12-20-2023

A Systems Approach to Process Design and Sustainability - Synergy via Pollution Prevention, Control, and Source Reduction

Emmanuel Apau Aboagye
Rowan University

A SYSTEMS APPROACH TO PROCESS DESIGN AND SUSTAINABILITY – SYNERGY VIA POLLUTION PREVENTION, CONTROL, AND SOURCE REDUCTION

by

Emmanuel Apau Aboagye

A Dissertation

Submitted to the
Department of Chemical Engineering
College of Engineering
In partial fulfillment of the requirement
For the degree of
Doctor of Philosophy
at
Rowan University
September 14, 2023

Dissertation Chair: Kirti M. Yenkie, Ph.D., Associate Professor and Graduate Program Coordinator, Department of Chemical Engineering, Rowan University

Committee Members:
Gerardo J. Ruiz-Mercado, Ph.D., Researcher, US Environmental Protection Agency, Cincinnati, OH
Kauser Jahan, Ph.D., P.E., Professor and Department Head, Department of Civil & Environmental Engineering, Rowan University
Joseph F. Stanzione, III, Ph.D., Professor and Director of AMMI, Department of Chemical Engineering, Rowan University
Kevin D. Dahm, Ph.D., Professor, Department of Chemical Engineering, Rowan University

Dedications

I dedicate this work to my dad, mum, brother (Kwame), and to my sisters (Afia and Awo) for their steadfast support throughout my academic studies. Special dedication to Prince, Franca, Obed, Kwabena Darko, Prophet Alex, and Rev. Gerald for their relentless support.

Acknowledgments

I want to express my sincerest gratitude to my advisor, Dr. Kirti M. Yenkie, for her unwavering support and guidance throughout my studies. I would like to thank Dr. Gerardo J. Ruiz-Mercado, Dr. Kauser Jahan, Dr. Joseph F. Stanzione, III, and Dr. Kevin Dahm for serving on my dissertation committee. A special thank you to my fellow graduate students: Swapana Jerpoth, Fred Ghanem, and Austin Lehr for their invaluable input during my time as a Ph.D. student. A big thank you to my undergraduate students that I mentored through our engineering clinic projects over the years: Phuong Le, Maya Desai, Carley Tran, Quint Kearns, Liela Clarke, Brandon Jarrett, Ethan Schumaker, Chris Cavalieri, Tyler Heritage, Elizabeth Dellorco, and Anne-Marie Zamor for their assistance. I would like to express my gratitude to past colleagues at the Sustainable Design and Systems Medicine Lab, especially Dr. John Chea and Jake Stengel for their immense contributions. Thank you to my friends Dr. Prince Atsu, Dr. Stephen Boakye-Ansah, Wilfred Darko, Solomon Okine, Daniel Randy, Elizabeth Toffey, and Francisca Essuman-Quansah for their unwavering support.

Abstract

Emmanuel Apau Aboagye
A SYSTEMS APPROACH TO PROCESS DESIGN AND SUSTAINABILITY – SYNERGY VIA POLLUTION PREVENTION, CONTROL, AND SOURCE REDUCTION
2023-2024
Kirti M. Yenkie, Ph.D.
Doctor of Philosophy in Chemical Engineering

Historically, process design prioritized efficiency and profitability, often overlooking environmental and societal implications. However, given the global challenges like climate change and resource scarcity, there is a growing emphasis on embedding sustainability into process design. Adopting a systems-oriented approach provides a comprehensive view, spanning from raw material acquisition to end-of-life product management. Such an approach not only identifies potential sustainability challenges but ensures that solutions foster both environmental responsibility and economic viability. In this study, a comprehensive framework for designing industrial systems is introduced, aiming to encompass the entire lifecycle impacts of chemical processes. The research initially delves into two end-of-life scenarios: solvent recovery (as a pollution reduction intervention) and wastewater treatment systems (as a pollution control intervention). Employing graph-theoretical methods and multi-objective optimization, a thorough systems analysis which incorporates Ecological footprint and Emergy analysis, coupled with economic assessment is presented. Furthermore, a Machine Learning (ML) model (as a source reduction option) is developed to predict the cradle-to-gate impacts of chemicals. Merging the insights from this ML model with the end-of-life scenarios offers a comprehensive systems strategy, advocating for a sustainability-focused approach during the early stages of process design.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

List of Figures

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

List of Tables

Chapter 1

Introduction

Text and figures are reproduced and adapted with permission from Aboagye, E. A.; Chea, J.D.; Yenkie, K.M.; Systems Level Roadmap for Solvent Recovery and Reuse in Industries – A Review. *iScience* 2021 https://doi.org/10.1016/j.isci.2021.103114, and materials in preparation for publication.

1.1 Motivation

The increasing awareness of environmental issues has led industries to reassess their operations in light of sustainability. Industrial processes, notably solvent recovery [1] and wastewater treatment [2], are areas of growing concern given their significant environmental footprints. Recognizing this, the body of literature related to the sustainability assessment of industrial processes has grown exponentially over the past few decades (see Figure 1).

The chemical process industry is a significant contributor to modern life, producing a diverse array of products such as plastics, fuels, pharmaceuticals, and petrochemicals [3]–[8]. Therefore, the concept of sustainability in industrial processes is multifaceted, encompassing economic, environmental, and social dimensions [9], [10]. Traditionally, industrial processes have primarily focused on economic profitability. However, the call for sustainable development and recent climate change crises have prompted industries to consider environmental impacts alongside economic performance. Consequently, decision-makers in industrial processes now face the daunting task of simultaneously optimizing a multitude of often-conflicting objectives. This complexity has led researchers

and practitioners alike to seek out innovative approaches to sustainability assessment and decision-making in industrial processes.

As illustrated in Figure 1, in recent years there has been a marked increase in research focused on the sustainability assessment of chemical process design.

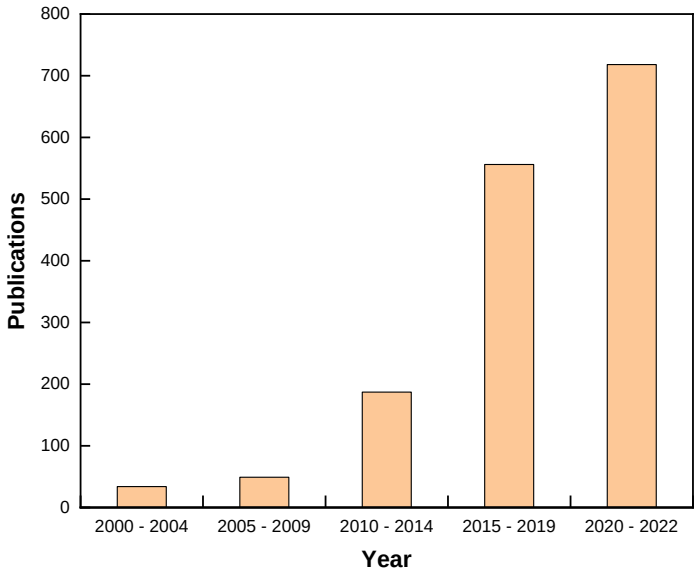


Figure 1. Web of Science (WoS) publications on sustainability assessment of chemical process design

Given that many chemical industries were constructed several years ago, prioritizing cost with no consideration for environmental impacts was the main driving factor. However, Environmental Impacts Assessment (EIA) [11], [12], which is a key aspect of sustainability [13] where the potential environmental consequences of a process or product is evaluated is crucial to achieving greenness of industrial processes. These environmental consequences take into consideration several factors such as air-, water-, soil- quality, biodiversity, and social factors. To perform any EIA, Life Cycle Assessment

(LCA) is the prevailing methodology utilized to quantitatively measure the environmental consequences of a product, process, or service [14]. However, performing LCA of processes is very challenging, especially at early-stage process synthesis.

This dissertation provides an in-depth exploration of the application of graph theory, optimization, and machine learning techniques for the sustainability assessment of industrial processes. In an era where industries are facing increasing pressure to reduce environmental impacts and promote sustainable practices, a comprehensive understanding of the trade-offs between economic viability and environmental sustainability is crucial. Through a set of multidisciplinary methodologies, this research aims to provide robust decision-making tools to assess, design, and operate more sustainable industrial systems. The methodologies developed can be used to recover valuable resources from waste streams and further facilitate EIA at early-stage process synthesis. Thus, the hypothesis for this dissertation is that ***computational methods aid in early-stage process design for efficient systems with low cost and enhanced sustainability***. Three main objectives are addressed in this dissertation to achieve the specified hypothesis namely, (1) *wastewater treatment network design via simultaneous cost and sustainability assessment approach*, (2) *multi-objective approach to solvent recovery systems design*, and (3) *predicting life cycle impacts of chemicals – a machine learning approach*.

1.2 Background

The United Nations Commission on Environment and Development describes sustainability as fulfilling current needs without jeopardizing the capability of future generations to address their own requirements [15]. However, the 2022 report from the United Nations Environment Programme (UNEP) suggests that global efforts are trailing

behind in maintaining the desired global temperature rise below 2˚C, and ideally, 1.5˚C, which is crucial for mitigating the effects of climate change [16]. To meet the targets set by the Paris Agreement [17], the UNEP report emphasizes a 30% reduction in greenhouse gas emissions. As industries evolve, there is a pressing need to design new processes with a sustainability-focused approach. It is essential for the chemical sector to take a lead role, exploring ways to significantly reduce emissions, aiding global efforts to counteract climate-related challenges. Life Cycle Assessment (LCA) is one of the ways to quantify the environmental footprint of processes, products, or services.

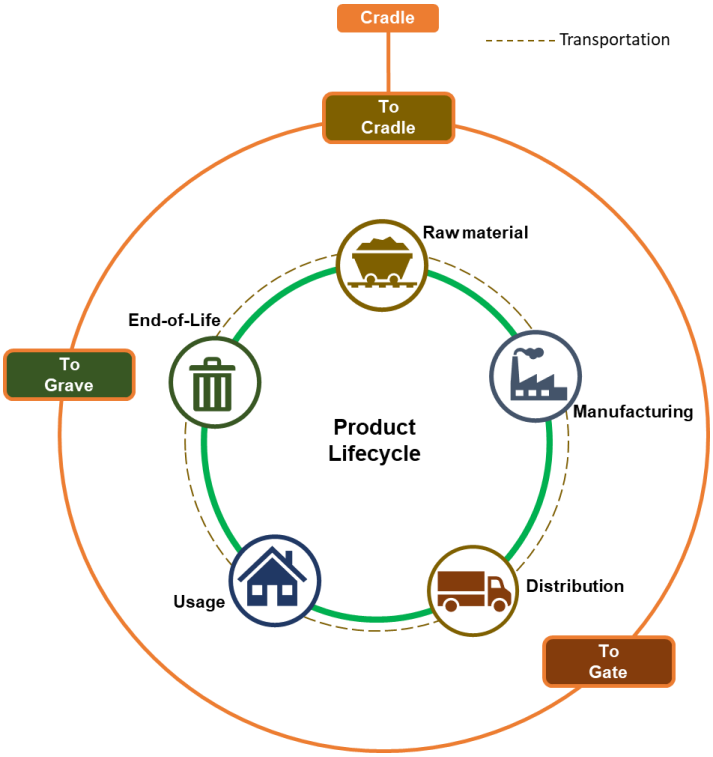


Figure 2. The product life cycle comprises various stages, each of which involves material transportation. LCA can be conducted using one of three primary models: cradle-to-gate, cradle-to-grave, and cradle-to-cradle

In a typical product life cycle, five distinct stages can be identified: (i) raw material extraction, (ii) product manufacturing, (iii) distribution, (iv) usage, and (v) disposal as shown in Figure 2. The disposal phase marks the end-of-life of the product and various scenarios can be considered, such as recycling, reuse, or disposal to landfills. Conducting a Life Cycle Assessment (LCA) generally involves four fundamental steps, which include goal and scope definition, Life Cycle Inventory (LCI) analysis, impact assessment, and interpretation [14], [18]–[21]. Depending on the scope of the LCA, various analysis frameworks can be implemented, namely, cradle-to-gate, cradle-to-grave, gate-to-gate, or cradle-to-cradle. As highlighted in Figure 2, cradle-to-gate analysis starts from raw-material extraction to product distribution. Cradle-to-grave assessment ends at the disposal phase.

1.2.1 Resource Recovery: End-of-Life Scenarios

In an era of limited resources and increasing environmental concern, resource recovery [22] has emerged as a vital strategy to ensure sustainability. Thus, waste is no longer considered as an endpoint but as a valuable resource for recovery and reuse of key chemicals. This paradigm shift promotes circular economy [23], where the waste goes through various transformations into a useful resource.

Resource recovery refers to the extraction of useful materials and energy from waste. The goals are to reduce the burden on natural resources, mitigate environmental impact, and achieve economic benefits from waste valorization. This approach aligns with the principles of circular economy, which aims for a closed-loop system minimizing waste and making the most of resources.

There are different types of End-of-Life (EoL) for various products [24]. For example, if we consider solvents, their typical EoLs include incineration, selling to third-parties, or on- or off-site disposal. Furthermore, these solvents can also end up in wastewater streams based on their use. The solvents can be recovered through various techniques such as distillation, adsorption, membrane separation, extraction, among others.

Implementing resource recovery as an EoL scenario faces numerous challenges. These include technical barriers, economic feasibility, regulatory issues, and public acceptance [25], [26]. Technical challenges include the diverse nature of waste, which requires tailored treatment and recovery solutions. Economically, the initial investment and operating costs of recovery facilities may be high compared to the use of virgin materials. However, the cost can often be offset by the value of the recovered resources and reduced waste disposal costs. On the regulatory front, stringent environmental standards demand highly efficient recovery processes to ensure that the effluents meet the required specifications.

Nevertheless, the opportunities offered by resource recovery are immense. Besides the environmental benefits, resource recovery can significantly improve resource efficiency, enhance energy security, and create job opportunities. Furthermore, the integration of advanced digital technologies, such as machine learning, can enhance the efficiency and adaptability of recovery processes [27]. Wastewater treatment and solvent recovery as EoL scenarios represent significant strides towards a circular economy. These approaches not only mitigate environmental impact but also unlock the value of waste as a resource.

1.2.2 Wastewater Treatment Networks

Wastewater treatment process is indispensable to ensuring environmental safety and public health, reducing the potential impacts of hazardous substances in the ecosystem. Wastewater treatment networks (WWTNs) use a systematic series of process units to treat and recover resources from domestic/industrial/agricultural wastes [28]. These treatment networks are designed to remove or neutralize contaminants and produce effluents that can be safely released into the environment or used in secondary applications. Additionally, some of these networks can recover valuable substances, such as solvents, from the waste streams [29].

WWTNs vary significantly depending on the wastewater characteristics. Despite the effectiveness of conventional WWT processes, they are not always capable of fully removing certain chemicals, such as active pharmaceutical ingredients (APIs), endocrine disruptors, or emerging contaminants such as per- and polyfluoroalkyl substances (PFAS).

The design of these networks is a complex task requiring the consideration of various factors, such as the type and concentration of contaminants, regulations, recovery potential, economic viability, and environmental impacts [30], [31]. Therefore, the field of process systems engineering (PSE) plays a pivotal role in creating efficient, effective, and adaptable wastewater treatment networks [32], [33]. Moreover, the selection and sequencing of technologies, the balance between recovery and treatment, and the overall network layout require meticulous planning and optimization. In this context, the field of graph theory [34] provides a robust framework for WWTN design. It enables the systematic enumeration of all possible networks and shortlist the feasible ones when

subject to certain input conditions and recovery requirements thereby facilitating the optimization of cost, resource recovery, and sustainability metrics.

However, the use of wastewater treatment networks as an EoL option is not without challenges. Technically, the diverse nature of industrial waste makes it challenging to develop a one-size-fits-all solution [35]. Regulatory-wise, stringent environmental laws necessitate that effluents meet certain quality standards, demanding highly efficient treatment processes. Nevertheless, the potential benefits offered by wastewater treatment networks as an EoL option for even solvents, are driving research and development efforts in this direction. The wastewater treatment network as an EoL scenario offers a sustainable solution to industrial waste management. Its successful implementation necessitates advanced process engineering and sustainability assessment approaches to optimize the performance, economics, and environmental impact of these networks [2]. With growing sustainability concerns, such initiatives will undoubtedly play a significant role in the future of industrial waste management.

1.2.3 Solvent Recovery

Solvents are indispensable to a wide array of industrial procedures, finding significant usage in industries such as pharmaceuticals, food, cosmetics, nutraceuticals, biofuels, paints, and fine chemicals [36]–[38]. Notably, the pharmaceutical industry utilizes solvents extensively for active pharmaceutical ingredient (API) purification and refinement [39], [40]. The ever-increasing demand for solvents has, however, led to a substantial surge in waste generation. For instance, the pharmaceutical industry generates approximately 25-100 kg of waste per kg of product [41]. This waste issue primarily stems from industrial process inefficiencies and flawed solvent selection criteria, often leading to

excessive solvent usage for achieving the desired product purities and quantities. Traditional waste management approaches in many industries have largely revolved around incineration, offsite, and onsite disposal techniques. However, these strategies pose significant challenges concerning the emissions produced, safety protocols, waste solvent handling, and the subsequent impact of these solvents on the ecosystem. For example, the annual Disability-Adjusted Life Years (DALY) linked to transportation for offsite disposal is estimated to fall between 0.35 and 35.03 [42]–[44], highlighting the considerable human health implications associated with these conventional disposal methods. Given the escalating trends in waste solvent generation, there is a pressing need for process intensification methods, such as solvent recovery, to address the growing environmental, health, and safety concerns.

Different aspects exist for a systems level solvent recovery design. Figure 3 shows the various aspects that can contribute to an efficient design. Economics is the main driving factor for most design problems. However, other factors such as process synthesis routes, solvent waste characteristics, environmental impact assessment and quality by design or control, can have a significant impact on how solvent recovery systems are designed efficiently, as shown in Figure 3.

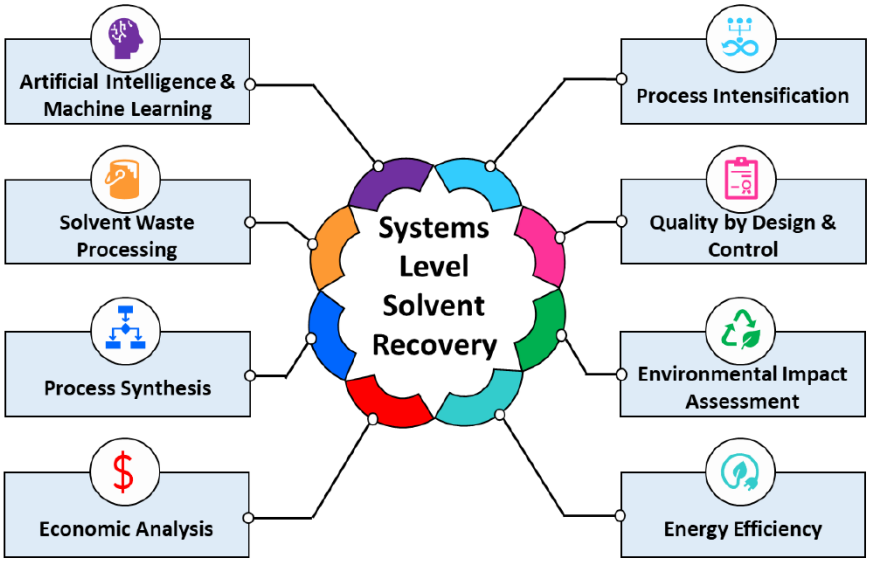


Figure 3. Various aspects that can contribute to an efficient design of solvent recovery processes

For example, in the pharmaceutical space, Quality by Design (QbD) is of paramount importance [45]. Therefore, practicing QbD means finding alternative ways to recover the solvent for reuse rather than resorting to conventional treatment methods.

1.2.4 Process Systems Engineering and Sustainability Assessment

Future demand for process design research is anticipated to be influenced not only by traditional profit-driven motives, but also significantly by initiatives focused on sustainability. Thus, leveraging Process Systems Engineering (PSE) in the sustainability assessment of chemical processes is critical, offering a unique perspective towards addressing Sustainable Development Goals (SDGs).

PSE deals with the design, operation, and optimization of chemical and industrial processes [3], [46]. It employs mathematical and statistical models, computational

algorithms, and control strategies for efficient process development, thus, improving productivity and reducing cost. However, the confluence of PSE and sustainability assessment provides a robust blend of multifaceted approach to sustainable development.

PSE provides the technological and operational strategies for enhancing process efficiency through optimization strategies, thus helping to minimize the environmental footprint of the process. This synergistic association enables a systems-level perspective that helps to consider the entire life cycle of the process – from resource extraction to end-of-life. Thus, through PSE, the often-conflicting interest between economics and sustainability assessment can be solved via a multi-objective optimization [47] problem formulation approach. Therefore, in this work, we leverage the power of PSE to solve multi-objective optimization problems within the areas of wastewater treatment network and synthesis of solvent recovery pathways.

1.2.5 Role of Machine Learning in Sustainability Assessment

The inception phase of process design, referred to as early-stage design, involves the preliminary examination undertaken after the delineation and detailing of the product [48]. This phase is distinguished by a greater degree of freedom attributed to the copious process options available [49]. The abundance of these options can be ascribed to the primary objectives of early-stage process synthesis, which predominantly include the identification of various raw materials along with their intermediate compounds, determination of process conditions, and selection of suitable equipment. Further characterizing this stage involves a considerable degree of experimentation and the frequent use of trial-and-error techniques, necessitated by the need to refine and test different synthesized routes [46], [50]. Concurrently, this phase of design is marked by the

limited availability of data [51], making the assessment of process sustainability a formidable challenge.

Recent developments within the chemical process sector have underscored the necessity for the incorporation of sustainable design principles as highlighted earlier. Instead of treating sustainability as merely a constraint to be optimized, it is progressively being viewed as a goal to be achieved. This transition towards a goal-oriented approach is crucial, as the consideration of sustainability issues during the early design stages aids in the synthesis of processes that are more efficient. This, in turn, minimizes waste and emissions, promotes the use of renewable energy sources, and reduces the overall environmental impact of the manufacturing process.

