

## Poste de thèse à IFP Energies nouvelles (IFPEN) en Sciences chimiques

### Paramétrage du modèle thermodynamique SAFT par apprentissage machine

L'utilisation de la biomasse apparaît comme une alternative prometteuse pour la synthèse de diverses familles de produits chimiques à forte valeur ajoutée, qui peuvent être utilisés pour la fabrication d'autres produits. Les charges issues de la biomasse et leurs dérivés sont principalement composés de molécules oxygénées, or les modèles et méthodes traditionnellement utilisés dans l'industrie pétrochimique ont historiquement été développés pour restituer les propriétés d'hydrocarbures. La conception de nouveaux procédés de transformation nécessite des méthodes d'estimation rapide et précise des propriétés physico-chimiques d'intérêt pour l'industrie chimique.

Les modèles thermodynamiques - équation d'état (EoS), en particulier l'EoS PC-SAFT (pour « Perturbed Chain Statistical Associating Fluid Theory ») - et ses dérivés sont largement utilisés pour calculer les propriétés des fluides pour des conditions ciblées de température et de pression. Ce modèle est fondé sur une thermodynamique statistique forte qui assure de bonnes capacités d'extrapolation. Son application nécessite la connaissance de paramètres spécifiques au fluide à étudier. Des contributions de groupes ont été développées pour prédire certains de ces paramètres, mettant ainsi en évidence le lien entre la structure des fluides considérés et les valeurs des paramètres associés.

Dans ce contexte, nous proposons ce sujet pour mettre en œuvre des méthodes basées sur la science des données afin de développer de nouvelles approches pour paramétrer l'EoS PC-SAFT, et ainsi étendre son champ d'application notamment aux nouvelles filières énergétiques. La combinaison envisagée de méthodes d'apprentissage machine et d'un modèle thermodynamique vise à améliorer la capacité prédictive de chacun de ces modèles pris individuellement. Plusieurs approches seront étudiées : (i) des approches consistant à prédire les paramètres de l'EoS au moyen de modèles basés sur la QSPR (« quantitative structure property relationship »), et (ii) une approche combinant la physique et l'apprentissage machine.

**Mots clefs:** équation d'état, apprentissage machine

<b>Directeur de thèse</b>	Pr. De HEMPTINNE, (ORCID : 0000-0003-1607-3960), Ingénieur de recherche, Département de Thermodynamique et Modélisation moléculaire, <a href="mailto:jean-charles.de-hemptinne@ifpen.fr">jean-charles.de-hemptinne@ifpen.fr</a>
<b>Ecole doctorale</b>	ED388 Chimie Physique et Chimie Analytique de Paris Centre, <a href="http://ed388.upmc.fr/">ed388.upmc.fr/</a>
<b>Encadrant IFPEN</b>	Dr. CRETON Benoît (ORCID: 0000-0002-3287-877X), Ingénieur de recherche, Département de Thermodynamique et Modélisation moléculaire, <a href="mailto:benoit.creton@ifpen.fr">benoit.creton@ifpen.fr</a>
<b>Localisation du doctorant</b>	IFP Energies nouvelles, Rueil-Malmaison, France
<b>Durée et date de début</b>	3 ans, début au cours du 4ème trimestre 2023. Salaire brut annuel > 25k€.
<b>Employeur</b>	ProSim, Toulouse, France.
<b>Qualifications</b>	Master en statistiques/sciences des données, en sciences chimiques/physiques ou en génie chimique.
<b>Connaissances linguistique</b>	Maîtrise de l'anglais et du français ou forte volonté d'apprendre le français
<b>Autres qualifications</b>	Thermodynamique, chémoinformatique, apprentissage machine

Pour postuler, merci d'envoyer votre lettre de motivation et votre CV à l'encadrant IFPEN indiqué ci-dessus.

#### IFP Energies nouvelles

IFP Energies nouvelles est un organisme public de recherche, d'innovation et de formation dont la mission est de développer des technologies performantes, économiques, propres et durables dans les domaines de l'énergie, du transport et de l'environnement. Pour plus d'information, voir [notre site web](#).

