontières du système et l'unité fonctionnelle influencent considérablement les évaluations. Par exemple, il convient de clarifier si la chaîne d'approvisionnement en biomasse en amont est prise en compte. Une fois l'objectif et le champ d'application correctement définis, la chaîne d'approvisionnement, les technologies et les procédés de transformation doivent également être décrits.

L'architecture de données est directement inspirée de la construction de l'architecture *Big Data* et se compose de cinq sous-étapes : (i) collecte et extraction des données, (ii) enrichissement et

stockage des données, (iii) traitement des données, (iv) analyse des données (brutes) et (v) visualisation des données (brutes). Cette étape peut être automatique, semi-automatique ou manuelle et elle utilise des méthodes issues de l'ingénierie des connaissances ou de ML.

La troisième étape consiste à compléter l'inventaire du cycle de vie en utilisant les données de procédés de l'étape précédente. L'inventaire est complété par des bases de données gratuites ou commerciales telles que la base de données EcoInvent contenant des données d'arrière-plan. Ensuite, une ou plusieurs méthodes de calcul des impacts doivent être déterminées conformément à la première étape qui intègre la nature de l'étude et du système. Les impacts de la durabilité voire les sphères de dommages sont calculés. A l'issue de cette étape, le résultat principal est une structure [procédés : biomasses : impacts] difficilement analysable.

L'étape de visualisation et d'analyse des résultats comprend des méthodes dérivées de ML pour aider à l'analyse de cette matrice. En partant de la littérature statistique, les techniques traditionnelles de réduction des dimensions (*Multi-Dimensional Scaling*) et de regroupement non supervisé (*k-means*) sont combinées pour extraire des informations pertinentes. Un modèle de prédiction du rendement en glucose basé sur un réseau neuronal artificiel a également été utilisé, puis des diagrammes de dépendance partielle (PDP) et des diagrammes d'espérance conditionnelle individuelle (ICE) peuvent être utilisés pour visualiser l'interaction entre les entrées et les sorties. Enfin, l'utilisateur (expert) analyse les points regroupés en clusters pour les relier à des procédés/impacts significatifs. La visualisation des groupes de données aidera le chercheur ou les ingénieurs R&D dans la décision finale suivant l'objectif de leur étude. L'approche est testée sur la comparaison des procédés de prétraitement de la paille de riz et de la canne de maïs pour la production de glucose.

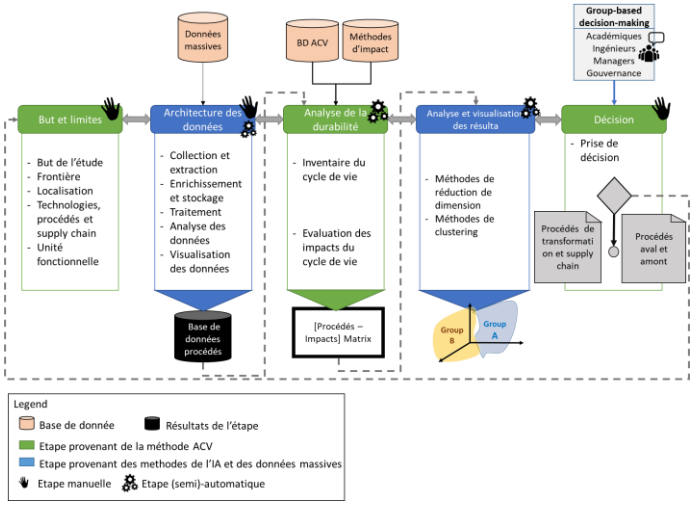

Figure 1 : Descriptif général de l'approche de couplage ML/ACV

Belaud, J.-P.; Prioux, N.; Vialle, C.; Sablayrolles, C., Big data for agri-food 4.0 : Application to sustainability management for by-products supply chain, *Computers in Industry*, 2019, 111, 41-50.

Davis, C. B.; Aid, G.; Zhu, B., Secondary Resources in the Bio-Based Economy : A Computer Assisted Survey of Value Pathways in Academic Literature, *Waste and Biomass Valorization*, 2017, 8(7), 2229-2246.

Grimaud, G.; Perry, N.; Laratte, B., Decision Support Methodology for Designing Sustainable Recycling Process Based on ETV Standards. *Procedia Manufacturing*, 2017, 7, 72-78.

Morales-Mendoza, L. F.; Azzaro-Pantel, C. ; Belaud, J.-P. ; Pibouleau, L. ; Domenech, S.,. An integrated approach combining process simulation and life cycle assessment for eco-efficient process design. *Computer Aided Chemical Engineering*, 2012, 30, 142-146.