Over the past two decades, considerable research has been dedicated to the application of Machine Learning (ML) techniques to enhance the sustainability of chemical processes. As evidenced by Figure 4, there has been an exponential increase in the volume of publications concerning the application of ML to chemical process design. However, it is noteworthy that its application in the context of early-stage design has not been adequately explored. This discrepancy emphasizes the need for further research in this domain to facilitate the design of innovative and environmentally friendly systems within the chemical process space.

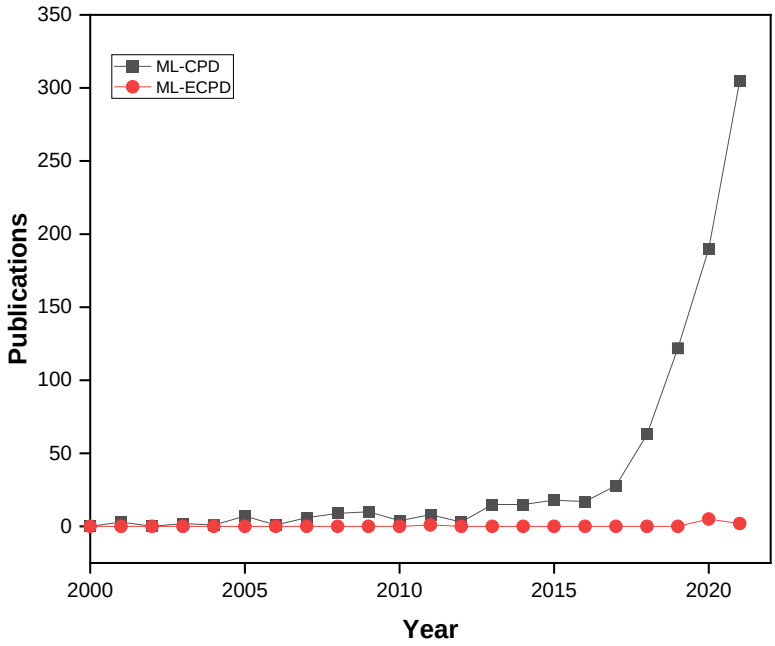


Figure 4. Publication contributions for Machine Learning for Chemical Process Design (ML-CPD) and Machine Learning for Early-stage Chemical Process Design (ML-ECPD) from Web of Science

1.3 Synergistic Approach of Wastewater Treatment, Solvent Recovery, and ML for LCIA Predictions

Wastewater treatment, fundamentally, is an exercise in pollution control. It involves the removal of pollutants and contaminants from wastewater before it is released back into the environment [28], [52]. This process is crucial in preventing waterborne pollutants from damaging aquatic ecosystems and endangering public health. By treating wastewater effectively, we are not just addressing the symptoms of environmental degradation but actively working towards maintaining the ecological balance. This aspect of environmental management is vital in controlling the direct impacts of industrial and domestic activities on natural water bodies.

Solvent recovery, in contrast, is a proactive approach that aligns with the principle of pollution prevention. This process involves capturing and reusing solvents used in industrial processes, thereby reducing the demand for new solvents and minimizing solvent-related waste [1], [39], [53]. By recovering and reusing solvents, industries can significantly decrease their environmental footprint. This approach not only reduces the volume of hazardous waste but also conserves resources and energy that would otherwise be expended in the production of new solvents. Solvent recovery exemplifies a shift from a traditional, linear economic model to a more sustainable, circular model where resource efficiency and waste reduction are prioritized.

The integration of machine learning in predicting Life Cycle Impact Assessments (LCIA) marks a paradigm shift towards source reduction. Machine learning algorithms [54]–[56] can analyze complex datasets to predict the environmental impacts of products or processes over their entire life cycle. This predictive capability is instrumental in identifying potential environmental hotspots and facilitating source reduction strategies. By leveraging machine learning, industries can optimize their processes to minimize resource use and waste generation from the very beginning. This approach aligns with the principle of source reduction, which is the most effective way to minimize environmental impact.

The unification of these three objectives - wastewater treatment as pollution control, solvent recovery as pollution prevention, and machine learning for LCIA predictions as source reduction - creates a comprehensive and synergistic approach to environmental management. Each element complements the others, forming a multi-layered strategy that addresses environmental issues at different stages:

1. Preventive measures: Solvent recovery and machine learning for LCIA predictions work at the upstream level, preventing pollution before it occurs. Solvent recovery reduces the generation of hazardous waste, while machine learning helps in designing processes and products that are environmentally friendly from the outset.

2. Control measures: Wastewater treatment acts as a control measure, dealing with pollutants that have already been generated. This is crucial for mitigating the immediate impacts of industrial and domestic effluents on the environment.

3. Data-driven decision making: The use of machine learning in LCIA predictions also facilitates data-driven decision-making, enabling industries and policymakers to evaluate the long-term impacts of their actions and make informed choices about resource utilization and waste management.

Thus, the integration of wastewater treatment, solvent recovery, and machine learning for LCIA predictions represents a holistic approach to environmental management. This unified strategy not only addresses the immediate challenges of pollution control but also fosters a culture of pollution prevention and source reduction. By adopting this integrated approach, we can significantly enhance our efforts towards sustainable development and environmental conservation.

1.4 Thesis Structure

This thesis is organized into four main parts. In the first part of this research (Chapter 3), the application of superstructure/maximal structure-based optimization techniques is used to delve into the sustainability assessment of wastewater treatment networks using the P-graph framework. The complex nature of these treatment networks requires an intricate balance between treatment effectiveness, economic feasibility, and

environmental impact. Our approach enhances the process synthesis methodology with a graph-theoretic approach, namely P-graph, that is used to synthesize the wastewater treatment networks. The resulting feasible and optimal structures are then individually solved using non-linear programming algorithm in GAMS, a high-level modeling system for mathematical optimization. This combination of methods results in a more detailed and holistic view of the environmental impacts of the network, leading to better-informed and more sustainable design and decision-making.

Chapter 4 continues by exploring waste solvent recovery processes, a critical segment of industrial operations with significant sustainability potential. A superstructure-based optimization model is developed that integrates economic and environmental factors, delivering a comprehensive sustainability assessment of these processes. This model allows the systematic exploration of feasible process alternatives and pathways, revealing optimal configurations under a diverse set of operational, economic, and environmental constraints.

Chapter 5 takes a direction by employing machine learning techniques to predict life cycle impact assessment metrics, such as human health and global warming impacts, of chemicals during early-stage process synthesis. These metrics are pivotal to measuring the sustainability of industrial processes over their entire lifespan. By leveraging machine learning, forward-looking assessments can be made that will help industries and policymakers to forecast the long-term environmental consequences of their operations, thereby promoting sustainability.

The last part of the research (Chapter 6) presents the summary and conclusion of the key findings. It further gives some future work direction in the space of wastewater treatment and machine learning for better systems design and sustainability assessment.

Thus, this thesis represents interconnections between process systems engineering principles, and machine learning to address sustainability challenges. It delivers a comprehensive framework for sustainability assessment in industrial processes, illuminating the complex trade-offs between economic performance and environmental sustainability. This work seeks to make significant strides in the broader sustainability discourse, particularly as it relates to industrial operations.

Chapter 2

Materials and Methods

2.1 Systems Design Approach

In the dynamic and complex world of process design, the systems approach has emerged as a transformative paradigm, offering a comprehensive framework for addressing the multifaceted challenges of designing and managing processes [49], [57]. A systems approach is a methodology that views a process not as a mere aggregation of discrete parts, but as an integrated whole [58]–[61]. This perspective is grounded in the understanding that the performance and efficacy of a process are not solely dependent on its individual components, but rather on the interactions and relationships between these components. The approach is inherently holistic, emphasizing the importance of understanding the entire system, including its environment, objectives, and the complex interplay of its various elements [58].

One of the core principles of the systems approach is interconnectivity [58]. In process design, this principle mandates a recognition of the intricate web of dependencies and interactions within the system. Each component or operation within the process is seen not in isolation, but as part of a larger network, where changes or disturbances in one area can have cascading effects throughout the system. This interconnected perspective is crucial for predicting and managing the implications of process modifications, ensuring that optimizations in one part do not inadvertently lead to inefficiencies or problems in another.

Another critical aspect of the systems approach is its emphasis on the end-to-end process flow [57], [58]. Unlike traditional methods that might focus on optimizing individual operations or stages, the systems approach seeks to enhance the overall process flow, ensuring smooth and efficient transitions between different stages and minimizing bottlenecks. This approach is particularly advantageous in complex processes where multiple operations must be seamlessly integrated to achieve the desired output.

The systems approach also inherently incorporates a degree of flexibility and adaptability [59]. In a constantly evolving business and technological landscape, processes must be designed with an eye toward future changes and challenges. The systems approach allows for this adaptability, enabling processes to evolve and scale in response to new demands, technological advancements, or shifts in the market.

Furthermore, this approach fosters a culture of continuous improvement and innovation. By encouraging a broad, holistic view of the process, it opens up opportunities for innovative solutions that might be overlooked in a more narrow, segmented approach. This culture of innovation is crucial for maintaining competitiveness and efficiency in a rapidly changing world. Therefore, in this work, we leverage the idea of systems design to develop and solve mixed integer non-linear programming (MINLP) problems.

2.2 Sustainable Process Index (SPI) Methodology for Sustainability Assessment

Ecological footprints offer a comprehensive representation of the impact of a process on the land area. The utilization of land area as a metric provides meaningful insights when interpreting result outcomes. The principal advantage of implementing an ecological footprint is the ability to quantify the environmental load and stress associated with the process. Typically, ecological footprints take into account emissions affecting air,

water, and soil in relation to the process under review. Therefore, most ecological assessments inherently incorporate human health impacts.

The SPI [62]–[64], which is an ecological footprint, is able to map all these processes to land area, with the basic idea being that processes that uses larger land areas are less sustainable for the same product objective. The main advantage of using SPI is its ability to quantify the ecological pressures of the process, which is an often-neglected factor when performing sustainability assessments. Furthermore, the SPI helps tackle the insufficient interdependencies and relations between society and nature by presenting the effect of anthropogenic activities on the ecosystem and also justifies the fact that the available land area on earth is limited and must be used judiciously. Additionally, social impacts of the process are considered since the SPI inherently takes into consideration embedding the related process emissions into the ecosystem. Thus, the SPI presents a better way of relating the environment to human activities.

There are seven main land areas considered in quantifying the SPI of a process or product. These are (1) area needed for raw material production, (2) area needed for the energy production, (3) area needed for installation of equipment, (4) area needed for staff accommodation, and areas needed for embedding (5) water emissions, (6) air emissions, and (7) soil emissions. The summation of these partial areas gives the total arable area needed for that process. The areas can be grouped into two main parts namely: input areas and output areas, as shown in Figure 5.

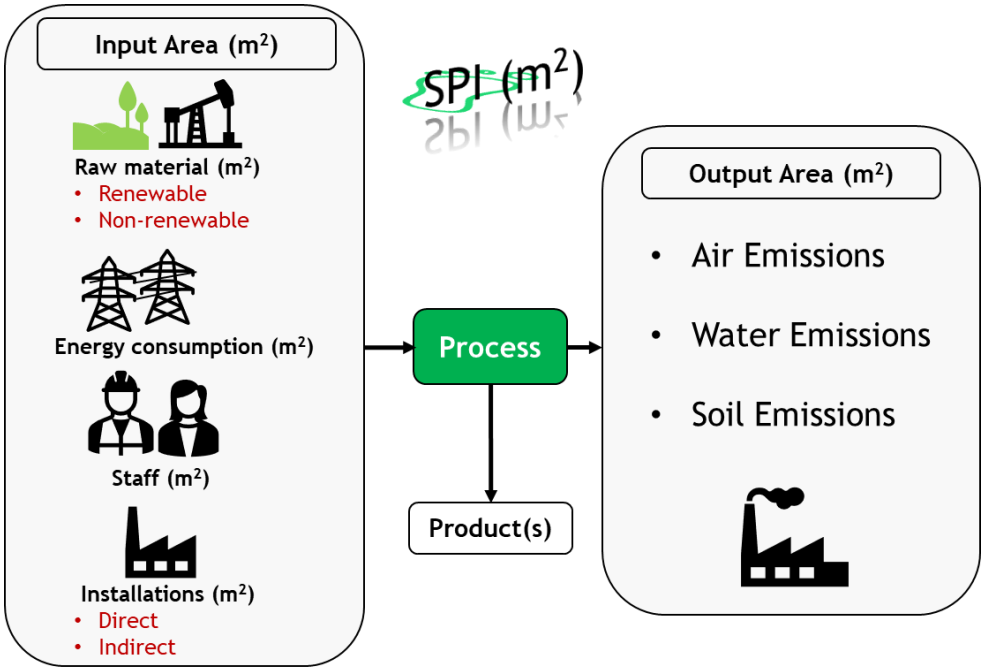


Figure 5. The various aspects to quantifying the sustainable process index

$$A_{tot} = A_R + A_E + A_I + A_S + A_p \quad (1)$$

Equation (1) shows the basic equation for calculating the SPI value, where A_R is the area needed for raw material production, A_E is the area needed for energy consumption, A_I is the area needed for installation of equipment, A_S is the area needed for to accommodate staff, A_P is the area needed to dissipate the air ,water and soil emissions sustainably into the ecosystem, and A_{tot} is the total area.

SPI is tailored for assessing industrial processes. It is particularly relevant when the goal is to evaluate and improve the sustainability of specific manufacturing processes or industrial activities. This approach is especially beneficial in industries where process efficiency and environmental impacts are critical considerations. In contrast, other

ecological footprints focus broadly on general consumption patterns, urban development, and national policies which might not provide the detailed insights required for process improvements in the industrial settings. Furthermore, SPI evaluates multiple environmental aspects of a process, including energy use, raw material consumption, emissions, direct and indirect installation of process equipment. This breadth ensures a holistic view and understanding of the environmental impacts. By covering such a wide array of environmental factors, SPI enables industries to identify and mitigate their most significant environmental impacts, rather than focusing on a single aspect like carbon footprint or water usage. Additionally, SPI provides a clear and process-specific tool for evaluating and improving industrial sustainability. Such specificity makes SPI an effective tool for developing industry-specific regulations and guidelines, as well as for companies looking to align with or even exceed environmental standards. Lastly, the SPI methodology can be customized and adopted to different industry-types or processes. This flexibility allows for more relevant and accurate assessments across various sectors, hence ensuring that the unique environmental challenges and opportunities of different industries are accurately represented and addressed.

In all SPI analysis in this thesis, Equation 1 is foundational model used to assess the ecological pressure of the system under consideration. Detailed mathematical models and parameter estimations for each component of the SPI is discussed in Chapter 3 and Appendix A.

2.3 Emergy Analysis Methodology for Sustainability Assessment

Ecological indicators provide valuable insight into the impact of human activities on the environment. However, they are often inadequate for accurately estimating the

energy demands of processes, a critical metric in chemical processes. Therefore, the integration of indicators that measure energy use is essential. Emergy, a quantitative methodology introduced by Odum in 1988 [65]–[67], has seen significant application over the years. This technique quantifies the flows of materials, energies, currencies, and services involved in the production of a product, and represents these flows as solar energy equivalents or solar-emjoules (sej).

Emergy represents the cumulative energy, in its various forms, required to bring about a transformation process that results in a particular product [68]. In entropy terms, Emergy can be viewed as a measure of the produced entropy along the entire supply chain of a process. The concept of Emergy is based on the principle that all forms of energy can be traced back to the solar energy that sustains life and processes on earth [69]. Hence, any product, service, or system, can be assigned an 'emergy value' that signifies the equivalent solar energy consumed throughout its production or supply chain. Figure 6 shows the various steps involved in Emergy accounting.

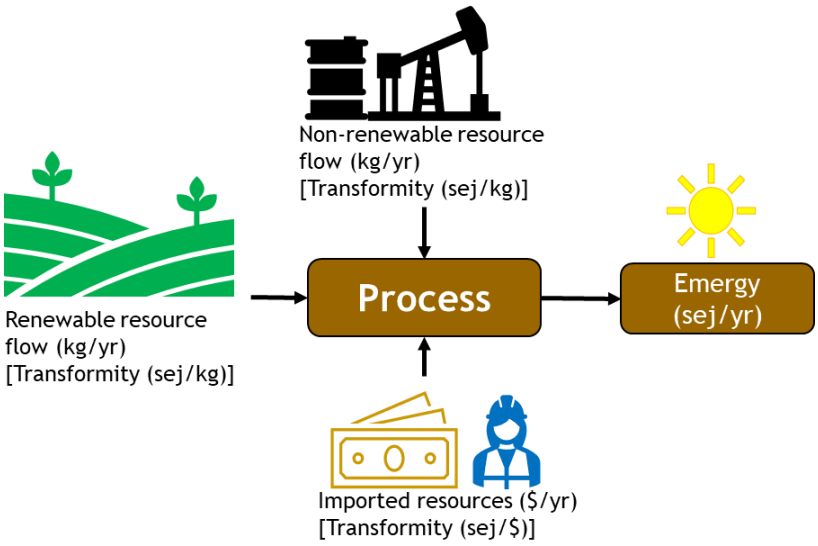


Figure 6. Various aspects to Emergy accounting

As shown in Figure 6, Emergy analysis delivers a quantified assessment of three significant footprints: Renewable Natural Resources (R), Non-Renewable Natural Resources (N), and Imported Resources (F). These footprints facilitate the calculation of Emergy Yield Ratio (EYR), Environmental Loading Ratio (ELR), and the Emergy Sustainability Index (ESI) [70]. The sum of Emergy illustrates the wealth of a system.

One major advantage of using this methodology is its ability to encapsulate both the qualitative and quantitative aspects of energy use [71]. It transcends beyond traditional life cycle assessment methods and economic evaluation by integrating all the direct and indirect energy flows contributing to a product or system. This helps to provide a comprehensive analysis of environmental impacts, especially in the area of resource utilization. Additionally, it is able to provide more accurate ‘energy hotspots’ in a process, especially in areas where energy usage is extremely high. Another significant advantage of

Emergy analysis is its ability to integrate economic and ecological evaluations. By converting different forms of energy and resources into a common unit (emjoules), Emergy facilitates an integrated assessment of both economic costs and ecological impacts. This harmonization is essential for formulating strategies that are not only economically viable but also ecologically responsible, thus addressing the often-competing interests of economic development and environmental preservation. Lastly, Emergy analysis is instrumental in promoting the principles of the circular economy. It highlights the importance of recycling and efficient resource use, identifying processes where waste minimization and resource reuse or recycling can be enhanced. This contributes to a more sustainable and circular approach to production and consumption, reducing waste and maximizing resource efficiency.

While Emergy analysis offers a significant benefit, it should not be used in isolation. It must be complemented with other environmental impact assessments indicators for a comprehensive picture of the sustainability of the process; hence, its integration with SPI in this work.

2.4 Multi-Objective Optimization

Optimization, as a critical aspect of decision-making and planning, traditionally focused on single-objective scenarios, where the aim was to find the best solution from a pool of feasible alternatives concerning one objective [47], [72]. However, real-world problems often require the simultaneous optimization of multiple, often conflicting, objectives. Such scenarios give rise to multi-objective optimization (MOO) [47], a subfield of optimization that addresses problems involving multiple objectives.

MOO encapsulates the essence of trade-off, compromise, and balance between conflicting objectives. The conflicting nature of the objectives means that improving one objective may worsen others, leading to a set of multiple optimal solutions, referred to as Pareto-optimal solutions. The Pareto-optimal front represents the list of non-dominated solution sets that achieve the best trade-offs among the objectives, with no solution being superior to any other without sacrificing at least one objective. There are several methodologies available for solving multi-objective optimization problems. The choice of a specific approach often depends on the problem context and the nature of the decision variables, constraints, and objective functions.

Some of the most prevalent methodologies include the weighting method, the ε-constraint method, goal programming, and evolutionary algorithms [47], [72].

2.5 Machine Learning for Sustainability Assessment

In light of the escalating complexity and multifarious nature of environmental challenges, it is crucial for Environmental Impact Assessment (EIA) professionals to have access to precise and high-performance tools that bolster their analytical abilities. These tools are instrumental in facilitating exhaustive assessments, enabling effective communication of outcomes, and promoting data-driven decision-making pertaining to the environmental impacts associated with industrial procedures. To cater to this demand, a plethora of modeling software platforms have been innovatively devised, each possessing distinct attributes and functionalities specifically designed for the execution of Environmental Impact Assessment and Life Cycle Assessment (LCA) analyses [73]. Some of the most established software platforms are SimaPro®, OpenLCA, Umberto, EcoChain, and GaBi [74], [75].

However, while these traditional methods are valuable for impact assessment, there are numerous limitations and drawbacks when applied to early-stage process design. The dynamic nature of early-stage process design, which is characterized by frequent changes and iterations makes impact assessment challenging [49]. Furthermore, these traditional methods focus on a comprehensive analysis, however, during the design phase, it may be more beneficial to identify and prioritize the most important environmental hotspots or improvement opportunities [76].

ML is primarily engaged in formulating algorithms and statistical models that empower computers to learn and formulate predictions, without explicit programming [77]. The main objective of ML systems is to create models that can generalize from known examples to make predictions and decisions in previously unseen situations [78], [79]. The three main types of ML systems are supervised learning, unsupervised learning, and reinforcement learning. Each type addresses different learning scenarios and problems and can be combined with other types for better predictions.

The following chapters in this dissertation will use either of these strategies or a combination of them for sustainability assessment.