IFPEN propose à tous les doctorants de participer à des séminaires et des formations qui leur sont dédiés. Pour plus d'information, merci de consulter nos [pages web dédiées](#).

## PhD position at IFP Energies nouvelles (IFPEN) in *Chemical sciences* **Parameterization of the SAFT thermodynamic model using machine learning**

The use of biomass appears as a promising alternative for the synthesis of a various families of strong value-added chemicals, which can be used in the manufacture of products. Product issued from the biomass are mainly composed of oxygenated molecules, but the models and methods traditionally used in the petrochemical industry have historically been developed to restore properties of hydrocarbons. The design of new manufacturing processes requires methods for rapid and accurate estimation of physicochemical properties of interest in the chemical industry.

Thermodynamic models - equation of state (EoS), in particular the EoS PC-SAFT (for "Perturbed Chain Statistical Associating Fluid Theory") and its derivatives are widely used to compute the properties of fluids for targeted conditions of temperature and pressure. This model is founded on strong statistical thermodynamics which ensures good extrapolation capacities. Its application requires knowledge of parameters specific to the fluid to be studied. Group contributions have been developed to predict some of these parameters, thus highlighting the link between the structure of the considered fluids and the values of associated parameters.

In this PhD subject, we propose to implement methods based on data science to develop new approaches to parameterize the PC-SAFT EoS, and thus extend its field of application to area of energy processes. The envisaged combination of machine learning and a thermodynamic model aims at improving the predictive capacity of each of these models taken individually. Several approaches will be investigated: (i) approaches consisting in predicting the EoS parameters by means of QSPR (quantitative structure-property relationship) based models, and (ii) an approach mixing physics and machine learning.

It should be noted that this work will be funded by ProSim, carried out at IFPEN in close collaboration with research groups at Université de Lorraine.

**Keywords:** Equation of states, Machine Learning

<b>Academic supervisor</b>	Pr. De HEMPTINNE Jean-Charles (ORCID : <a href="https://orcid.org/0000-0003-1607-3960">0000-0003-1607-3960</a> ), Research engineer, Thermodynamics and Molecular Simulation Department, <a href="mailto:jean-charles.de-hemptinne@ifpen.fr">jean-charles.de-hemptinne@ifpen.fr</a>
<b>Doctoral School</b>	ED388 Chimie Physique et Chimie Analytique de Paris Centre, <a href="https://ed388.upmc.fr/">ed388.upmc.fr/</a>
<b>IFPEN supervisor</b>	Dr. CRETON Benoît (ORCID: <a href="https://orcid.org/0000-0002-3287-877X">0000-0002-3287-877X</a> ), Research engineer, Thermodynamics and Molecular Simulation Department, <a href="mailto:benoit.creton@ifpen.fr">benoit.creton@ifpen.fr</a>
<b>PhD location</b>	IFP Energies nouvelles, Rueil-Malmaison, France
<b>Duration and start date</b>	3 years, starting not earlier than December 2023. Minimum gross annual salary > 25k€
<b>Employer</b>	ProSim, Toulouse, France.
<b>Academic requirements</b>	University master's degree in Statistics/Data sciences, Chemical/Physical sciences or Chemical Engineering
<b>Language requirements</b>	Fluency in English, and in French or willingness to learn French
<b>Other requirements</b>	Thermodynamics, Chemoinformatics, Machine Learning.

To apply, please send your cover letter and CV to the IFPEN supervisor indicated here above.

### About IFP Energies nouvelles

IFP Energies nouvelles is a French public-sector research, innovation and training center. Its mission is to develop efficient, economical, clean and sustainable technologies in the fields of energy, transport and the environment. For more information, see <https://www.ifpen.com>.

IFPEN offers a stimulating research environment, with access to first in class laboratory infrastructures and computing facilities. IFPEN offers competitive salary and benefits packages. All PhD students have access to dedicated seminars and training sessions. For more information, please see our [dedicated WEB pages](#).