Abstract

Decarbonization of the oil refining industry is essential for reducing carbon emissions and mitigating climate change. Co-processing bio feed at existing oil refineries is a promising strategy for achieving this goal. However, accurately quantifying the renewable carbon content of co-processed fuels can be challenging due to the complex process involved. Currently, it can only be achieved through expensive offline $_{14}$C measurements. To address this issue, with high-quality and large-scale commercial data, our study proposes a novel approach that utilizes data-driven methods to build inferential sensors, which can estimate the real-time renewable content of biofuel products. We have collected over 1,000,000 co-processing data points from refineries under different bio feed co-processing ratios and operational conditions—the largest dataset of its kind to our knowledge We use interpretable deep neural networks to select model inputs, then apply robust linear regression and bootstrapping techniques to estimate renewable content and confidence interval. Our method has been validated with four previous $_{14}$C measurements during co-processing at the fluid catalytic cracker. This novel methods provides a practical solution for the industry and policymakers to quantify renewable carbon content and accelerate the transition to a more sustainable energy system.

Introduction

Our society's dependence on fossil fuels is one of the main causes of climate change, and its continued use has led to many problems, such as increased extreme weather events, ecological damage, forest fires, and glacier melting. According to the International Energy Agency (IEA), in 2020, fossil fuels accounted for approximately 80% of the world's primary energy consumption [1]. Achieving the net-zero goal and decarbonizing the economy requires a fundamental shift from a fossil-fuel-based economy to a renewable energy-dominated economy [2].

The current pace of decarbonization by building from scratch is accelerating (such as solar photovoltaic system, wind farms, and electric cars), but still much slower than expected. Thus, to further accelerate the process in a cost effective way, working with the current fossil fuel industry is a quicker and reliable path forward. Co-processing is a technology that enables oil refineries to utilize into biorefineries by co-processing low-carbon-intensive feedstocks, such as bio-crudes made from forest, mill residues [3], [4], microalgae, municipal sludge, and municipal waste [5]. These materials can undergo processes such as hydrothermal liquefaction to yield biocrude, which is then suitable for further processing [6], [7]. It has a lower carbon footprint than conventional petroleum refining since burning bio feed-derived fuel is carbon neutral (although producing the biofuel will have some greenhouse gas emissions).

Co-processing oleochemical feedstocks have been fully commercialized by various refiners around the world, either in the hydrotreater [8] or at the fluid catalytic cracker [9]. This provides valuable experiences for refiners to co-process biocrudes made from waste feedstocks in the future. As the world demands lower carbon intensity fuels, the renewable content of these fuels will become another essential property, much like sulfur, nitrogen, and flash point are for the fuels produced today. This means that fuel producers will need to consider the renewable content of their products as a critical factor. Furthermore, policies like the Low Carbon Fuel Standard (LCFS) incentivize refiners to co-process renewable feedstocks with conventional crude oil [10]. To ensure the effectiveness of policies like LCFS in influencing the oil sector, it is crucial to estimate accurately the reduction in carbon intensity achieved through co-processing renewable feedstocks with conventional crude oil.

The task of accurately tracking the renewable content of fuels resulting from co-processing of biogenic feedstocks with fossil fuels poses significant challenges [3], [11]. A key issue in this regard is the uneven distribution of green molecules across different fractions combined with the limited availability of effective techniques for quantifying the renewable content of each stream. One common approach for tracking the renewable content from fossil molecules is the isotopic analysis, specifically radiocarbon dating using the

isotope [9], [12]. This technique, widely used in the field of archaeology to determine age, can be adapted to trace the origin of carbon molecules. Biofuels were produced in recent years while fossil fuels were formed hundreds of millions of years ago. This property makes

a useful tool for quantifying the renewable content of biofuels. However, if multiple measurements are needed, the equipment setup and maintenance costs can be very high (thus only a few dedicated labs can perform such tests). Additionally, this method cannot provide continuous measurements which means the test results only provide a snapshot of the operation. Another potential solution is the development of online renewable content monitoring systems to provide continuous measurements. However, the implementation of such equipment requires extensive evaluation before oil refiners can fully adopt them.

The co-processing process is highly complex, involving multiple interrelated unit operations and control loops. The process exhibits high-dimensional, nonlinear, and dynamic characteristics. The emergence of artificial intelligence technology has brought new ideas and solutions for the monitoring and optimization of this type of industrial processes [13], [14], [15]. Artificial intelligence technology can learn and analyze a large amount of historical data to establish accurate soft sensor models, thereby achieving real-time monitoring and prediction of process status and performance [16], [17], [18].