Chapter 3

Synthesis of Wastewater Treatment Networks via Simultaneous Cost and Sustainability Assessment Approach

Text and figures are reproduced and adapted with permission from Aboagye, E. A.; Burnham, S. M.; Dailey, J.; Zia, R.; Tran, C.; Desai, M.; Yenkie, K. M. Systematic Design, Optimization, and Sustainability Assessment for Generation of Efficient Wastewater Treatment Networks. *Water* 2021, *13* (9), 1326. https://doi.org/10.3390/w13091326; Pimentel, J., Aboagye, E., Orosz, Á., Markót, M.C.,Cabezas, H., Friedler, F., Yenkie, K.M., 2022. Enabling Technology Models with Nonlinearities in the Synthesis of Wastewater Treatment Networks based on the P-graph Framework. Computers & Chemical Engineering 108034. https://doi.org/10.1016/j.compchemeng.2022.108034; and Aboagye E.A., Pimentel J., Orosz Ákos, Cabezas H., Friedler F., Yenkie K.M., 2021, Efficient Design and Sustainability Assessment of Wastewater Treatment Networks using the P-graph Approach: A Tannery Waste Case Study, Chemical Engineering Transactions, 88, 493-498, https://doi.org/10.3303/CET2188082

3.1 Background

Minimizing wastewater (WW) generation is a crucial aspect of pollution prevention. WW streams usually contain various hazardous pollutants that must be reduced to acceptable concentrations or limits before disposal. As water has no substitutes, there is an urgent need to treat WW for reuse, proper disposal, or recycling. In the past two decades, researchers have explored the systematic design and optimization of wastewater treatment (WWT) plants or networks [2], [33], [35], [80]–[83]. However, a comprehensive approach

for designing efficient WWT networks is still lacking [31] in regards to sustainability assessment.

The effectiveness of wastewater treatment processes depends on proper design, operation, and maintenance of treatment facilities. The environmental impacts of these processes include energy consumption and greenhouse gas emissions associated with the operation of wastewater treatment plants. Typically, for domestic wastewater treatment, the treatment process involves several steps, each designed to reduce different contaminants. The process starts with removal of debris, sand and grits. The primary settler and activated sludge are typically used for the treatment of such wastewater [52].

During the primary settling step, there is removal of solid contaminants. Wastewater is held in large basins where heavy solids can settle to the bottom, while oil, grease, and lighter solids float to the top. These substances are then physically removed. Primary settling is effective in removing large, suspended solids but not the dissolved organic and inorganic pollutants [28], [35], [52].

The secondary treatment step is primary biological, and targets dissolved organic matter that escapes the settling stage. The activated sludge process is a widely used method in this stage. It involves aerating the wastewater to encourage the growth of bacteria and other microorganisms. These microorganisms consume organic matter, effectively reducing the organic content of the wastewater. The mixture then goes to another settling tank to remove the microorganisms from the treated water [31], [84]. Further treatment is necessary to remove nutrients like nitrogen and phosphorus, which can cause eutrophication in water bodies, and such is accomplished using advanced treatment methods which include biological nutrient removal, such as filtration, or chemical

treatments. The treated water is usually disinfected to kill any remaining pathogens. Common disinfection methods include chlorination and ultraviolet light treatment.

The tertiary wastewater treatment stage, also known as advanced wastewater treatment [33], represents the final cleaning process that improves wastewater quality before it is reused, recycled, or discharged into the environment. This stage is particularly crucial in the treatment of domestic wastewater as it focuses on the removal of contaminants that primary and secondary treatments may not fully address. The primary objective of tertiary treatment is to elevate water quality to meet specific standards required for its intended final use, which may include agricultural irrigation, industrial processes, or replenishment of natural water bodies. A central aspect of tertiary wastewater treatment for domestic wastewater is the removal of nutrients, particularly nitrogen and phosphorus, which are significant contributors to eutrophication in water bodies [28], [52]. Eutrophication, the enrichment of water by nutrients, can lead to excessive growth of algae and other aquatic plants, disrupting ecosystems and degrading water quality [85], [86].

The process of nitrogen removal typically involves biological nitrification and denitrification [28], [52]. Nitrification is a two-step aerobic process where ammonia is first converted to nitrite and then to nitrate. Subsequently, in the anoxic conditions of denitrification, these nitrates are converted into nitrogen gas, which is released into the atmosphere. This biological process is often supplemented with physical and chemical methods such as ion exchange and membrane filtration to ensure effective removal.

In the case of phosphorus, it is generally removed through chemical precipitation [28], [52]. Chemicals like alum or iron salts are added to the wastewater to form insoluble compounds with phosphorus, which then precipitate out of the water. Biological

phosphorus removal can also be employed, where specific bacteria absorb phosphorus in excess of their metabolic needs and are subsequently removed from the water as part of the waste sludge. Figure 7 shows the traditional flowchart for municipal wastewater treatment where the technologies are predefined for the treatment process.

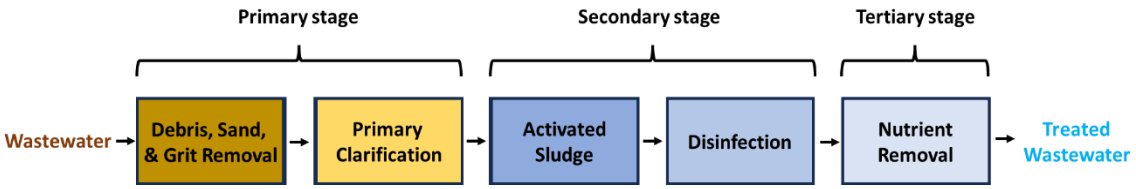


Figure 7. Flowchart showing the typical steps involved in municipal wastewater treatment

Another area that is increasingly being recognized for its potential is desalination [87]–[89] processes for wastewater treatment. In areas grappling with water scarcity, desalination offers a promising solution to augment water supplies by transforming treated wastewater into a resource suitable for various uses. The technology, while primarily aimed at salt removal, effectively eliminates a wide range of contaminants, making it an integral component of advanced wastewater treatment strategies.

The cornerstone of desalination in wastewater treatment lies in its two most prevalent technologies: Reverse Osmosis and Electrodialysis [87]–[89]. Reverse Osmosis employs a semipermeable membrane to filter out salts, bacteria, and other impurities. Electrodialysis, on the other hand, utilizes an electric potential to drive salt ions through membranes, effectively separating them from the wastewater. This method excels in

applications where specific ion removal is critical, such as in certain industrial wastewater treatments.

In the field of WWT, many researchers have made contributions to process design and optimization [32], [84], [90]–[94]. Optimizing the cost of treatment as well as sustainability assessment has been given some attention in recent years [2], however, few studies have focused on the sustainability assessment of wastewater treatment networks (WWTNs). Additionally, there has been no research regarding the integration of ecological footprints during the synthesis of wastewater treatment networks. Therefore, this chapter uses a superstructure-based approach to formulate a multi-objective problem for cost and sustainability assessment through an MINLP formulation and optimization. The sustainability assessment metric considered includes the sustainable process index (SPI), which is an ecological footprint indicator. This work implements a graph-theoretic approach called the P-graph framework [34] for the synthesis of WWTNs.

The design of wastewater treatment plants (WWTP) is a complex systems problem, owing to the diverse range of technologies available for the removal of different contaminants. The treatment process is performed in stages, as shown in Figure 8, with the classification of treatment technologies based primarily on the relative abundance of the contaminants, treatment efficiency, driving force for separation, and the type of contaminant being removed.

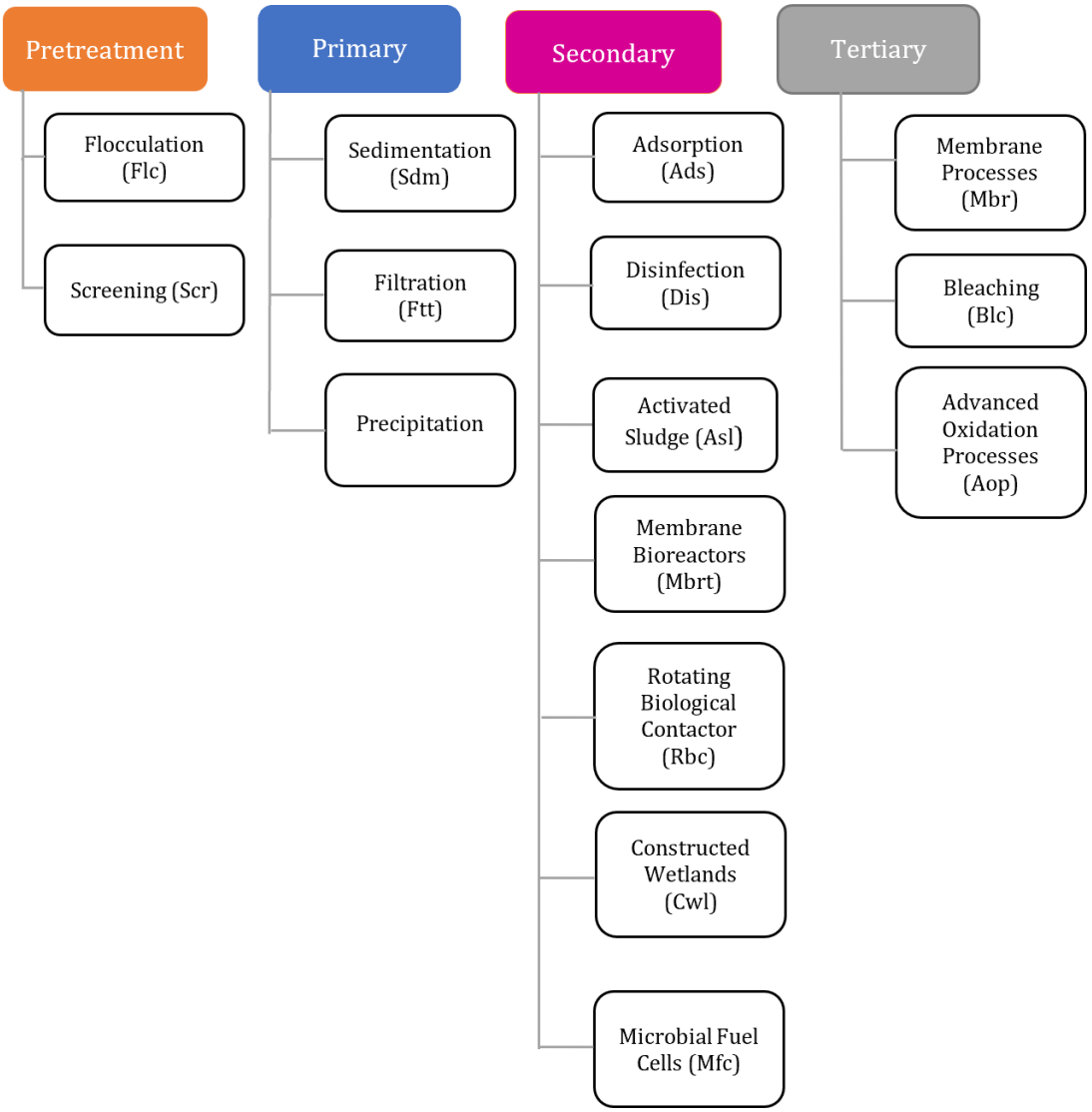


Figure 8. Typical technologies found in stage-wise WWT

3.1.1 Stages and Categories in Wastewater Treatment

Wastewater treatment (WWT) typically falls into three main categories: physical, biological, and chemical, as shown in Figure 9. Categories in this case describe the various processes and transformations that happen in a specific treatment technology during the treatment process. The physical category entails treatment technologies that typically use

separation criteria based on physical driving force. For example, sedimentation technology leverages the difference in particle densities for contaminant removal. The biological category are the technologies that use biological means and transformations to remove contaminants, such as activated sludge, rotating biological contactors, among others. Finally, the chemical categories are technologies that typically use some chemical transformations to remove contaminants. Examples of such technologies include advanced oxidation processes, disinfection, among others. It should be noted that most technologies use a combination of these categories for contaminant removal, however, there is always a predominant mode of contaminant removal for each treatment network.

These categories are used in a stage-wise approach to effectively remove contaminants based on their physical, biological, and chemical properties. During the pretreatment stage, a screening unit is used to remove large solids such as rags, cans, bottles, or anything that could clog pumps or pipes downstream. Additionally, flocculants are typically added to help suspended solid contaminants coalesce together and form heavier solids within the wastewater stream. The primary stage involves physical means, such as sedimentation and filtration, to remove solid contaminants. Most bacteria and microorganism treatment methods are found at the secondary stage, which is used to remove biological and chemical contaminants. The tertiary stage is usually used to adjust the pH and remove further contaminants that are still not within the desired standards. Typically, physical treatment processes are involved in the primary stage, while the secondary and tertiary stages include biological and chemical treatment processes.

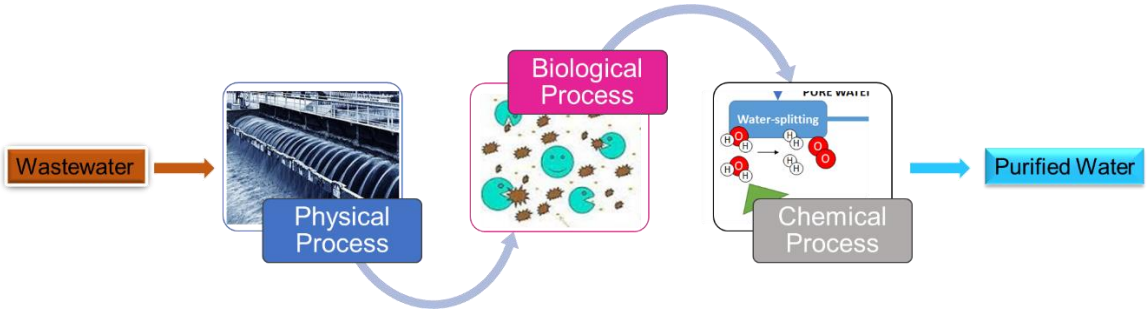


Figure 9. The three major types of categories for WW

3.1.2 Wastewater Treatment Technologies

Technologies selected at each stage of the treatment process are dependent on the contamination, efficiency of removal, and the driving force. These technologies can be grouped into the pretreatment, primary, secondary, and tertiary stages as shown earlier in Figure 8. The pretreatment stage typically facilitates solids removal during the primary stage. Additionally, screens are implemented during this stage to remove large solid particles which can clog the downstream technologies. Technologies with lower efficiencies are generally included in the primary stage, which is focused on the removal of solid contaminants (usually are present in large quantities). The secondary stage comprises technologies that can effectively remove both biological and chemical contaminants. The tertiary stage includes the most efficient technologies for biological and chemical contaminants, and solid contaminants that may have escaped the pretreatment, primary, and secondary stages. However, in some cases, the tertiary stage can be omitted if the purity requirements are not stringent. The illustrated diagram in Figure 8 portrays the sequential stages of WWT and some technologies employed in each stage. Abbreviated names consisting of three to four letters are used to represent the various technologies in

the diagram. These abbreviations will be adopted in the process flow diagrams and superstructure representations henceforth. Below is a discussion of some of the treatment technologies.

3.1.2.1 Pretreatment Technologies. Flocculation (Flc) is a process that involves the addition of a substance known as flocculants to a mixture, which leads to the clustering of particles to form larger agglomerates [52], [95]. The particles tend to become destabilized following the introduction of flocculants, which causes them to stick together and coagulate [52]. The suspended coagulated solids are further clustered together to form rapid-settling flocs. These flocs are then removed during the primary treatment stage [96].

3.1.2.2 Primary Treatment Technologies. Sedimentation (Sdm) is a unit operation used to separate solid components in a liquid mixture based on their density differences, with less dense component rising to the top and the heavier ones sinking to the bottom [28], [52], [97]. Separation is achieved by allowing the mixture to settle over a period of time, during which gravity causes natural separation. Sedimentation is primarily employed in the treatment process to remove large solid particles or flocs generated during flocculation.

Filtration (Granular) (Ftt): In WWT, suspended solids (SS) are commonly removed using granular filtration, which targets particles with sizes ranging from 10-100 microns. Prior to filtration, WW is often subjected to flocculation to promote the formation of larger particles. The filtration unit typically employs a bed of granular filter media, which functions in layers. Larger particles are removed at the top layer, while smaller particles are captured at the lower layers[52], [98], [99]. The efficiency of this process depends on factors such as the filter media, filtration velocity, particle concentration, and the

physicochemical properties of the SS. Different filter media types used for this process include gravel, coarse and fine sand [100].

3.1.2.3 Secondary Treatment Technologies. Adsorption (Ads) is a well-established process in which an adsorbate stream is made to flow through a solid adsorbent, leading to the accumulation and deposition of the adsorbate onto the surface of the adsorbent [101]. Several types of adsorbents can be used for this process, including activated carbon, synthetic polymers, and silica-based adsorbents. Activated carbon, in particular, is often preferred due to its cost-effectiveness and its ability to undergo thermal regeneration after use [52].

Disinfection (Dis) is a critical process employed to eliminate harmful bacteria and other contaminants present in WW. Three widely used disinfection techniques are chemical, physical, and radiation-based methods [102]. Chemical disinfection techniques include chlorination, ozonation, as well as acid and alkaline treatments [103]. Physical disinfection techniques rely on heating, ultraviolet (UV) irradiation, filtration, and settling. On the other hand, radiation techniques employ electromagnetics and acoustics [104]. Chlorine is the most commonly used chemical disinfectant. While light/heat disinfection is also effective, it can be expensive when large volumes of WW require treatment [105], [106].

Rotating Biological Contactor (Rbc) is a biological treatment process that involves the use of basins containing large circular disks, which are mounted on horizontal shafts that rotate slowly through WW streams. The disks are typically divided into various compartments by baffles within a single basin or reservoir [107], [108]. Biomass removal efficiency is dependent on the speed of rotation and disc diameter. Rbc is primarily used to

remove biological oxygen demand (BOD), phosphorus, nitrates, and some suspended solids during the WW treatment process [108].

Membrane Bioreactor (Mbrt) consists of a bioreactor with immersed membranes. The membranes are constructed from various materials such as plastic and ceramics. Compressed air is passed through the system to scour the membrane. This technology is effective at removing organic carbon and other nutrients [29], [109]. An increase in aeration rate can help prevent the formation of biofilms on the surfaces of the immersed membranes [110]. The membrane bioreactor operates in a similar manner to activated sludge, but without the need for a secondary clarification unit. It is considered a promising and environmentally friendly treatment technology due to the high quality of effluent produced [111].

Constructed wetlands (Cwl) utilize the concept of vegetation, soil, and microorganisms in the treatment of wastewater. The vegetation within the wetland facilitates absorb nutrients such as nitrates from the wastewater [112]. As wastewater passes through soil layers in constructed wetlands, there is some sedimentation and filtration [113]. The choice of plant species is a crucial factor in determining the efficacy of the constructed wetland in treating wastewater [114], [115]. Constructed wetlands are advantageous in that they can be used in small or remote areas and do not involve high operating costs [116].

Microbial fuel cell (Mfc) technology involves the degradation of substrates present in wastewater by microorganisms. In this process, microbes oxidize the organic substrates in the anode chamber, leading to the generation of electricity as a result of the transport of electrons to the cathode chamber of the cell. The hydrogen ions move through a semi-

permeable membrane to the cathode chamber, where they combine with supplied oxygen to form purified water [117], [118]. The utilization of wastewater as a fuel source provides the advantage of simultaneously purifying water and generating electricity in microbial fuel cells. However, the high capital cost associated with this technology represents the primary obstacle to its commercialization [119].

Activated sludge (Asl) is a process used mainly for the treatment of biological contaminants. The objective is to convert biodegradable organic substances into more stable compounds. The traditional method employed in the activated sludge process is the suspended growth process, which can be either anaerobic or aerobic [28], [52]. The activated sludge process commonly involves three fundamental components: a reactor for microorganism suspension, a liquid-solid separation chamber, and a recycle system [120]. Numerous industrial and utility firms have adopted the activated sludge technology for treating their wastewater streams [121]–[123].

3.1.2.4 Tertiary Treatment Technologies. Advanced oxidation processes (Aop) refer to the use of oxidizing agents to oxidize contaminants in wastewater. Hydroxyl radical-based and sulfate radical-based processes are some examples of such processes [124]. These radicals react with the organics in the wastewater, leading to their decomposition [125]. Examples of advanced oxidation processes include Fenton, photo-assisted Fenton, catalytic ozonation, photocatalysis, and the combination of hydrogen peroxide with ozone or with other agents [126]–[128].

Membrane processes (Mbr) refer to pressure-based filtration technologies that utilize semi-permeable membranes and particle sizes for separation. The prominent membrane technologies are microfiltration, ultrafiltration, nanofiltration, reverse osmosis,

dialysis, and electrodialysis [129]. Generally, membrane technologies with smaller pore sizes tend to be more expensive. However, these wastewater treatment systems can achieve an efficiency of approximately 99% [52], [130]–[132].

Bleaching (Blc) is a technology employed in the wastewater treatment (WWT) process to remove impurities and enhance the color of the treated wastewater. It is typically implemented at the final stage of treatment. The wastewater streams that commonly require bleaching are those generated by the pulp-and-paper industry[133], [134]. Chlorine in the form of hypochlorite and ozone are the usual bleaching agents utilized in this technology.

3.1.3 Wastewater Contaminants and their Classification

Wastewater streams require appropriate treatment technologies based on their unique characteristics, such as total suspended solids (TSS), total dissolved solids (TDS), pH, phosphates, nitrates, oil and grease contents, and heavy metals, among others [28], [135], [136]. Contaminants in WW streams can be categorized as biological or chemical oxygen-demanding components. Biological oxygen demand (BOD) is assigned to components like fatty acids, proteins, alcohols, and sugars and indicates the amount of oxygen required for microorganisms to eliminate contaminants [52], [97], [137]. Chemical oxygen demand (COD) is used to measure the presence of chemically treated contaminants, including those that require BOD [138]. The ratio between COD and BOD can determine the amount of oxygen required to oxidize contaminants to acceptable levels in the WW.

3.2 Superstructure Optimization Approach to Wastewater Treatment

The selection of a suitable technology for each treatment stage to meet purity requirements while minimizing the overall objective function is a complex decision-making process. This design complexity, therefore, requires the application of systems engineering tools to

determine the most appropriate technology for a given WW stream. One approach to addressing this problem is through the use of a superstructure generation and optimization method [35], [139], [140]. Many researchers have employed this superstructure optimization technique for process synthesis problems such as integrated process water networks [141], distributed water treatment and supply [142], and dynamic influent WWTNs [35]. Superstructures are very effective ways of selecting WWT technologies.

3.3 The P-graph Approach to Wastewater Treatment Network Synthesis

The P-graph framework is a systematic, graph-theoretic method for process design. Established by Friedler et al. in 1992 [34], [143], [144], it uses a distinct set of axioms and algorithms, alongside a bipartite graphical depiction of process units. This structure enables comprehensive handling of design problems. Guided by the P-graph axioms, a maximal structure (or superstructure), is generated from the initial set of materials and operating units. This process, inherently ruling out structural inconsistencies and incomplete structures, provides a clear advantage over traditional superstructure-based optimization methods by preventing errors and reducing unnecessary complexity. Moreover, algorithms of the P-graph framework capitalize on the structural properties of the network to manage binary terms and expedite optimization solutions. This is achieved by minimizing the problem size and enhancing the bounding step. The methodology also yields the set of n-best designs in a ranked order. In contrast, traditional methods often require further model modifications such as integer-cuts to obtain a ranked set of solutions, leading to additional time and computational effort.

The P-graph framework rests on 3 main foundations. The first foundation is characterized by two main node types: M-type and O-type [34]. The M-type nodes,

symbolized by circles on the P-graph, represent the materials or streams under consideration in the process design. On the other hand, the O-type nodes, depicted as horizontal bars, correspond to the operating units that manage the transformation of these materials. These distinct nodes are interconnected through arcs, which elucidate the directionality of the material flow within the structure.

The second foundational principle involves a collection of combinatorial axioms that define the structural prerequisites for a viable process [34]. A system of operating units that adheres to these axioms is termed a "combinatorially feasible structure" or "solution structure." As such, the pursuit of the optimal, or a range of the best feasible processes, is restricted to these combinatorially feasible structures, resulting in a significant narrowing of the search domain. The following are the axioms of the P-graph framework and the corresponding explanations.

Axiom 1: Every product must be represented as a material node in the structure of a feasible process [34].

This axiom underlines a fundamental principle in process engineering. In the context of process engineering, a "material node" refers to a point within the process where a specific material or product is identified and accounted for. This includes raw materials, intermediates, final products, and by-products. The axiom emphasizes that every desired product of a process must have a corresponding material node. This requirement is not merely a matter of record-keeping or systematic organization but is crucial for the operational feasibility of the process. If a desired product is not represented as a material node, it implies that the process lacks a clear pathway or mechanism for its production.

Axiom 2: A material node of the structure of a feasible process has no ancestor if and only if it represents a raw material [34].

This statement not only defines what constitutes a raw material within a process but also delineates the boundaries between the process and the external environment. The concept of an "ancestor" in this context relates to the origin of a material within the process structure. According to the axiom, a material node that has no ancestor is classified as a raw material. This essentially means that raw materials are inputs that enter the process from outside and are not products or intermediates of any process steps. They are the starting points, the inputs from which the process begins. The axiom underscores the role of raw materials as interfaces between the process and the outside world. This perspective is crucial for understanding how processes are designed and operated. Raw materials are inputs that are external to the process – they are not generated within the system but are sourced from the environment. In practical terms, the axiom provides a clear guideline for identifying and managing raw materials in a process. It implies that any material required for the process that cannot be produced internally must be sourced externally. This understanding aids in the efficient design and optimization of processes, ensuring that all necessary inputs are accounted for.

Axiom 3: Each operating unit appearing on the structure on the structure must be an element of the set of operating units of the problem definition [34].

This axiom essentially states that in designing a process, only those operating units explicitly identified at the outset of the problem should be used as components of the process. This axiom has far-reaching implications for how engineers approach the design and optimization of industrial processes. At its core, this axiom establishes a foundational

rule for process design: the set of operating units considered for a process must be predefined and fixed at the beginning of the design phase. Operating units refer to the various equipment and components used in a process, such as reactors, separators, heat exchangers, and others. The axiom demands that the design and synthesis of a process should only incorporate those units that have been identified and defined as part of the initial problem set. This approach ensures a structured and focused design process, as it limits the scope of the design to a predefined set of tools and capabilities. This axiom also serves to streamline the problem-solving process in process synthesis. By establishing clear boundaries on what operating units can be used, it helps prevent the complication of the design process with too many variables or choices. This focus is particularly important in complex industrial processes where the myriad of potential options could otherwise lead to analysis paralysis or suboptimal solutions due to an overabundance of choices.

Axiom 4: For each operating unit node of the structure of a feasible process, there is a path leading from this node to a material node of the structure representing a product [34].

This statement asserts that every operating unit within a process should contribute, either directly or indirectly, to the production of a desired product. In any industrial process, operating units such as reactors, separators, or distillation columns serve specific functions in converting raw materials into desired products. The axiom underlines that each of these units must be part of a pathway leading to the creation of a product. This means that every unit should have a clear and justified role in the overall process; it should be involved in processing, transforming, or otherwise contributing to the production of an end product. The rationale behind this axiom is to prevent redundancy and ensure efficiency in process design. By stipulating that each operating unit must be part of a productive pathway, the

axiom eliminates the inclusion of superfluous or non-contributory units. This approach is critical in optimizing resource utilization, reducing costs, and improving the overall efficiency of the process.

Axiom 5: For each material node of the structure of a feasible process, there exists at least one operating unit in the process, for which the material represented by this material node is an input or an output [34].

This axiom essentially states that every material identified in a process must be linked to at least one operating unit as either an input or an output. This axiom is pivotal in ensuring that every material in a process is accounted for and actively participates in the process. The rationale behind this axiom is to prevent the occurrence of redundant or idle materials within the process, which can lead to inefficiencies and inconsistencies. By stipulating that each material must be associated with at least one operating unit, the axiom ensures the coherence and logical flow of the process. This approach is essential for the efficient management of materials, ensuring that they are utilized effectively and contribute to the desired outputs of the process.

The third foundational element consists of algorithms developed based on the combinatorial axioms. These algorithms leverage the structural details of the problem, producing results that refine the synthesis process and address significant challenges in algorithmic process design. Specifically, the MSG (Maximal Structure Generation) algorithm autonomously creates a thorough and precise superstructure, known as the "maximal structure". This incorporates the operating units and materials as designated by the designer. Meanwhile, the SSG (Solution Structure Generation) algorithm lists all structures within the maximal structure that adhere to the P-graph axioms. These structures,

termed “combinatorially feasible structures”, can then be individually optimized, producing either a comprehensive or a top-ranked list of solutions.

Thus, these capabilities of P-graph can be leveraged for an efficient design of wastewater treatment network due to their systems nature. The superstructure developed for both pharmaceutical and municipal case studies presented in this work uses conventional approach for the synthesis analysis, while the tannery and coffee wastewater case studies implement the P-graph approach for the synthesis problems.

3.4 Sustainable Process Index for Wastewater Treatment Network Assessment

As discussed earlier, the commonest way to evaluate the environmental sustainability of processes is through LCA. However, there are other methodology that can be used which not, only consider the emissions from the process, but also take into account the source and type of raw material and energy usage, the area needed to accommodate the actual process plant, and further accommodate staff. Thus, in this work we incorporate the SPI methodology for sustainability assessment of wastewater treatment networks. Equation (1) gives the basic equation for estimating the SPI.

A_R considers the area needed to provide both renewable and non-renewable raw materials. Equation (2) gives the area needed for renewable raw material production, A_{RR}.

$$A_{RR} = \frac{F_{RR}}{y_{RR}} \tag{2}$$

where F_{RR} is the flow of the renewable material. Since wastewater is predominantly water, the average precipitation rate and seeping ratio are used to estimate the annual yield of rainfall, y_{RR}. From the United States’ National Oceanic and Atmospheric Administration

(NOAA), the average precipitation rate from 2009 to 2019 is 0.8105 m/y. Multiplying this value by the density of water (1000 kg/m^3) and a seeping ratio of 0.3 gives a value of 243.2 kg/m^{2}y as the yield of water. For the non-renewable raw material, A_{NR}, Equations (3) and (4) are used.

$$A_{NR} = \frac{F_{NR} E_{D,NR}}{y_{NR}} \tag{3}$$

$$E_{D,NR} = \frac{0.95\ C_N}{C_E} \tag{4}$$

Here, F_{NR} is the flowrate of the non-renewable raw material used in the wastewater treatment, $E_{D,NR}$ is the energy required to supply one kilogram of the material in question, C_N is the price of the material, and C_E is the unit price of energy. The total raw material area is estimated using Equation (5)

$$A_R = \sum_{i=1}^{n} A_{i,RR} + \sum_{j=1}^{m} A_{j,NR} \tag{5}$$

where i and j are the materials n and m are the total number of renewable and non-renewable raw materials, respectively.

A_E considers the area needed to provide energy for the treatment process. Equation (6) is used for the estimation.

$$A_E = \sum_{i=1}^{n} \frac{F_{i,E}}{y_{i,E}} \tag{6}$$

Here, the $F_{i,E}$ is the energy used from the i energy source (hydro, nuclear, solar, etc.), while the $y_{i,E}$ is the energy yield from the corresponding energy source.

For area needed for installation, A_I, two components are considered. The first is the area needed for direct installation of the of the various technologies. This area is estimated based on the size of the equipment. The second is the area needed for indirect installations such as piping and instrumentation. Similar to using the energy demand of the non-renewable raw material area, an estimate is made based on the cost of installation, plant life, and cost of electricity. Equations (7) – (9) are used for these estimates.

$$A_{II} = \frac{E_{D,II}}{y_{II}} \tag{7}$$

$$E_{D,II} = \frac{0.54\, C_I}{C_E Y} \tag{8}$$

$$A_I = \sum_{i=1}^{n} \left(A_{i,DI} + A_{i,II}\right) \tag{9}$$

where $A_{i,DI}$ is the area occupied by technology i, and $A_{i,II}$ is the corresponding indirect installation associated with the specific technology. Equation (10) is used to estimate the area needed to accommodate staff.

$$A_S = \frac{N_S}{y_S} \tag{10}$$

Here, N_S is the total number of workers or staff at the treatment plant per year, while y_S is the number of workers per m^2 in a year. All these areas are considered as "input" areas to the process as shown in Figure 5 in section 2.1.

The "output" area, in this case comprises the area needed to dissipate the process emissions into the air, water, and soil compartments. To estimate the dissipation areas, the rate of renewal of each compartment, c (R_c), the allowable concentration of each emission component, m, into the compartment, c ($C_{m,c}$), and the amount of component, m emitted $F_{m,c}$. For the rate of regeneration for the soil compartment, R_{soil}, the rate of soil renewal, which averages 0.00022 m/y in the United States is used assuming the soil is of type loamy, with a density of 1300 kg/m^3 for a 50% pore space. Multiplying the density by the rate of soil generation gives an R_{soil} value of 0.2926 kg/m^2y. For the rate of renewal of the water compartment, the yield of water in Equation (2) is used. The case is different for the air compartment. Equations (11) – (13) are used for the area estimation. Here, the $A_{P,c}$ is the area for each ecological compartment, while A_P is the total area for emission dissipation. Detailed models developed for each compartment can be found in Appendix A

$$S_{c,m} = R_c \, C_{c,m} \tag{11}$$

$$A_{P,c} = \sum_{m=1}^{n} \frac{F_{c,m}}{S_{c,m}} \tag{12}$$

$$A_P = \sum_{c=1}^{n} A_{P,c} \tag{13}$$

3.5 Cost Optimization and Post-Optimization Analysis for SPI

Two case studies are considered for this analysis. The first entails a municipal wastewater treatment while the second is a pharmaceutical treatment problem. Below is a detailed description of the case studies, starting with superstructure synthesis, mathematical models for cost optimization, and SPI assessment of the optimal treatment pathways.

3.5.1 Problem Specification and Solution Methodology

A generic superstructure was developed for both case studies comprising all possible connections from the inlet wastewater stream to the final treated water. Pretreatment is the first stage which is made up of a flocculation (Flc) unit. In the second stage two primary treatment technologies are considered, namely, Sedimentation (Sdm) and granular Filtration (Ftt). Adsorption (Ads), Activated Sludge (Asl), Rotating Biological Contactors (Rbc), Membrane Bioreactor (Mbrt), and Disinfection (Dis) are the secondary stage technologies. The tertiary stage comprises of Advanced Oxidation Processes (Aop), Bleaching (Blc), and Membrane Processes (Mbr). Allocation of a technology to a stage is based on the efficiency of technology and wastewater characteristics. Depending on the effluent specifications, some stages can be skipped, hence a Bypass (Byp) is provided at each stage to accommodate that scenario. The

developed superstructure is shown in Figure 10, where each stage is linked to the next by Mixers (Mxr) and Splitters (Splt).

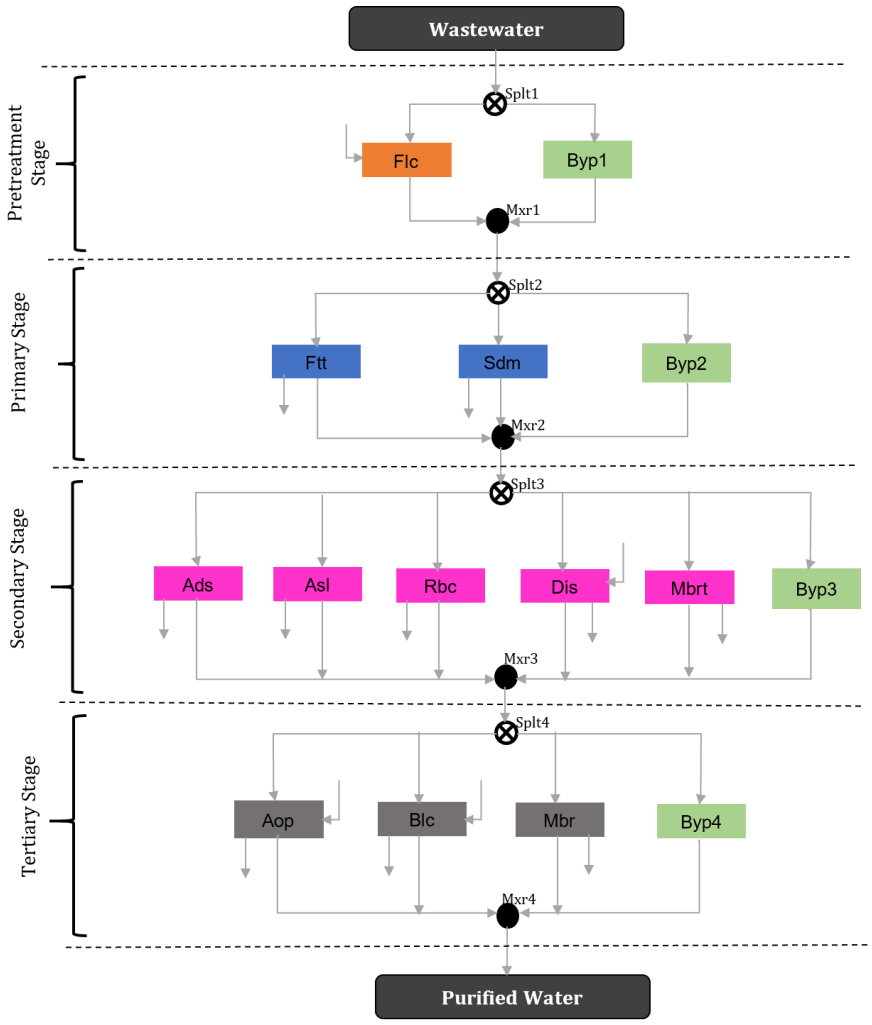


Figure 10. Generic superstructure for municipal and pharmaceutical wastewater treatment

The next step is to develop surrogate models for each technology considered in the superstructure. In building the mathematical models for each technology, mass and energy balances, design capacity constraints, capital cost, and operating cost are considered. The

mass balances include the constituents of the wastewater and other raw materials required for specific technologies. The capital cost is annualized over a 30-year period. The operating cost is made up of raw material, consumables, labor, utilities, and "other" (maintenance and overhead) costs. The raw material cost entails the cost of flocculant for the Flc unit, sodium hypochlorite for Blc, liquid chlorine for Dis, and ozone for Aop. Granulated Activated Carbon (GAC) for the Ads unit and filters for the Mbr unit are the major consumables for the analysis. In all technologies, it is assumed that the energy usage is by electricity. Detailed mathematical models for each technology in the superstructure, together with their parameters can be found in Appendix A. Additionally, all the analysis for SPI can be found in Appendix A.

Wastewater streams typically contain several waste constituents. For example, Table 1 shows typical contaminants found in a municipal wastewater.

Table 1

Typical Contaminants Found in Municipal Wastewater Streams ([28], [31], [52], [132])

Contaminant	Concentration	Units
Acids/Chlorides	5	mg/L
COD	68–272	mg/L
BOD	100–400	mg/L
Settable Solids	250–450	mg/L
Lead	30–80	mg/L
Zinc	1.0	mg/L
Nickel	0.04	mg/L
Copper	40–100	mg/L
Specialized Chemicals	<0.5	µg/L
Nitrogen	20-85	mg/L
Phosphorus	5-15	mg/L

Therefore, to simplify the optimization problem, the contaminants were grouped into three main categories, namely, solids, metals, and chemicals. The solids comprised of mainly settleable and suspended solids. The metals include heavy metals commonly found in wastewater streams such as lead, zinc, nickel. The chemical contaminants include chlorides, acids, organics, and inorganic compounds. One of the primary reasons for grouping BOD and COD together is to simplify the model. Both BOD and COD are

indicators of organic pollution in water, with COD encompassing a wider range of organic compounds than BOD. In the context of an MINLP model, which is already complex due to its nonlinear and integer characteristics, simplifying the model by aggregating similar variables can significantly enhance computational feasibility. This approach reduces the number of variables and constraints, thereby making the model more tractable and solvable within a reasonable timeframe. Furthermore, in many wastewater treatment scenarios, operations are designed to target a broad spectrum of contaminants. While BOD and COD have distinct chemical implications, their treatment often involves overlapping processes and technologies. By aggregating these parameters, the model more closely mirrors real-world treatment scenarios where specific treatments target a range of organic pollutants, rather than individual constituents. Additionally, the aggregation into a single category indicates a focus on the overall treatment efficiency rather than on specific pollutants. This approach is particularly relevant when the primary goal is to assess the general performance of a wastewater treatment network rather than its effectiveness in removing specific types of contaminants. Table 2 shows the inlet concentrations and effluent specifications for the municipal case study.

Table 2

Municipal Case Study Contaminant Composition and Effluent Specifications

Contaminants	Inlet Concentrations (g/m^3)	Outlet Specifications (mg/m^3)
Solids (settleable)	200	≤2
Metals (Pb, Cu, Zn, Ni)	0.1	≤0.005
Chemicals (acids, chlorides, organics, inorganics)	1	≤0.001

For pharmaceutical wastewater, they vary not only in composition but also in quantity and season depending on raw materials used in the manufacturing process [138],[145]. Additionally, the wastewater streams usually contain Active Pharmaceutical Ingredients (APIs) such as Acetaminophen, Dextromethorphan HBr, Guaifenesin, among others as shown in Table 3. Therefore, in the pharmaceutical case study, there is an extra contaminant component, namely, APIs, as shown in Table 4.

Table 3

Some APIs Found in Pharmaceutical Wastewater Streams ([29], [132])

Contaminants	Concentration (g/m^3)
Acetaminophen	32.5
Dextromethorphan HBr	1.0
Guaifenesin	20.0
Phenylephrine HCl	0.5

Table 4

Inlet Composition and Outlet Specification for Pharmaceutical Wastewater Case Study

Contaminants	Entering Stream (g/m^3)	Purity Specifications (mg/m^3)
Solids	10	≤ 2
Metals	0.01	≤ 0.005
Chemicals	44	≤ 5
Pharmaceutical (APIs)	0.4	≤ 0.02

The problem is formulated as a Mixed Integer Nonlinear Programming (MINLP), which is a subset of Discrete Programming problems. In formulating the MINLP problem, binary variables are used, where technologies selected are assigned a value of "1" (Yes),

while the unselected ones are assigned a value of "0" (No). MINLP optimization are typically used in situations where a "Yes" or "No" decision is to be made, hence its implementation in this work. The optimization problem is formulated in the General Algebraic Modeling Systems (GAMS) language, and the global optimization solver used is the Branch and Bound Navigator (BARON) [147], with the objective of minimizing the total cost of treatment as shown in Equation (13).

$$CCTPC = CCTAC + CCTRM + CCTCS + CCTLC + CCTUC + CCTOC \quad (13)$$

Here $CCTPC$, is the total treatment cost, while $CCTAC$, $CCTRM$, $CCTCS$, $CCTLC$, $CCTLC$, $CCTUC$, and $CCTOC$ are the annualized capital cost, raw material, consumable, labor, utility, and "other" costs, respectively.

Once the optimization is completed, the final step is to perform sustainability assessment of the optimal pathway using the SPI methodology as described in section 3.4. Below is the systematic evaluation framework for the treatment problem:

1. Determine technology, parameters, inlet stream composition, and outlet effluent stream specifications
2. Generate a superstructure that identifies and considers all treatment technologies
3. Develop mathematical models for each technology
4. Formulate an MINLP problem to minimize the treatment cost, while satisfying effluent specifications

5. Perform sustainability assessment for the optimal pathway using SPI. Additionally perform a comparative analysis of the incurred ecological burden with and without treatment

3.5.2 Results and Discussion

The selected optimal pathway for the municipal wastewater treatment is show in Figure 11, where Flc, Sdm, Ads, and Blc are the technologies selected.