The soft sensor is a data-driven artificial intelligence method designed to predict and monitor key performance indicators in the production process [17], [18]. Soft sensors are usually built using machine learning algorithms such as artificial neural networks, support vector machines, and decision trees. They have become an essential tool for process monitoring and control, leading to improved efficiency, product quality, and reduced downtime. It takes existing external variables as inputs and builds mathematical models to predict and estimate the values of some key process parameters or variables, thus enabling real-time monitoring and control.

Our industrial partner, Parkland Refining Ltd, is currently co- processing oleochemical/lipid feedstocks such as tallow, canola oil, and tall oil that reduce the carbon intensity (CI) of the various fuels that they produce. This is similar to what other oil companies around the world, such as BP in Washington State, Preem in Sweden, and ENI in Italy are doing to reduce their carbon footprint. Parkland is expanding its co-processing to 5500 barrels per day [19]. The availability of a significant amount of industrial data generated through commercial operations provides unique opportunity to build more reliable and robust AI models that can effectively track the renewable content during FCC co-processing.

In the sections that follow, we will detail the innovations in our study. These aspects differentiate our work from typical research in the field, provide valuable contributions and new perspectives to the existing research.

For the theoretical innovations, this study is pioneering in its domain, emphasizing the critical role of artificial intelligence in decarbonizing the oil refining industry. This study proposes an approach that can contribute to the innovative fusion of interpretable deep neural networks, robust linear regression and bootstrapping techniques to achieve high accuracy in estimating the real-time renewable content of liquids produced, signifying a leap towards the continuous, real-time monitoring of renewable carbon. The efficacy is further validated with four

measurements employed during co-processing at the FCC. With an average error rate below 4%, we have demonstrated that AI can play a significant role in enhancing the precision of renewable carbon tracking.

For dataset contribution, we have amassed over 1,000,000 co-processing data points from refineries. This dataset, to our best knowledge, is the largest of its kind. While traditional laboratory-scale studies are often limited to a few or tens of samples, our dataset stands out in its depth and breadth. By encapsulating varied operational conditions such as temperature and catalyst activity, along with diverse bio-oil inclusion ratios, our data is pivotal in driving forward significant advancements in the domain of bio feed co-processing.

In practical implications, our findings provide oil refineries with an invaluable toolset to quantify renewable carbon accurately. This can help policymakers and stakeholders better understand the performance of bio feed in renewable energy production, informing decisions on promotion and investment in the industry. Furthermore, our method offers substantial cost reductions for refineries (millions of dollars) and ensures companies secure optimal governmental incentives through accurate renewable carbon measurements. This precision not only provides financial benefits but also propels the broader shift towards renewable energy sources.

Section snippets

Co-processing at the FCC

Co-processing is a process that involves the simultaneous treatment of liquefied biomass and petroleum-based feedstocks (such as crude oil or natural gas liquids) in a refinery. In this process,the liquefied biomass feedstock is derived from the bioprocessing of solid biomass materials, such as agricultural waste, forestry residues, crop residues, and organic waste, to yield compatible intermediates suitable for refinery operations. These intermediates are then treated to meet essential

AI for renewable carbon tracking

The integration of Artificial Intelligence (AI) into the field of bio feed co-processing has the potential to provide new and innovative solutions for tracking the renewable carbon. In this work, we will focus on the application of soft sensors for modeling and predicting renewable carbon in real-time. By utilizing deep learning for feature selection and combining bootstrapping and robust linear regression for building robust soft sensor models, we aim to use AI to track renewable carbon

Data gathering and preprocessing

In this section, we utilized commercial-scale co-processing data from the Parkland refinery in Burnaby, British Columbia, Canada, for our case study. It is important to highlight the value and uniqueness of the data we have collected for our study. This is a commercial dataset that provides a large volume of high-quality data from continuous co-processing operations at a commercial scale. Its commercial nature ensures that the data reflects real-world scenarios and practical considerations,

Conclusion

In conclusion, this study presents a novel AI-based method for tracking renewable carbon during bio feed co-processing, utilizing large-scale commercial data, interpretable deep neural networks, robust linear regression, and bootstrapping. Our approach efficiently and accurately estimates real-time renewable carbon content in produced liquids, potentially saving oil refineries millions of dollars annually by eliminating the need for expensive

measurements. This research holds significant