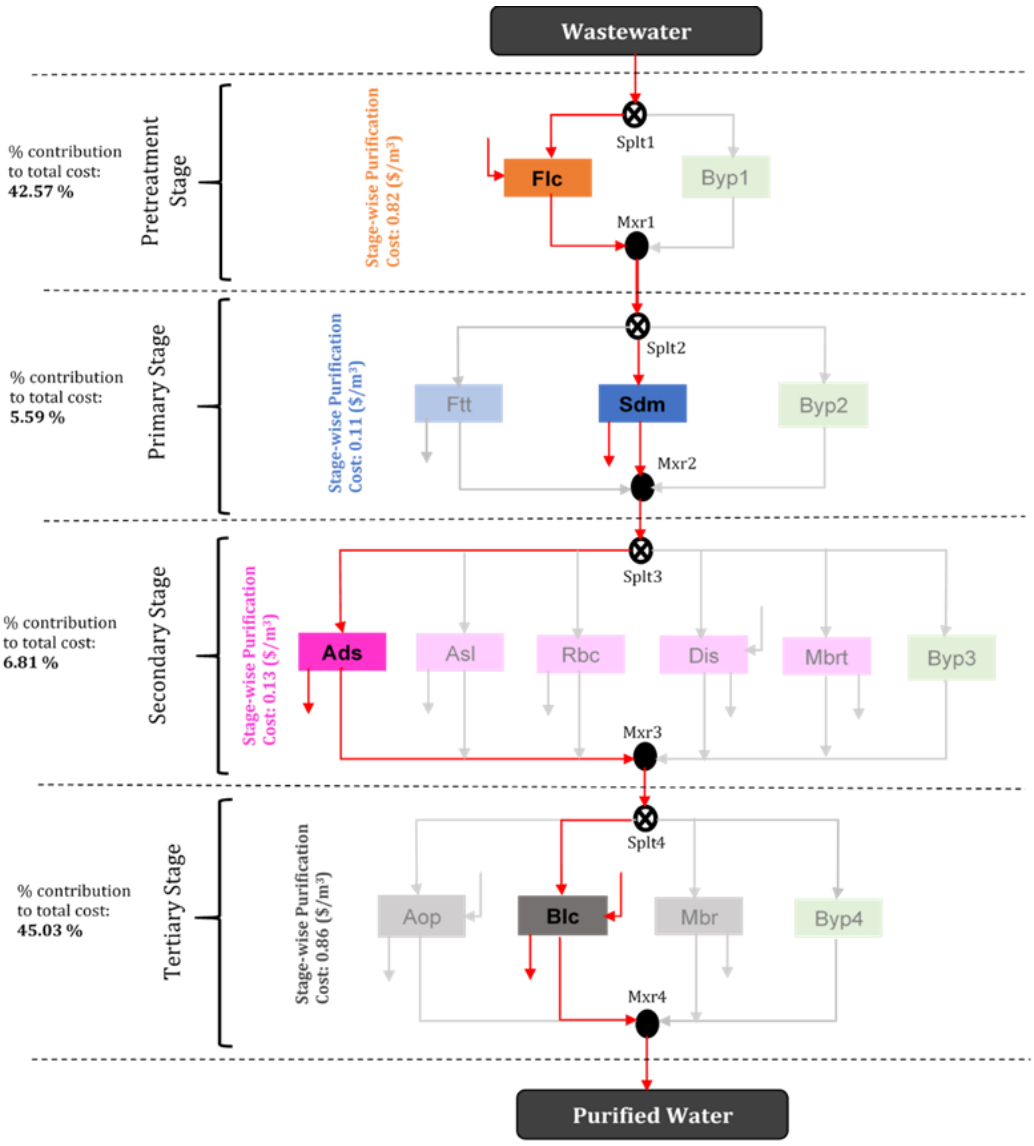


Figure 11. Optimal pathway for municipal wastewater treatment.

The total treatment cost is $1.52 million/y (1.92 $/m^3) with both pretreatment and tertiary stages dominating the cost as shown in Figure 11 by the stage-wise percentage cost distribution. This is due the high unit cost of the flocculant and hypochlorite used in the flocculation and bleaching units, respectively. Annually, the flocculation unit requires 796,356 kg of flocculant, and the bleaching process necessitates 1,307,164 kg of sodium

hypochlorite. The adsorption unit consumes 275,436 kg of GAC per year. The high consumption is attributed to the assumption of no regeneration for the GAC. Thus, regenerating the GAC could lower the yearly requirement and associated costs, however, most of the cost is dominated by the pretreatment and tertiary stages, hence cost reduction prioritization methods should be given to the technologies involved at these stages. Due to the high concentration of solids in the municipal wastewater, pretreatment stage was implemented to significantly improve the removal of the solid contaminants during the primary stage operation. The yearly expense for treating the wastewater, based on 7,920 hours of annual operation, amounts to 1.92 $/m^3. With this amount, the comprehensive operating cost - which encompasses material, consumable, labor, utility, and other associated costs - for the treated wastewater stands at 1.65 $/m^3.

Figure 12 shows the optimal pathway for the pharmaceutical wastewater treatment problem with Flc, Ftt, and Ads the selected technologies. In this case, the tertiary stage was bypassed with the primary stage being the highest cost contributor. The total treatment cost for the same yearly hours is 3.44 $/m^3, with an operating cost contribution of 3.16 $/m^3.

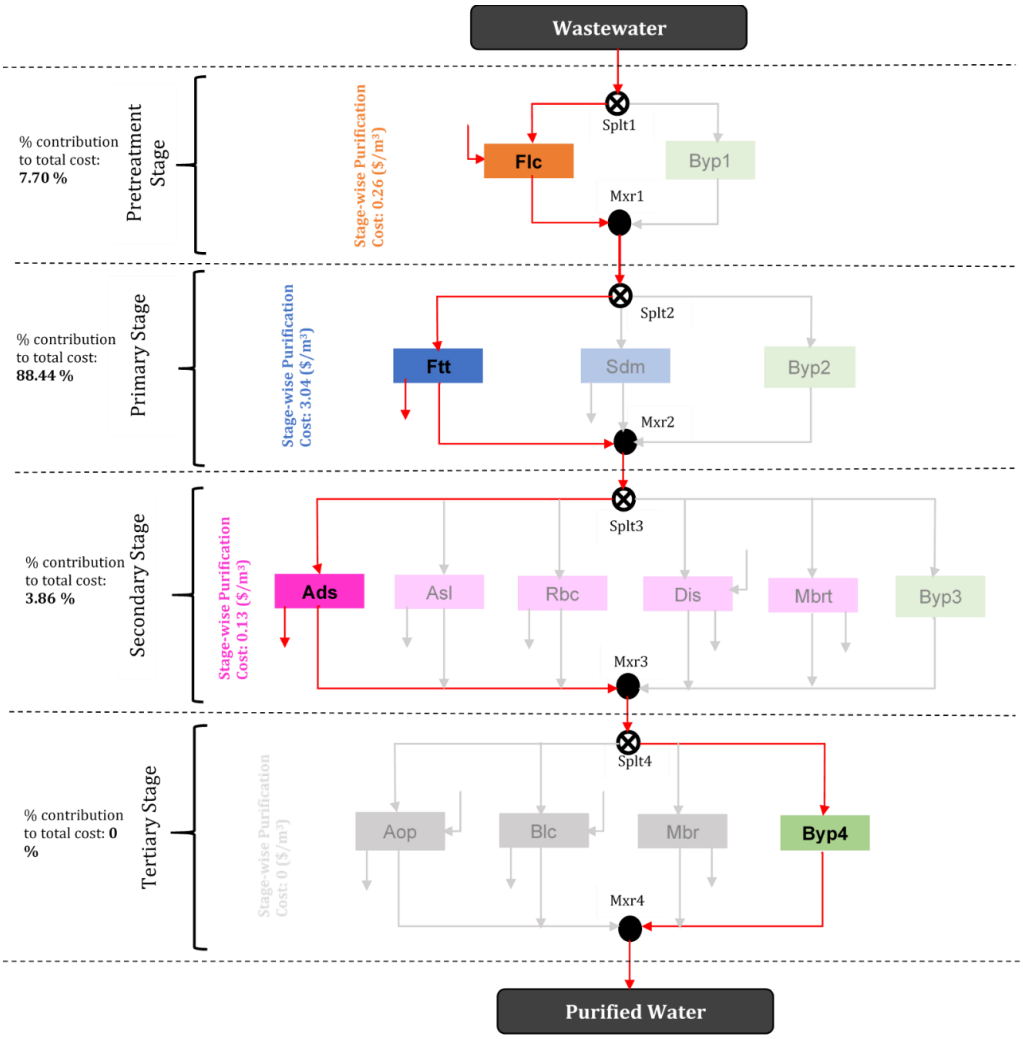


Figure 12. Optimal treatment pathway for pharmaceutical wastewater treatment

Often, stakeholders are interested in the next best choice aside from the optimal cost. Therefore, by applying the method of integer-cuts, the next sub-optimal pathway can be determined. Table 5 gives a summary of the ranked options for both case studies. From Table 5 it can be noted that for the primary stage, filtration was preferred to sedimentation due to the need for technologies with higher efficiencies required to meet the effluent specifications for the pharmaceutical waste stream. Additionally, due to the complex nature

of pharmaceutical wastewater streams, the cost of treatment becomes very expensive even for the optimal value. Thus, finding ways to reduce the amount of API in the waste streams can help improve the cost of wastewater treatment in these industries.

Table 5

Results Summary for Second and Third Best Options

Treatment Option	Treatment Network Pathway		Cost ($/m^3 WW)	
	Municipal	Pharmaceutical	Municipal	Pharmaceutical
First best (optimal)	Flc-Sdm-Ads-Blc	Flc-Ftt-Ads	1.92	3.44
Second best	Flc-Sdm-Dis-Blc	Flc-Ftt-Dis	5.89	7.45
Third best	Flc-Ftt-Dis-Blc	Flc-Ftt-Blc	8.56	20.80

The area needed to embed soil emissions has the highest percentage contribution of 77.7% to the municipal wastewater treatment. This is due to the low allowable concentrations. Given the presence of various metals in the wastewater stream, lead (Pb) was selected as the representative metal for quantification in this category. This decision was informed by the fact that Pb possesses the most restrictive annual allowable concentration in the soil compartment. Consequently, by addressing the area concerning lead, we inherently account for the other metallic contaminants. The raw material area is the next highest contributor with a value of 14.3%. Nevertheless, this percentage is heavily

influenced by the volume of wastewater processed annually. Notably, the spatial requirement for installation has the minimal impact on the overall SPI value. When contrasting the total spatial imprint for wastewater treatment with that of direct disposal (without any prior treatment), there is a marked 99.8% escalation in the SPI. Consequently, from an environmental perspective, it is more favorable to undergo wastewater treatment prior to its disposal. In our calculations for the area required for direct disposal, we exclusively considered the space needed to incorporate the wastewater into the soil and water compartments. This assumption is deemed justifiable, as other SPI categories only become significant in the presence of a treatment process.

For pharmaceutical treatment, the release of chemical contaminants into the water compartment of the ecosystem is the predominant factor influencing the SPI value. This prominence arises from the limited annual permissible concentration combined with an elevated flow rate of chemical contaminants to the water compartment. In contrast, the spatial requirement for installation has the least impact on the composite SPI value, suggesting that the treatment pathway exerts a minimal environmental burden in the overarching sustainability assessment. When considering strategies for pharmaceutical wastewater treatment, it's clear that reducing contaminant concentrations at their origin is pivotal. This perspective is reinforced by the fact that a substantial 93.9% of the SPI value stems from how contaminants disperse throughout various ecosystem compartments. In the absence of treatment, the area needed to manage wastewater in an environmentally conscious manner increases by 91.7%. In essence, bypassing treatment and introducing pharmaceutical wastewater directly into the ecosystem incurs a 91.7% greater environmental toll.

For a clearer picture, Figure 13 presents the SPI values from our case studies. It is noteworthy that the SPI value, even after pharmaceutical wastewater treatment, remains notably elevated compared to the direct disposal of standard municipal wastewater.

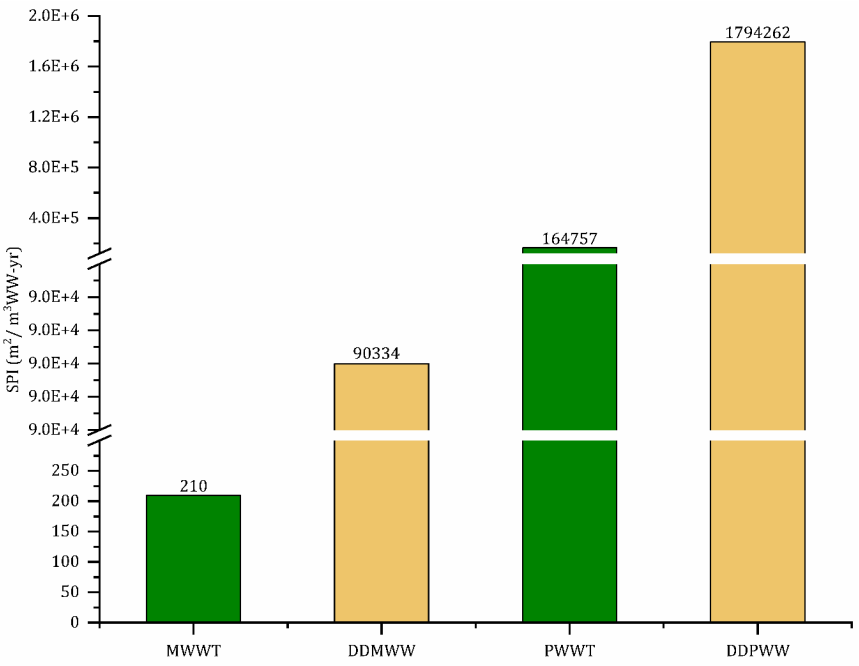


Figure 13. Consequential SPI for the various scenarios (MWWT – municipal wastewater treatment; DDMWW – direct disposal of municipal wastewater; PWWT – pharmaceutical wastewater treatment; DDPWW – direct disposal of pharmaceutical wastewater)

3.6 Structural Complexity Using the P-graph Methodology for SPI Analysis

In this section, the P-graph framework is used to first generate the maximal structure (superstructure). The solution structure generator (SSG) algorithm is then used to enumerate all the structurally feasible networks. Thus, the use of P-graph guarantees analyzing only structures that are feasible, thus, narrowing the search space for the optimal structure. Additionally, the use of P-graph can help generate the n-best feasible structures

without manually implementing integer-cuts for the next feasible options. The operational principles and the typical efficiencies of contaminant removal as discussed previously were used for the synthesis problem.

3.6.1 Problem Specification and Solution Methodology

In synthesizing the superstructure, the primary, secondary, and tertiary stages were considered in this analysis. For the structure of the primary stage, filtration has greater efficiency compared to sedimentation. Thus, a comprehensive structure may incorporate sedimentation followed by filtration. Therefore, the first stage can be structured in two main ways: a sedimentation process followed by filtration, or solely a filtration unit. The notion of placing filtration before sedimentation is not logical given the higher efficiency of filtration. Furthermore, while filtration typically succeeds sedimentation, it does not necessitate a mandatory flow through the filtration process. The flow retains the flexibility to bypass the filtration subsequent to sedimentation and proceed directly to the secondary stage for additional treatment. Figure 14 show the operations and connections for the primary stage.

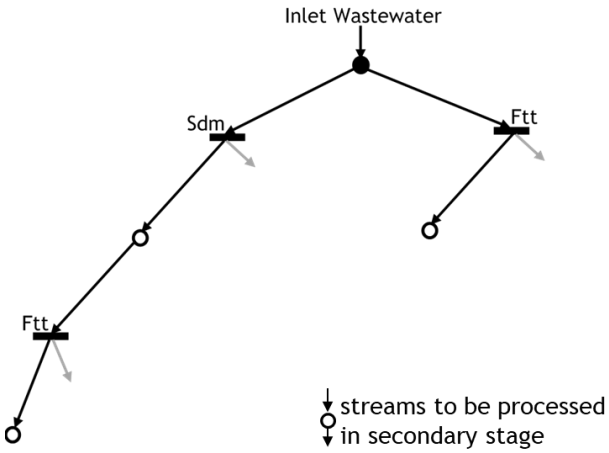


Figure 14. Primary stage structural complexity

For the secondary stage, one new technology was added to the technologies considered in section 3.5, namely, microbial fuel cell (Mfc). The same idea and principle used to generate the structure for the primary stage was used to generate a similar structure for the secondary stage as shown in Figure 15.

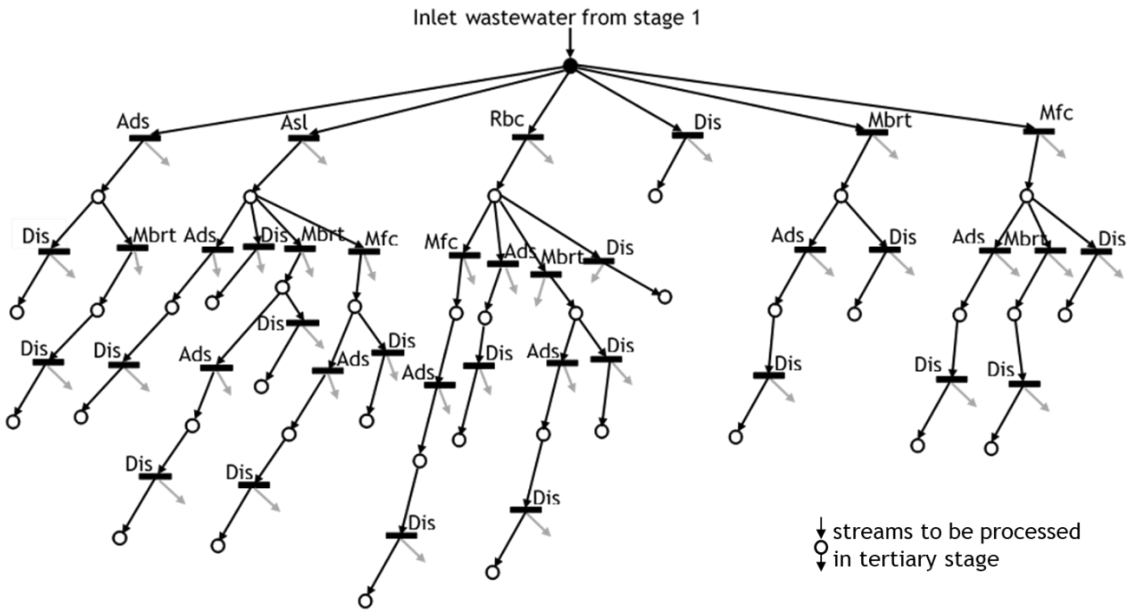


Figure 15. Secondary stage structural complexity

For the tertiary stage, two more technologies were included, namely, ion exchange (Inx) and ultrasonic (Uls). The same formulation used in stage one was used as shown in Figure 16.

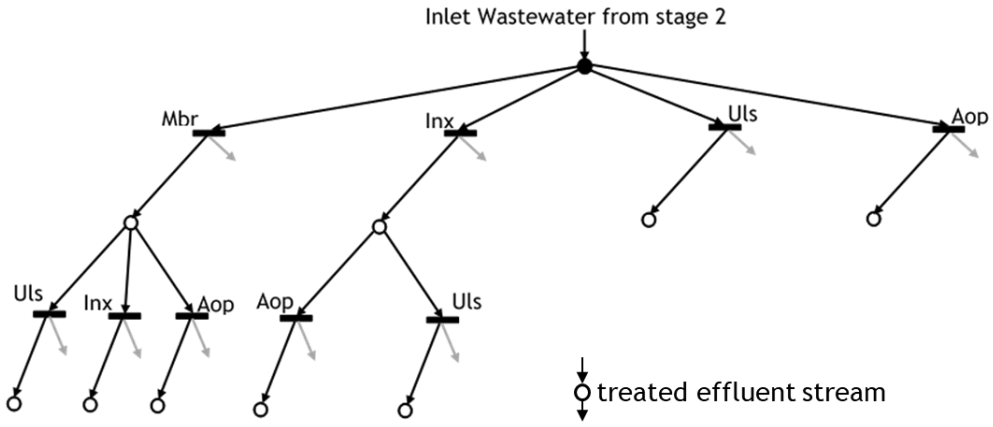


Figure 16. Tertiary stage structural complexity

It should be noted that each “streams to be processed” from preceding stage, has access to the whole structure of the next stage.

Upon implementing the P-graph approach in this study, the analysis identified 3, 38, and 9 plausible outlet streams for stages 1, 2, and 3, respectively. This led to the utilization of 1,720 operating units to model the feasible treatment pathways, along with an equivalent number of auxiliary units (1,720) to represent the mixing operations. Additionally, 3,447 M-type nodes were deployed to delineate the materials within the structure. The formulated optimization problem comprises 3,447 material balance inequalities and 3,440 continuous variables, which are used to represent the sizes of the various units. To indicate the inclusion or exclusion of units within the structure, an equivalent number of binary variables are incorporated. Despite the exponential increase in problem complexity attributed to these binary variables, they are effectively managed through the combinatorial algorithms inherent to the P-graph framework. For the purposes of this work, it is posited that the costs associated with the operating units are fixed-charge

linear functions, contingent on the flow rates of the streams entering the operations. Furthermore, the mixing operation is only considered in the final stage, immediately preceding the final product. This assumption is integral in maintaining the linearity of the overall model.

A tannery case study is considered for this analysis. The tannery sector stands as a predominant contributor to wastewater generation. In 2016, the United States (US) processed an estimated 30,000,000 tons of rawhide [148]. Given that each ton of rawhide processing yields 30 – 35 m^3 of wastewater [149], the annual tannery wastewater production in the US can be approximated to be between 900,000,000 and 1,050,000,000 m^3. The intricate nature of tannery wastewater can be traced back to the array of chemicals employed during the transformation of the rawhide into leather, introducing a spectrum of pollutants. Specifically, the tanning phase is the primary source of chromium, ammonium, and chloride salts in the wastewater, whereas the beamhouse is chiefly responsible for elevated organic and sulfide content. The untreated discharge of tannery WW, with its elevated concentrations of organics, sulfides, suspended solids, and notably chromium, poses significant health and environmental risks [150].

To simplify the analysis, contaminants were organized into five principal categories. Total Suspended Solids (TSS) encompassed all suspended solids in the sample. The metal contaminants category predominantly included heavy metals, specifically Mn^{2+}, Zn^{2+}, As^{3+}, Pb^{2+}, Cd^{2+}, Ni^{2+}, Co^{2+}, and Fe^{2+}. Although Chromium (VI) (Cr^{6+}) is generally classified as a heavy metal, it was isolated from this category due to its significant presence as a challenging pollutant in the tannery industry, stemming from the recalcitrant nature of its removal. The chemical contaminants category primarily consisted of anions, including

SO_4^{2-}, PO^{4-}, NO^{3-}, and Cl^-. For each contaminant category, the pollutant with the most stringent permissible limit was employed as the outlet specification. Table 6 presents the assumed inlet and outlet contaminant concentrations. The entering flow rate of the wastewater (WW) was assumed to be 100 m^3/h.

Table 6

Inlet and Outlet Specification for Case Study

Contaminant	Inlet Concentration (mg/L)	Outlet Specification (mg/L)
TSS	258	≤100
Metal	250.20	≤0.1
BOD/COD	5,958.62	≤30
Chromium	23.07	≤0.01
Chemical	3,459.32	≤30

3.6.2 Results and Discussion

Figure 17 showcases the first three potential structures. The best-fit structure incorporates technologies like sedimentation (Sdm), rotating biological contactor (Rbc), microbial fuel cells (Mfc), ion exchange (Inx), and advanced oxidation processes (Aop). Stage #1, #2, and #3 refer to primary, secondary, and tertiary treatments, in that order. The notable variation among the structures is in how the flow is divided for the process. For instance, structure #1 diverts 22,308 kg/h of wastewater, which is 24.77% of the total liquid stream coming from the ion exchange tech. It is evident from Figure that the wastewater

skipping the initial stage in structures #2 and #3 is quite minimal, at around 0.08% and 0.076%. This likely happens because the contaminant concentration in the waste stream remains high at the primary treatment stage. As we move on to the output from secondary and tertiary stages, they have had a significant amount of contaminants removed, allowing for more considerable flow splitting.

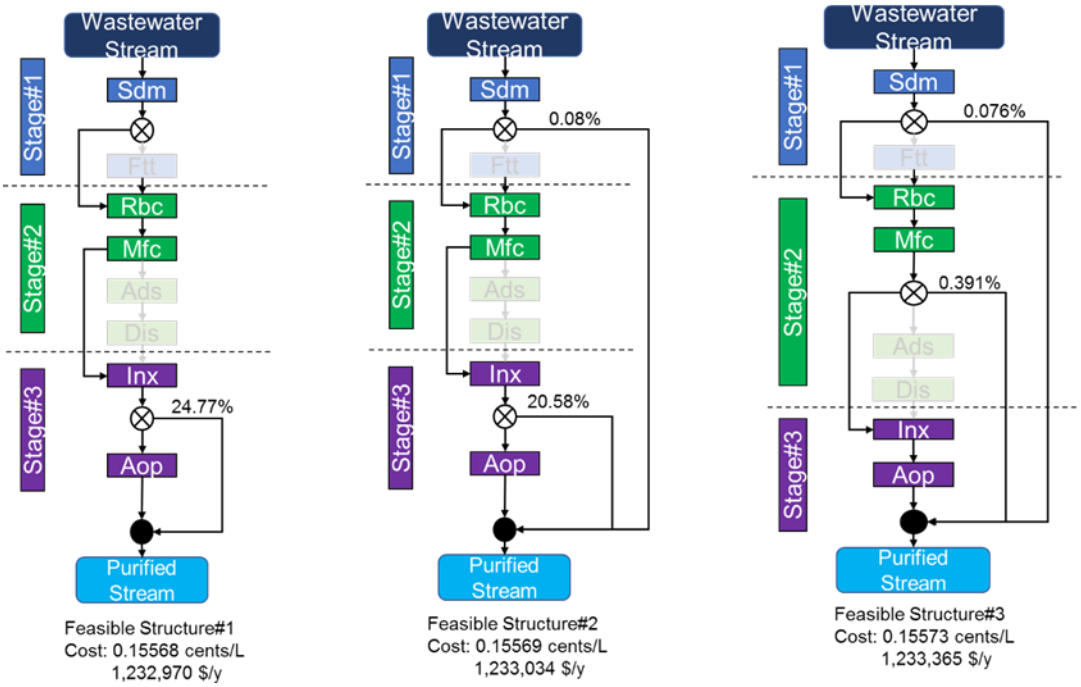


Figure 17. First three feasible structures showing selected technologies, the cost of treatment for 1 L of wastewater treated, and annual cost

From a stagewise cost distribution, the tertiary treatment stage stands out as the most significant contributor, ranging between 65.72% and 65.73%. The secondary stage contributes the least, at 3.93% for all the structures in Figure 1, while the primary stage floats around 30.34 – 30.35%. Diving deeper into the tertiary stage expenses, it becomes

clear that the Aop unit is the primary cost contributor, accounting for 93.47% of the entire tertiary treatment stage. Worth noting is that in every scenario, the chromium outlet specification was met. A closer look at the technologies revealed that the ion exchange unit effectively removed more than 90% of the chromium present in the inlet.

In the process of estimating the SPI for the treatment procedure, it is assumed that the contaminants removed are disposed via land and are thus characterized as emissions to soil. Concurrently, the purified wastewater stream, following treatment, is anticipated to be discharged into natural water bodies, and is hence classified as emission to water. The parameters, as outlined in Section 3.5, concerning the rates of renewability for both the soil and water compartments of the ecosphere, are utilized to compute the area necessary to effectively dissipate these water and soil emissions. Only the areas needed to dissipate soil and water emissions were considered for this analysis since the other areas made insignificant contributions to the overall SPI.

Table 7

SPI for the First Four Feasible Networks

Rank	Feasible Structure	SPI (km^2/m^3)	Cost ($/y)
#1	Sdm-Rbc-Mfc-Inx(24%Byp)-Aop	20.395	1,232,970
#2	Sdm(0.08%Byp)-Rbc-Mfc-Inx(20.58%Byp)-Aop	20.375	1,233,034
#3	Sdm(0.076%Byp)-Rbc-Mfc(0.391%Byp)-Inx-Aop	20.382	1,233,365
#4	Sdm-Rbc-Mfc-Inx-Aop	20.382	1,233,366

Table 7 presents the SPI values for each cubic meter of wastewater treated, corresponding to the feasible structures depicted in Figure 17. While feasible structure #1 is the most cost-effective, it simultaneously exhibits the highest SPI value compared to the other three structures. Consequently, when evaluating purely on the basis of cost, structure #1 emerges as the preferred choice. However, when considering ecological impact, structure #2 is more favorable. The variations observed in SPI values can be attributed to the distinct inequality constraints applied to the outlet stream. It's noteworthy that while all the structures met the stipulated purity standards, certain structures demonstrated superior contaminant removal efficiencies.

3.7 P-graph for Multiple Output Wastewater Treatment Streams and SPI Analysis

To help improve the sustainability of industrial processes, it is paramount to find ways to reduce water usage. The problem arises when trying to treat wastewater with multiple outlet streams, with each output stream meeting certain specification for reusability or disposal. In that regard, the synthesis problem becomes extremely challenging. Thus, having a wastewater treatment system that can produce multiple outlet streams can be advantageous, especially in the event of implementing recycling streams. Furthermore, there is a reduction in both capital and utility expenses as a singular treatment system is adept at facilitating the purification and recovery of multiple streams. In that regard, the determination of the set of n-best solutions can be a requirement from the decision maker.

The resulting problem from modeling different technologies oftentimes is a set of non-linear models to be solved. This is due to the difficult nature of guaranteeing a global

optimum after the optimization due to the nonlinearities within the system. However, by initially generating all the structurally feasible networks from the maximal superstructure, there is a guarantee that a global optima solution exists once each feasible network is numerically optimized. By using P-graph, all the structurally feasible networks can be generated using the SSG algorithm, and for each feasible network a detailed mathematical model can be developed to solve the resulting NLP problem. One advantage, in addition to being able to guarantee a global solution, is the reduction of structures that need to be solved, thus, improving computational efficiency. Furthermore, using P-graph helps in handling the technology selection part of the MINLP problem, making it easier to solve the resulting NLP problems. Additionally, being able to identify and rank n-best or all feasible networks based on either cost, or a sustainability metric presents the stakeholders a better wholistic view for decision making.

3.7.1 Problem Specification and Solution Methodology

A coffee case study is developed for this analysis. The soluble coffee industry is one of the highest consumers of water, consequently generating high volumes of wastewater [151]–[153]. For process intensification purposes and water conservation issues, it is advantageous to recycle water. Additionally, certain unit operations such as cooling towers [154] can benefit significantly from recycling due to less stringent water requirements for their operations. Table 8 shows the characteristics of the wastewater stream considered for this analysis, with a flowrate of 1,324,894 L/day.

Table 8

Wastewater Characteristics from Soluble Coffee Processing

Contaminant	Units	Feed Conditions
COD	mg/L	1140
TSS (Turbidity)	NTU	22
Conductivity	µS/cm	940

The objective is to have two outlet streams with one stream being used in a cooling tower operation while the other stream is discharged into a water body. The recycled stream is to have 80% of the contaminants removed, while the discharge stream specification is based on the discharge guidelines set by the United States Environmental Protection Agency (USEPA). Table 9 gives the inlet and outlet specifications for this case study.

Table 9

Coffee Wastewater Case Study Specifications. TBR – To Be Recycled; TBD – To Be Discharged

Component	Inlet Wastewater Flowrate (L/min)	Outlet Stream Specification (L/min)	
		TBR	TBD
COD	1.05	≤ 0.21	≤ 0.87
TSS	0.01	≤ 0.01	≤ 0.24
Conductivity	0.48	≤ 0.09	≤ 0.47

Three stages are considered in the synthesis of the maximal structure. The primary stage consisted of sedimentation and membrane processes for the removal of the TSS. Rotating biological containers, membrane bioreactors, and activated sludge were considered for the secondary stage to primarily remove COD, while advanced oxidation processes and ion exchange were considered for the tertiary stage to reduce the conductivity of the effluent streams to the required specifications. Figure 18 shows the maximal structure generated for the coffee case study.

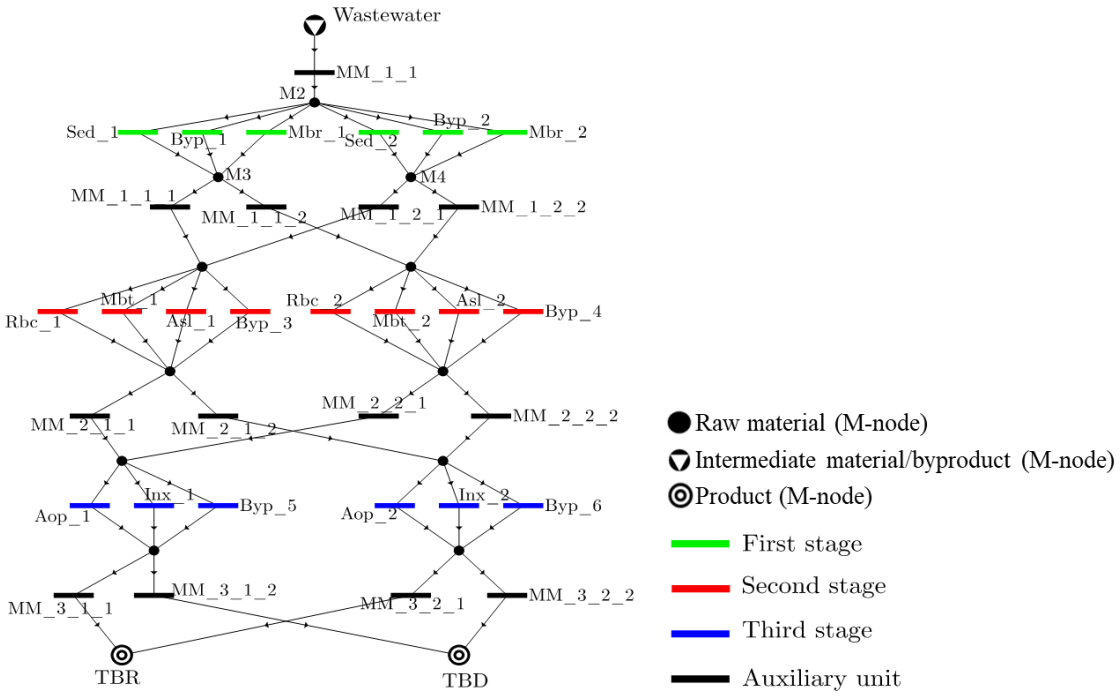


Figure 18. Maximal structure for coffee case study. TBR – To Be Recycles; TBD – To Be Discharged.

3.7.2 Results and Discussion

Using the MSG algorithm, over 300,000 possible network structures were generated. With the implementation of the SSG algorithm, 151,848 combinatorially feasible structures were identified from the maximal structure, thus reducing the structural search space by about 50%. Based on optimization of each feasible structure 2,779 were numerically feasible. The remining structures are either “technically infeasible” or the “solver failed to converge to a solution”. For the solutions which are “technically infeasible”, the structure fails to satisfy the constraints of the problem. One of the major constraints set for the analysis is to select structures with an SPI footprint lower than 1000 m^2. Thus, the problem can further be constrained for structures with even lower SPI values.

For the structures that the “solver failed to converge to a solution”, the solver provided no results.

If this synthesis problem was treated as a standard MINLP problem, the solver may halt for problems designated as “solver failed to converge to a solution” and the solution process may have been terminated. However, as can be seen from the analysis, these structures have no bearing on the results, once the problem was eventually transformed into NLP - another major advantage of using this approach for the treatment network synthesis. Figure 19 shows the structure (sedimentation-activated sludge-ion exchange) with the least SPI of 31.5 m^2. One non-intuitive observation is to bypass 25% bypass of the wastewater stream at the primary stage. Most of the SPI contribution is allocated to both the primary and secondary stages, with the tertiary stage contributing to only about 2% of the area.

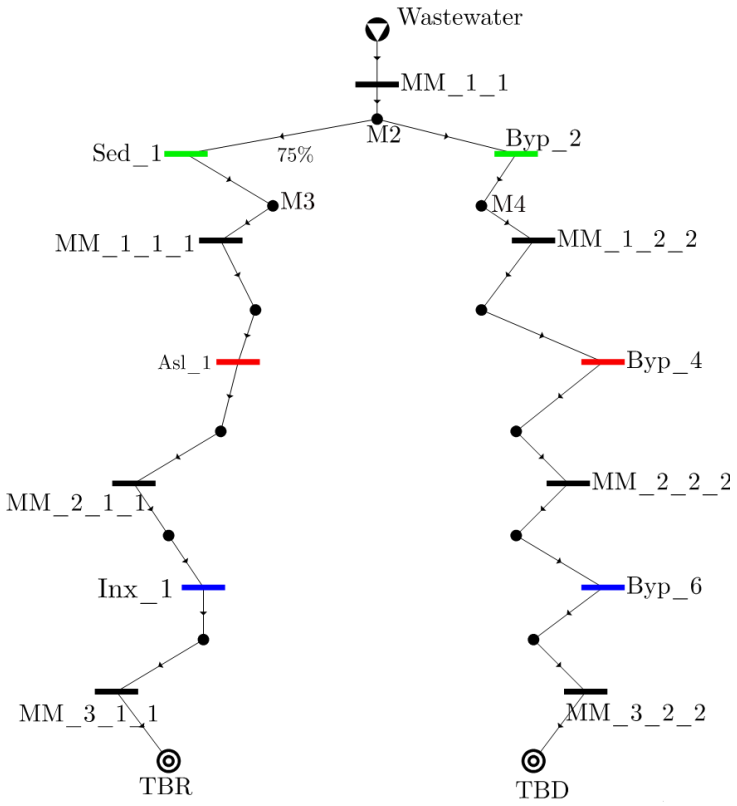


Figure 19. Structure with the least SPI

Figure 20 shows the technology preference for each stage based on both cost and SPI. Due to filter area requirements, selection of Mbr tend to increase the SPI of the overall treatment process, as shown in Figure 20 (a). From Figure 20 (b) it can be noted that there is equal preference between Rbc and Asl. One interesting oberservation also from the secondary stage is that none of the numerically feasible structures included Mbrt. This is due to the high SPI associated with its operation, hence, none of the structures with Mbrt could meet the 1000 m^2 constraint limit set. For the tertiary stage, there is much preference for Inx.

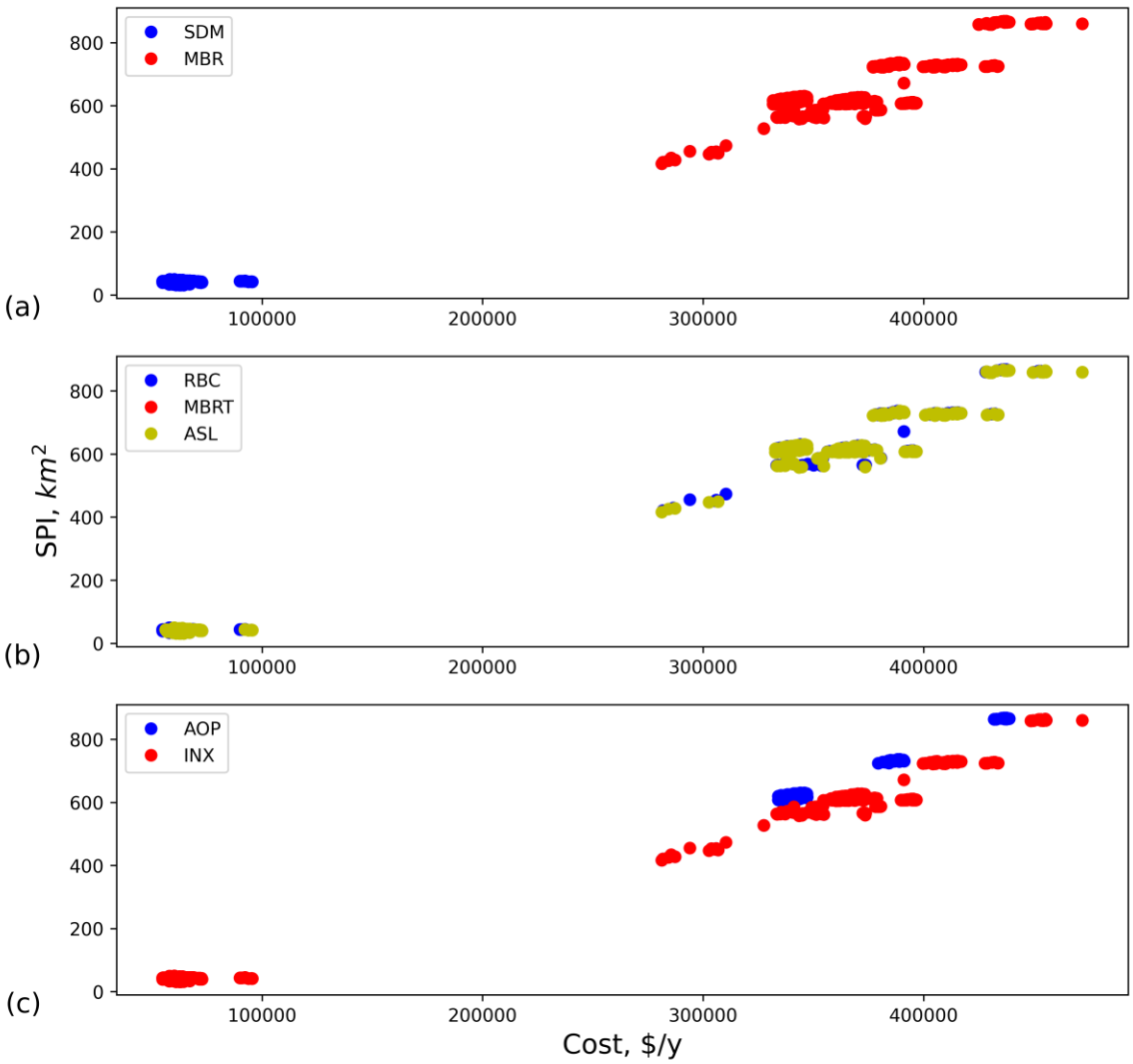


Figure 20. Distribution of technology selection for each stage. Subplots (a), (b), and (c) are for primary, secondary, and tertiary stages, respectively

3.8 Conclusions

Using the P-graph framework for wastewater treatment presents a better understanding of the synthesis problem. Non-intuitive solutions can be easily identified with this approach. Technologies that require the use of raw materials tend to increase the

SPI of the process. Additionally, mass-intensive technologies tend to increase the SPI of the process compared to energy-intensive technologies for wastewater treatment systems.

Chapter 4

Multi-Objective Approach to Solvent Recovery Systems

Text and figures are reproduced and adapted with permission from Aboagye, E. A.; Chea, J.D.; Yenkie, K.M.; Systems Level Roadmap for Solvent Recovery and Reuse in Industries – A Review. *iScience* 2021 https://doi.org/10.1016/j.isci.2021.103114 and Aboagye, E.A., Chea, J.D., Lehr, A.L., Stengel, J.P., Heider, K.L., Savelski, M.J., Slater, C.S., Yenkie, K.M., 2022. Systematic Design of Solvent Recovery Pathways: Integrating Economics and Environmental Metrics. *ACS Sustainable Chem. Eng*. 10, 10879–10887. https://doi.org/10.1021/acssuschemeng.2c02497

4.1 Background

The execution of solvent recovery processes brings with it several complexities. Within the context of our contemporary market-driven economy, cost is a critical factor influencing industrial policies, determining the feasibility of integrating solvent recovery into industrial processes. Furthermore, the selection of appropriate technology for solvent recovery is a challenging task, given the abundance of different technologies performing similar functions. The aim is to choose a technology that achieves the required specifications at the lowest cost, which necessitates systematic evaluation. The integration of a sustainability metric to quantitatively assess the environmental impact of solvent recovery processes is a relatively underexplored area. Existing metrics, such as the E-factor [41], quantify waste production per kilogram of product manufactured. The American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable (ACS-GCIPR) has adopted the Process Mass Intensity (PMI)[36] to evaluate the environmental efficiency

of pharmaceutical processes. However, these metrics are largely mass-dependent and often overlook energy demands along the supply chain of processes. One promising solution is the use of Emergy, an assessment that quantifies the total available energy expended in the transformation processes required to produce a product. Additionally, as discussed in Chapter 2, SPI is a better representation of the ecological burden of a process. Therefore, to have a holistic perspective to solvent recovery, there is the need to simultaneously optimize for both cost and sustainability.

4.1.1 Superstructure-Based Optimization

A superstructure optimization approach is used to capture all the technologies, mixers, splitters, streams, and connections. Four main stages were considered for the recovery process. The first stage comprised solid removal technologies, while the second stage (recovery stage) comprised technologies for liquid separations. The third and fourth stages (purification and refinement stages) are also made up of liquid separation technologies but of higher efficiencies. Additionally, bypass streams are included to eliminate non-essential stages. Binary variables were implemented for technology selection, with the summation of all binary variables at each stage equating to 1 meaning only a single technology can be selected at a stage. Figure 21 shows the proposed superstructure synthesized for the recovery process.

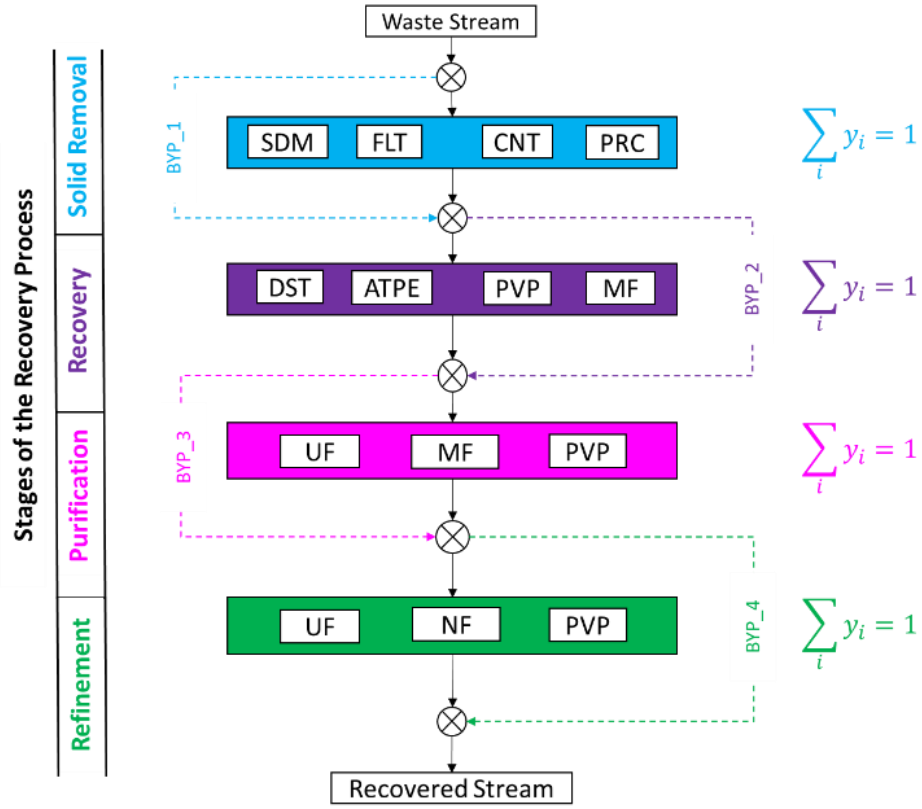


Figure 21. Superstructure for waste solvent recovery showing technologies, streams, and bypasses (SDM – Sedimentation; FLT – Filtration; PRC – Precipitation; CNF – Centrifugation; DST – Distillation; ATPE – Aqueous Two-Phase Extraction; PVP – Pervaporation; MF – Microfiltration; UF – Ultrafiltration; NF – Nanofiltration; BYP_1,2,3,4 – Bypasses 1,2,3, and 4

4.1.2 Multi-Objective Optimization

Multi-objective optimization (MOO) problem, which involve tackling two or more decision criteria that often compete against each other, is used in this analysis. These problems are quite common in real-world scenarios, especially when juggling multiple goals. For instance, in engineering tasks, it is typical to find challenges where there is a need to balance costs, risks, profits, efficiency, sustainability, and safety. This often means that there is no single ‘best’ solution, but rather a range of good solutions that offer different

trade-offs. MOPs can be broadly categorized into two types: generating-based methods and preference-based methods. Generating methods try to find a range of solutions that work well for the problem, resulting in what is known as a Pareto set. This set is super useful because it lets us weigh the pros and cons of different solutions. On the other hand, preference-based methods, like Goal Programming (GP), work a bit differently. Here, the presumption is that the decision-maker has predetermined targets that must be satisfied. In the context of this study, GP is firstly employed. The rationale for this choice arises from the inherent challenges in determining specific upper and lower bounds when multiple technologies compete for selection. Such bounds are essential for deploying generating methods and their associated Pareto sets. GP works by setting some reasonable target values for the decision variables and then trying to minimize the deviation from these targets. For the analysis, each objective function is initially optimized (minimization in this case) and subsequently a range of +/- 30%, with incremental steps of 10%, is employed to ascertain the 'goal' values. In the second part, the ε-constraints method. Since the three objective functions are all minimization problems, each objective is firstly optimized (minimized) to obtain the lower bound. Then an incremental step of 10% is employed as the constraint values. It should be noted that the MOP is implemented for the specialty case study as the complexity of the case study mimics realistic problems. The overarching objective in this research is the minimization of cost, SPI, and Emergy functions. Given the non-linear nature of the mathematical models, which encompass both continuous and binary variables, a Mixed-Integer Nonlinear Programming (MINLP) is formulated for the MOO problem. The Branch-And-Reduce Optimization Navigator (BARON, v 19.12.7)

[147] solver in General Algebraic Modeling Systems (GAMS, v 30.3.0) software was used to solve the problem.

4.1.3 Emergy Analysis for Solvent Recovery

The mathematical models for SPI have already been discussed in section 3.4, hence in this section we discussed the models for Emergy. As discussed in section 2.2, there are three main aspects to quantifying Emergy, namely, renewable energy resources (R), non-renewable natural resources (N), and imported resources (F). To estimate R, N, and F, the flow of the specific resource is multiplied by its corresponding transformity. Transformity is the amount of solar energy (expressed in solar emjoules) required to produce one joule or kg of energy or product [65]. It measures the energy quality from the sun that is used in transformations. Equation (14) is used to estimate the renewable resource Emergy.

$$R = \sum_{i}^{n} Q_{R_i} T_{R_i} \tag{14}$$

Here, Q_R (J/yr or kg/yr or $/yr) is the flow of the renewable resource, T_R (sej/J or sej/kg or sej/$) is the transformity for that resource, while i is the resource and n is the total number of renewable resources. In this analysis, the mass of cooling water and steam are assumed to be renewable resources. For the non-renewable Emergy, Equation (15) is used for the estimation.

$$N = \sum_{i}^{n} Q_{N_i} T_{N_i} \tag{15}$$

Here, Q_N (J/yr or kg/yr or $/yr) is the flow of the non-renewable resource, T_N (sej/J or sej/kg or sej/$) is the transformity for that resource, while i is the resource and n is the total number of non-renewable resources. The non-renewable resources comprised material of construction of each technology, electricity usage, and the waste solvent. For the imported resources, Equation (16) is used for the assessment.

$$F = \sum_{i}^{n} Q_{F_i} T_{F_i} \tag{16}$$

where, Q_F (J/yr or kg/yr or $/yr) is the flow of the imported resource, T_F (sej/J or sej/kg or sej/$) is the transformity for that resource, while i is the resource and n is the total number of imported resources. The imported resources are assumed to be the purchased cost of materials and services, namely, annualized capital, labor, utility, maintenance, and overhead costs. Thus, the total Emergy, T_{Em} (sej/y), is given by Equation (17)

$$T_{Em} = R + N + F \tag{17}$$

These three aspects lead to the quantification of Emergy Yield Ratio (EYR), Environmental Loading Ratio (ELR), and the Emergy Sustainability Index (ESI). EYR, calculated as the total Emergy divided by the imported resources, reveals the economic reliance of the process on imported elements. ELR, computed by dividing non-renewable and imported resources by renewable resources, indicates the environmental pressure exerted by economic activities on the process. Since this analysis considers both the direct and indirect

energy inputs and inputs of material resources, Emergy analysis is far-reaching in terms of environmental accounting. Appendix B contains all the detailed models for the Emergy analysis.

4.2 Case Studies

Pharmaceutical and Specialty chemicals case studies were studied for illustrative purposes. The pharmaceutical case study considers a binary waste stream as a motivating case study while the specialty case is more complex with four contaminants. Below is the solution strategy employed for each case study:

Step#1: Formulate the mathematical models (mass and energy balances, capacity equations, cost equations, SPI equations, Emergy equations)

Step#2: Specify the input and output stream requirements and other parameters

Step#3: Perform simulation using mixed-integer non-linear programming in selecting technologies and quantifying the cost, SPI, and Emergy

Step#4: Define 'goals' for each objective and perform optimization for each combination of 'goals'

Step#5: Use the ε-constraint method for the optimization

Step#6: Perform a similar analysis for incineration

Step#7: Compare solvent recovery with incineration and decide

4.2.1 Pharmaceutical Waste Stream: Problem Statement

Isopropanol (IPA) serves as a critical solvent in the pharmaceutical industry [155], [156], with widespread applications owing to its versatile properties. One notable example of its utility is observed in the synthesis of celecoxib, an active pharmaceutical ingredient (API) that is a key component of the arthritis medication known as Celebrex [157].

4.2.2 Results and Discussion

For this case study, the model comprises 347 equations and 336 variables, of which three are discrete. The model achieves an optimality gap of 0.0001 and takes 15.54 seconds to arrive at a solution. Figure 22 provides a detailed view of the Emergy and SPI distributions for the optimal pathway based on cost optimization, denoted as PVP-UF. The Emergy value of the waste solvent is notably high, approximately 96.6%. This elevated value is attributed to the extensive efforts required to convert resources into solvents, as evidenced by its sizable transformity and increased annual flow rate of waste solvent. In contrast, the Emergy associated with construction materials is the lowest, primarily because the pervaporation and ultrafiltration technologies require smaller capacities. The significant difference in Emergy content between the waste solvent and other components underscores the importance of minimizing waste at its source. This approach is essential for enhancing the environmental sustainability of processes.

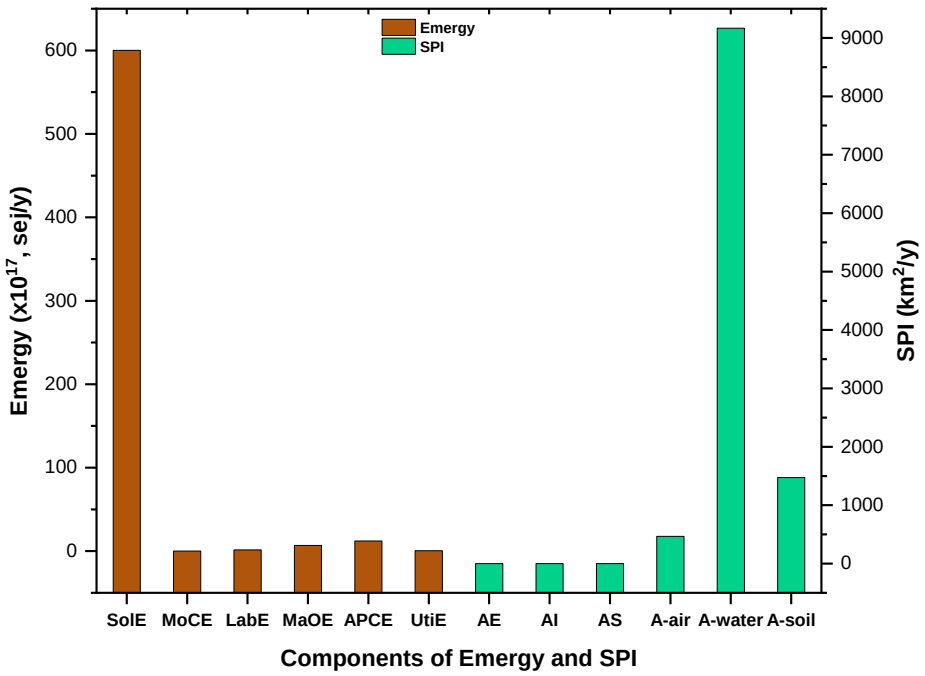


Figure 22. Emergy and SPI breakdown for pharmaceutical case study. [AE, Area for energy consumption; AI, Area for installation; AS, Area for staff usage; A-air, - Area needed to embed air emissions; A-water, - Area needed to embed water emissions; A-soil, - area needed to embed soil emissions; SolE – emergy content of waste solvent; MoCE – emergy associated with material of construction of selected technologies; LabE – emergy associated with labor; MaOE – emergy due to maintenance and overhead cost; APCE – emergy due to annualized purchase cost of technologies; UtiE – emergy associated with utilities]

When evaluating the total Emergy required for incineration, it is found to be 2.88 times higher than the Emergy for solvent recovery for the same amount of waste solvent. The EYR, which measures the economic reliance on imports for the recovery process, stands at 29.58. An EYR value exceeding 10 in processes typically signifies reduced dependency on imports. In contrast, the EYR for incinerating an equivalent amount of solvent is 1, underscoring the reduced sustainability of incineration compared to solvent recovery in this study. Furthermore, the ELR value for solvent recovery is 2.41×10^4,

whereas for incineration it is 4.02×10^3. Such a result aligns with expectations, given that incineration demands a larger influx of renewable resources (notably oxygen from the air) than solvent recovery, which primarily requires cooling water and steam. The Emergy Footprint Intensity (EFI) for incineration is 7.45, a value greater than 1, indicating that the ecological burden surpasses the environment's carrying capacity, highlighting the ecological instability of the process. However, with the introduction of solvent recovery, the EFI is reduced by 25.6%. Moreover, the ESI for the recovery process is 77.9% greater than that of incineration, underscoring the enhanced sustainability of solvent recovery.

Upon evaluating the SPI analysis for the case study, it becomes evident that the area necessary to accommodate water emissions possesses the most significant footprint. This expansive requirement is primarily attributed to the stringent limits on the allowable concentrations of volatile organic carbons in water. Consequently, even minor quantities of organic carbon released into the water necessitate a large, designated area. The total SPI for incineration exceeds that of solvent recovery by 76.2%.

A sensitivity analysis reveals that the cost, SPI, and Emergy values remain within acceptable ranges at flow rates below 1000 kg/hr (the standard case). However, when the flow rate is doubled, the SPI and Emergy of the process surge notably by 49.9% and 44.3% respectively, as shown in Figure 23. This highlights the significant impact of waste solvent generation on the sustainability of the process, emphasizing its sensitivity to changes in flow rate.

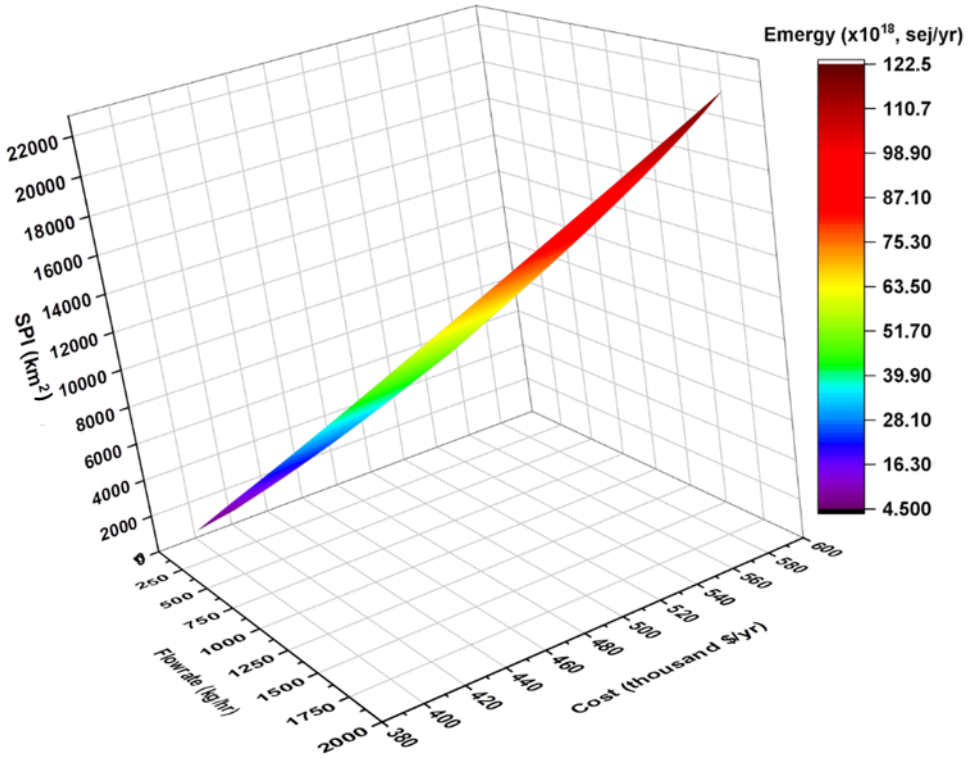


Figure 23. Sensitivity analysis showing how cost, SPI, and Emergy change with varying waste flowrate

4.2.3 Specialty Chemical Waste Stream: Problem Description

In the presented case study, the solvent waste stream consists of 21.3% dimethoxyethane (DME), 1.3% 1-ethoxy-1-methoxy ethane (EME), and 41.3%.

4.2.4 Results and Discussion

The model statistics for the specialty chemical case study encompass 825 equations, 731 continuous variables, 7 discrete variables, an optimality gap of 0.0001, and a solution duration of 96.81 seconds. Figure 24 displays the optimization results for various pathways related to this case study.

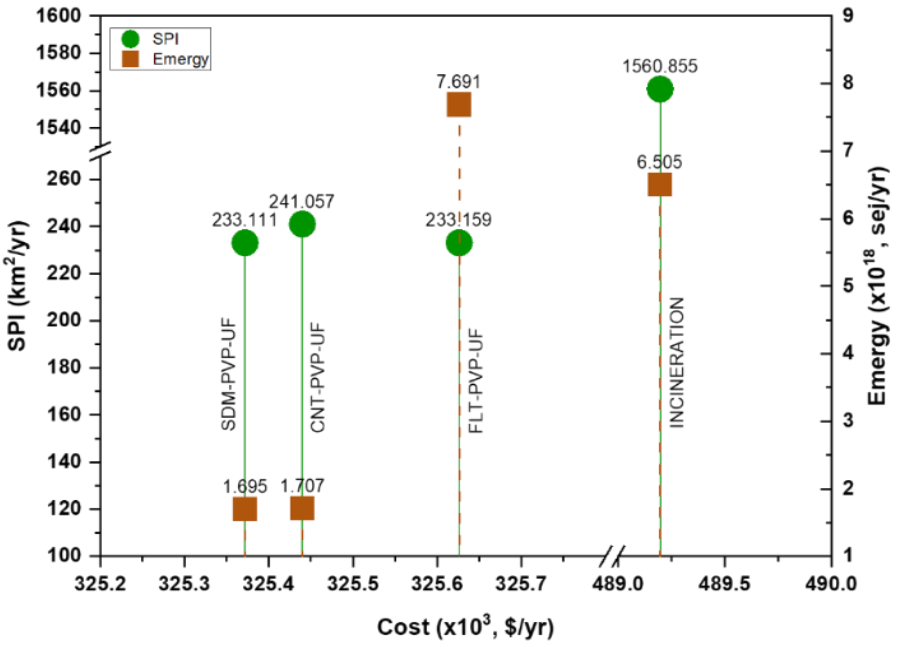


Figure 24. Feasible pathways compared to incineration for specialty case study

The sedimentation-pervaporation-ultrafiltration (SDM-PVP-UF) pathway demonstrates the lowest cost, SPI, and Emergy values, registering at $325,372/yr, 233.1 km^2, and 1.7×10^{18} sej/yr, respectively. Cost-wise, CNT-PVP-UF emerges as the second most favorable option, exhibiting a mere 0.02% increase from the optimal value, followed by the FTT-PVP-UF pathway, which presents a cost increase of 0.08%. The primary factor influencing this cost variation is the diverse capital and utility expenses associated with sedimentation, filtration, and centrifugation units. Regarding SPI, the FTT-PVP-UF pathway is more advantageous than CNT-PVP-UF due to the energy-intensive character of centrifugation, leading to an increased requirement for area to provide energy, thus resulting in an elevated SPI value. FLT-PVP-UF records the highest Emergy value, while incineration stands out as the pathway with the most substantial cost and SPI values.

Elevated emissions from the incineration process play a pivotal role in its higher SPI value. Additionally, the reliance on fossil-based fuel augments the area required for nonrenewable resource consumption, accounting for 42.3% of the total SPI value for the incineration process. Hence, the SPI offers valuable insights into the potential exploration of alternative solutions, such as the incorporation of renewable raw materials in the incineration process. Considering the economic reliance on imports, the EYR for incineration stands at 1.09, compared to 1.08 for the standard scenario (SDM-PVP-UF). However, the ELR for incineration exceeds the standard scenario by 95.0%. This pattern aligns with observations from the pharmaceutical case study, where the renewable resource input (oxygen, sourced from air) for the incineration process surpasses the input required for the recovery process, such as cooling water and steam. An increased ELR for the recovery process results in a diminished emergy sustainability index (6.80×10^{-6}) in comparison to incineration (1.37×10^{-4}), highlighting solvent recovery as a more environmentally friendly option.

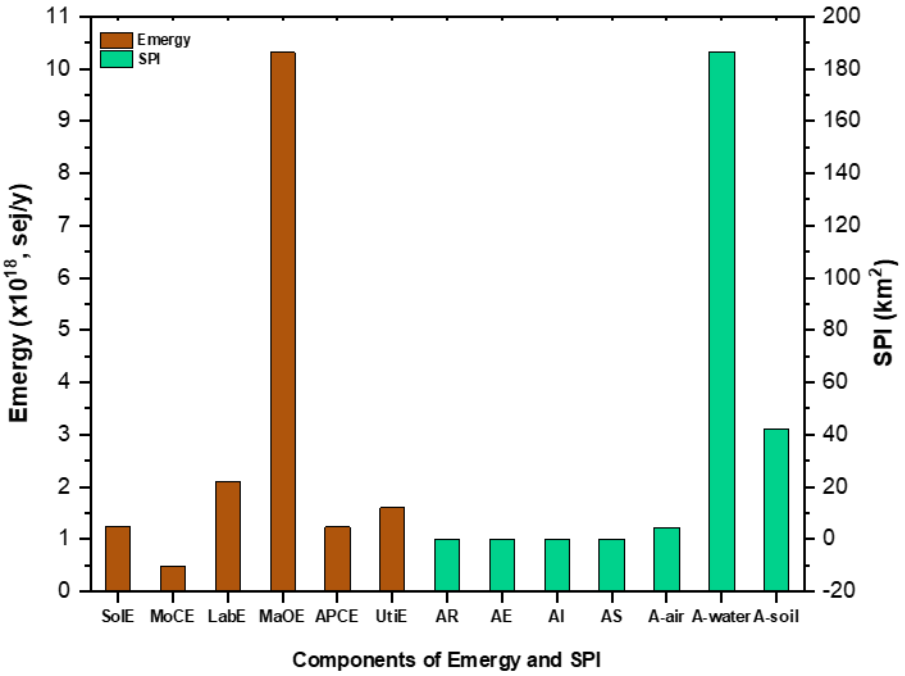


Figure 25. Breakdown of SPI and Emergy analysis for SDM-PVP-UF pathway for specialty case study

Figure 25 presents the SPI and Emergy distribution for the SDM-PVP-UF pathway. Unlike the Emergy analysis observed in the pharmaceutical case study, the maintenance and overhead costs are the predominant contributors to the overall Emergy. In terms of SPI value, the area required to assimilate water emissions remains the most significant contributor, a trend consistent with findings from the pharmaceutical case study.

A sensitivity analysis is conducted to assess the variations in cost, SPI, and Emergy in response to different waste flow rates as shown in Figure 26. In every scenario, the SDM-PVP-UF pathway emerges as the most favorable, primarily due to the reduced operating expenses linked with the sedimentation process. Variations in the process flow rate exert the most pronounced influence on the SPI, followed by its impact on Emergy. This

relationship stems from the direct correlation between flow rate and emissions, which subsequently affects the required area.

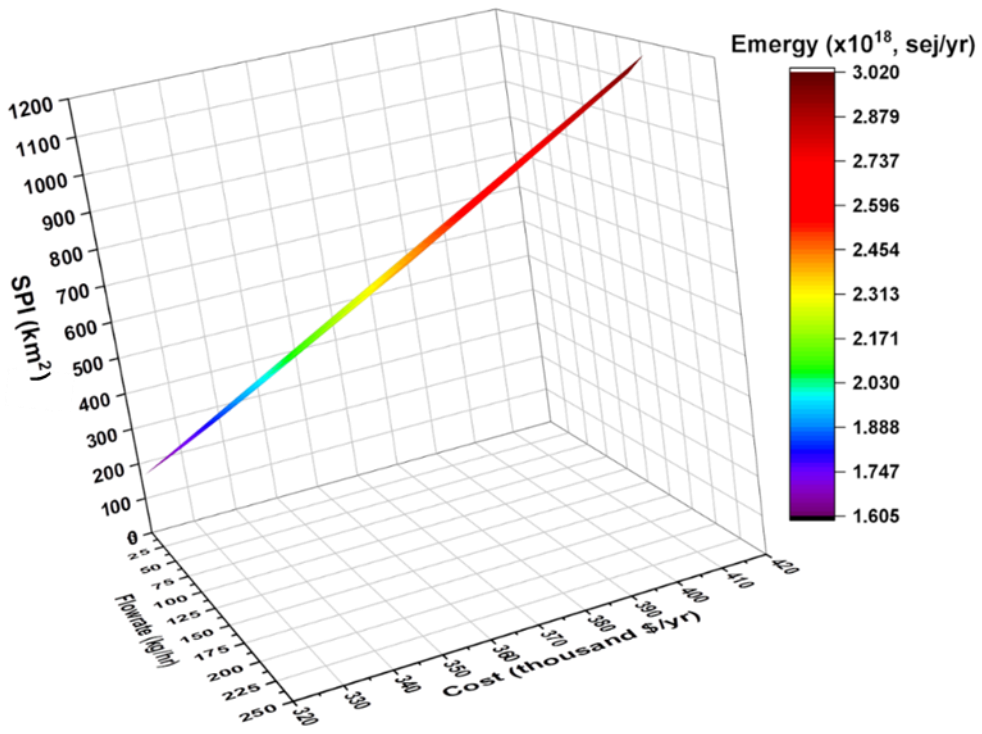


Figure 26. Sensitivity analysis showing how cost, SPI, and Emergy change with varying waste flowrate

Figure 27 shows the pareto chart generated for this case study using the constraint-based approach. Due to the energy intensive nature of CNT, the pathways shown with green circle correspond to CNT-PVP-UF. The data points indicated by the red circle correspond to SDM-PVP-UF. It can be observed that due to the bigger area needed by the SDM unit, the networks have higher SPI values. The datapoints with the blue circles

correspond to FLT-PVP-UF pathway. These pathways have higher cost association, and this can be attributed to the requirement of bigger filter areas, consequently, the higher cost.

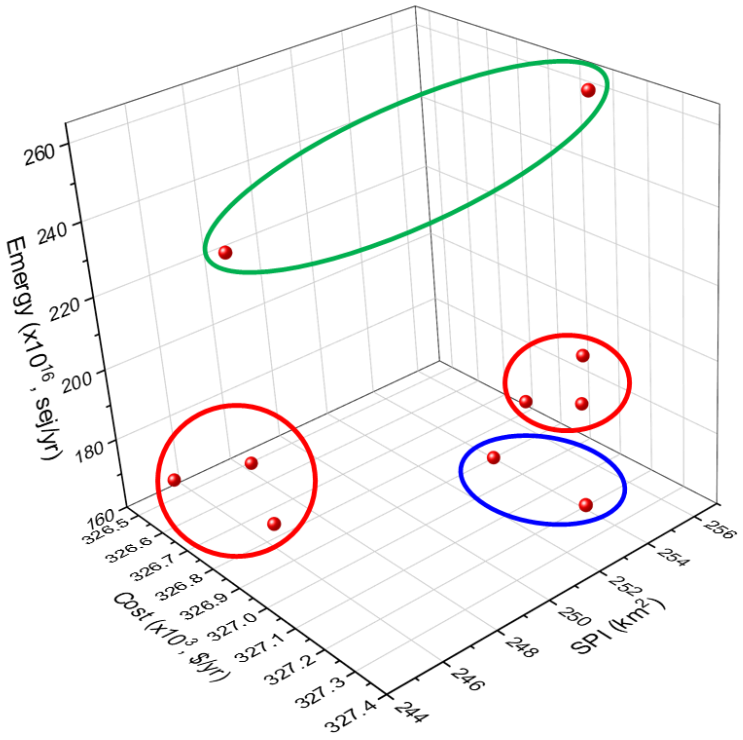


Figure 27. Pareto chart for specialty case study

4.3 Conclusions

In the presented framework, the scope encompasses economic considerations, the Sustainable Process Index (SPI), and Emergy metrics. Consequently, challenges related to solvent recovery have been reformulated into a multi-objective optimization problem, addressed using superstructure and mixed-integer nonlinear programming methodologies. Findings reveal that opting for solvent recovery over incineration can mitigate

approximately 76−85% of the ecological and energy burdens. It is noted that at lower flow rates, solvent recovery becomes less economically viable than incineration, given the associated capital and operational expenses. Moreover, when the flow rate is doubled in both case studies, there is a pronounced escalation in the SPI and Emergy footprints, underscoring the imperative to curtail solvent waste generation to further advance sustainability goals.

Chapter 5

Predicting Life Cycle Impacts of Chemicals: A Machine Learning Approach

Text and figures used for this Chapter are pending publication.

5.1 Background

In the face of urgent challenges posed by climate change and heightened environmental concerns, industries are under intense scrutiny regarding the environmental consequences of their operations, especially with metrics like Global Warming Potential (GWP) taking center stage. This brings into focus the critical role of early-stage process synthesis, the stage where potential operational processes are formulated and evaluated. Decisions taken during this phase set the tone for the environmental repercussions of the entire operation. It is within this scenario that Machine Learning (ML) offers a transformative approach. By incorporating ML at this nascent stage, industries can effectively leverage its capabilities for prompt, precise, and thorough assessments of sustainability.

One of the standout benefits of ML is its remarkable efficiency to handle systems without intuitive connections. Once appropriately trained, ML models can rapidly forecast sustainability metrics, enabling swift design adjustments and enhancements. Furthermore, the adaptability of ML allows it to seamlessly integrate with optimization strategies, assisting industries in designing processes that strike an optimal balance between environmental and economic concerns. From a financial perspective, ML proves invaluable. By pinpointing and tackling sustainability-related issues at the outset, industries can sidestep expensive alterations in later stages, resulting in significant cost savings. In

summary, integrating Machine Learning during the early-stage process synthesis marks a progressive step for industries committed to paving a sustainable path forward.

Thermodynamic attributes like enthalpy, entropy, and Gibbs free energy [158], [159] offer valuable information about the energy needs of a process, operational efficiency, and overall viability . These factors play a critical role in determining the energy consumption of a process, subsequently affecting key sustainability metrics such as Global Warming Potential (GWP) and the total carbon footprint. Conversely, molecular characteristics [160, p. 5], [161, p. 6], which encompass molecular weight, bond energies, and functional groups, give insight into the inherent qualities of chemical substances. These characteristics serve as indicators of the reactivity[162], potential toxicity [163], and environmental impact [164] of a chemical. Often times, information on both thermodynamic and molecular properties is accessible during the initial stages of process design. Thus, by constructing an ML model that uses thermodynamic and molecular descriptors as input features and sustainability metrics as outputs, predictions can be made for both new and existing chemicals that may not yet have established sustainability metrics, enabling a more comprehensive and informed approach to evaluating sustainability right at the early-stage synthesis.

In this chapter, ML models are developed to predict four sustainability metrics with molecular descriptors and thermodynamic properties as input features. The concept of ML is introduced where the fundamentals pertaining to developing an ML model is discussed briefly. The data acquisition regarding the sources of the thermodynamic and molecular properties of the chemicals are discussed. Further, the preparation of the data for model development is discussed. Various types of preprocessing methods used in this work is

discussed. Upon completion of the data preprocessing, the model development is discussed. Finally, the model evaluation and results are discussed together with the implementation of the model to a case. Figure 28 shows the various aspects for the ML implementation process.

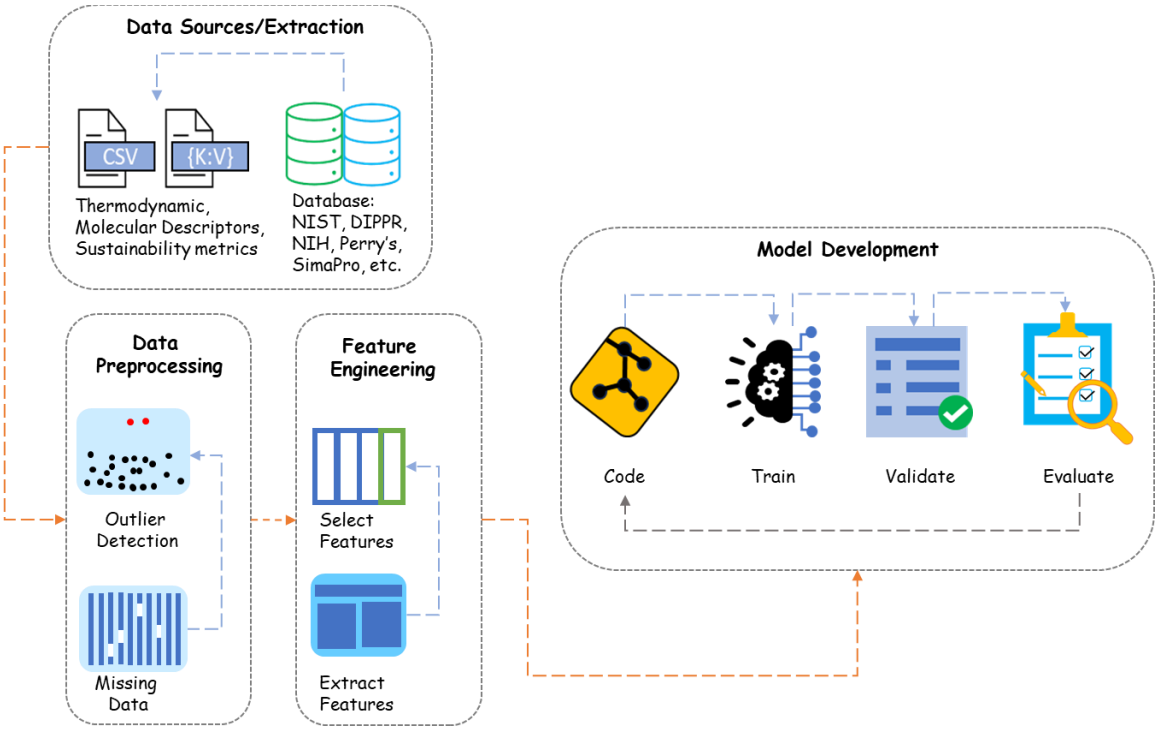


Figure 28. Schematic showing the stepwise approach to ML implementation

5.2 Fundamentals of Machine Learning

This section sub-section is dedicated to describing the fundamentals of ML systems, starting with the types of ML commonly encountered in the PSE space. Sub-section 5.2.2 introduces the concepts of feature selection and reducing high dimensional

data into a lower dimension, while retaining much information about the data. The last sub-section deals with ways to evaluate and validate a developed ML model.

5.2.1 Types of Machine Learning Algorithms

Supervised Learning (SL) [165], [166] refers to a subset of machine learning algorithms trained on labeled datasets. These datasets encompass both the input data (features) and corresponding output values (target labels). The primary objective of SL is to discern a function that maps inputs to outputs, enabling the model to predict accurately for previously unencountered data points [167]. Common applications of SL encompass classification and regression tasks, with prevalent methods including linear and logistic regression, support vector machines (SVM), ensemble models, and neural networks [54], [77], [168].

Contrastingly, Unsupervised Learning (UL) pertains to algorithms that process datasets lacking predefined labels or output values [77], [168]. The central purpose of UL is the identification of inherent patterns or structures within the data, which might manifest as clusters, relationships among variables, or concealed representations. Prominent applications of UL focus on clustering and dimensionality reduction, with techniques such as k-means clustering, hierarchical clustering, and principal component analysis (PCA) being widespread [169], [170].

Reinforcement Learning (RL) [171], another machine learning category, involves an agent that learns to make decisions within a particular environment based on feedback, which could be in the form of rewards or penalties. The ultimate goal for RL models is to optimize the agent's actions to cumulatively maximize rewards over time. RL is especially valuable in scenarios where the optimal solution is elusive and requires a degree of

explorative trial and error. Notable RL techniques include Q-learning, deep Q-networks (DQN), and policy gradients [171], [172].

5.2.2 Feature Selection and Dimensionality Reduction

Feature selection is crucial for enhancing the accuracy of predictive models by identifying the most relevant variables and simplifying the dataset. One approach, known as filter methods, assesses features individually using metrics like correlation and mutual information [173], [174]. These metrics gauge how each feature relates to the target variable, enabling the selection of the most strongly correlated features. Another approach involves iterative algorithms that evaluate feature subsets based on a given model's performance. Common iterative techniques include forward selection, backward elimination, and recursive feature elimination.

Dimensionality reduction techniques streamline large datasets, enhancing the efficacy of predictive models while conserving computational resources. Principal Component Analysis (PCA) [56], [175] is a widely used linear method that identifies principal components, or primary directions, showcasing the most data variation. While reducing data dimensions, PCA maintains much of the original data's variability. Beyond these, techniques like feature agglomeration and manifold learning, aggregate features based on similarity and retain intricate data relationships, addressing complexities beyond the reach of linear methods like PCA and SVD [56].

5.2.3 Model Evaluation and Validation

For accurate evaluation and validation of a predictive model, one method is to divide the data into training, validation, and test sets. The training set is used to train the model, the validation set is used for tuning hyperparameters, and the test set measures

efficacy of new data on the model [77]. This division helps ensure the model neither overfits nor underfits, offering a genuine reflection of its potential real-world performance. Cross-validation offers another evaluation approach: the training data is segmented into ‘k’ subsets, with the model trained and validated ‘k’ times, each time using a different subset as validation data [176]. The performance of the model is then averaged over these iterations for a consistent generalization estimate. Depending on the nature of the problem (regression or classification), various metrics, such as mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and R-squared, can assess the accuracy and reliability of the model.

5.3 Data Acquisition

A list of 350 common solvents is assembled covering a wide spectrum of molecules such as alcohols, esters, hydrocarbons, and ethers. The data set acquired is in two parts, the first part is the feature set data, while the second part is the label set data set. The feature set comprised of two types, the thermodynamic and molecular descriptor data. Thus, the feature set entails the chemical properties from which the model learns from. The label data is the one that the developed model tries to predict.

A total of 15 thermodynamic properties is acquired for each chemical. Some of the thermodynamic properties used in the model comprised critical temperature, critical pressure, critical volume, heat capacity, boiling point, standard Gibbs-free energy, among others. To acquire this data, the first step is extracting the SMILES string and chemical formula for each chemical. The SMILES string is generated from CIRpy (version 1.0.2), a Python library that serves as the interface for the Chemical Identifier Resolver (CIR) [177]. This library searches the National Institute of Health database for the structures of the

chemicals. In the next step, the extracted SMILES string is used to extract the respective thermodynamic properties using ‘chemicals’ (version 1.1.4) [178] and ‘thermo’ (version 0.2.26), two Python libraries which contain a database of an extensive compilation of pure and calculated chemical properties. The local databank found in both libraries is a compilation made from National Institute of Standards and Technology (NIST) [179], Design Institute for Physical Properties (DIPPR) [180], PubChem (by the National Institute of Health) [181], CRC Handbook, Perry’s Chemical Engineers’ Handbook , and various scientific papers and publications. Thus, over 20,000 chemicals and their corresponding thermodynamic properties are available as a local databank within these libraries.

For the molecular descriptor properties, RDKit (version 2023.3.3) [182], which is also an open-source Python library, is used to acquire 200 molecular descriptors for each chemical. RDKit is a comprehensive collection of cheminformatics toolkits which can be used to compute a wide range of molecular descriptors. It is commonly used in the cheminformatic space for drug discovery, and toxicological studies. Some of the molecular properties include molecular weight, carbon count, maximum partial charge, functional group, number of heterogeneous atoms, number of radical atoms, number of aliphatic rings, among others.

The last step in the data acquisition is extraction of the label data. For this, SimaPro® (version 9.4.0.2) [75] is used to gather sustainability metric data for cradle-to-gate of each chemical. SimaPro is an LCA software tool, which facilitates a detailed analysis of the life cycle of a product. A standout feature of SimaPro is its comprehensive database. This extensive resource includes a wide array of international datasets, detailing various aspects such as raw materials, different manufacturing processes, transportation

methods, and end-of-life practices. Four endpoint metrics are considered, namely, human health impact (HHI), ecosystem quality impact (EQI), global warming potential (GWP), and resource utilization impact (RUI) for each chemical. These four endpoint metrics are chosen due to decision-making relevance, ease of communication, and depth of analysis. Thus, using SimaPro, the impact per kilogram of each chemical for the listed sustainability metrics are assembled.

5.4 Data Preprocessing

The extracted data comes with a lot of inconsistencies, such as missing data, and outliers. Additionally, each feature has a different range of values. Therefore, to be consistent, and reduce the problem of overfitting or underfitting, the feature set data needs to be preprocessed before model development. Thus, data preprocessing is an ensemble of techniques aimed at transforming raw data into a format more suitable for model development [56]. The first step is to find ways to replace missing label data because the label data had 85 out of the 350 datapoints missing. Ideally, it is best to remove rows within the dataset that have missing label data, however, ML models thrive on large datasets, hence the need to find ways to replace the missing label data. The k-Nearest Neighbors (kNN) [183] method of data imputation is used for this analysis. The idea with kNN is to identify ‘k’ neighboring points or samples within the dataset that are similar or close in space. The space is normally a Euclidean distance or Manhattan distance space. The mean value of the ‘k’ neighbors is used to replace the missing data.

Upon completion of the missing data imputation, the next step is to determine outliers. This is done based on the label data, since the aim of the prediction is based on the label data. The first step is to normalize the data about the mean and standard deviation,

resulting in the z-score. Once normalized, the datapoints that fall beyond the three standard deviations from the mean is considered as an outlier [184]. The original label dataset did not follow are normal distribution, however, since the z-score outlier detection criteria works on the assumption that the data follows a normal distribution, a log transformation is initially implemented (see Figure 29) on the label data before the application of the outlier removal. The resulting dataset is then reverted to the original.

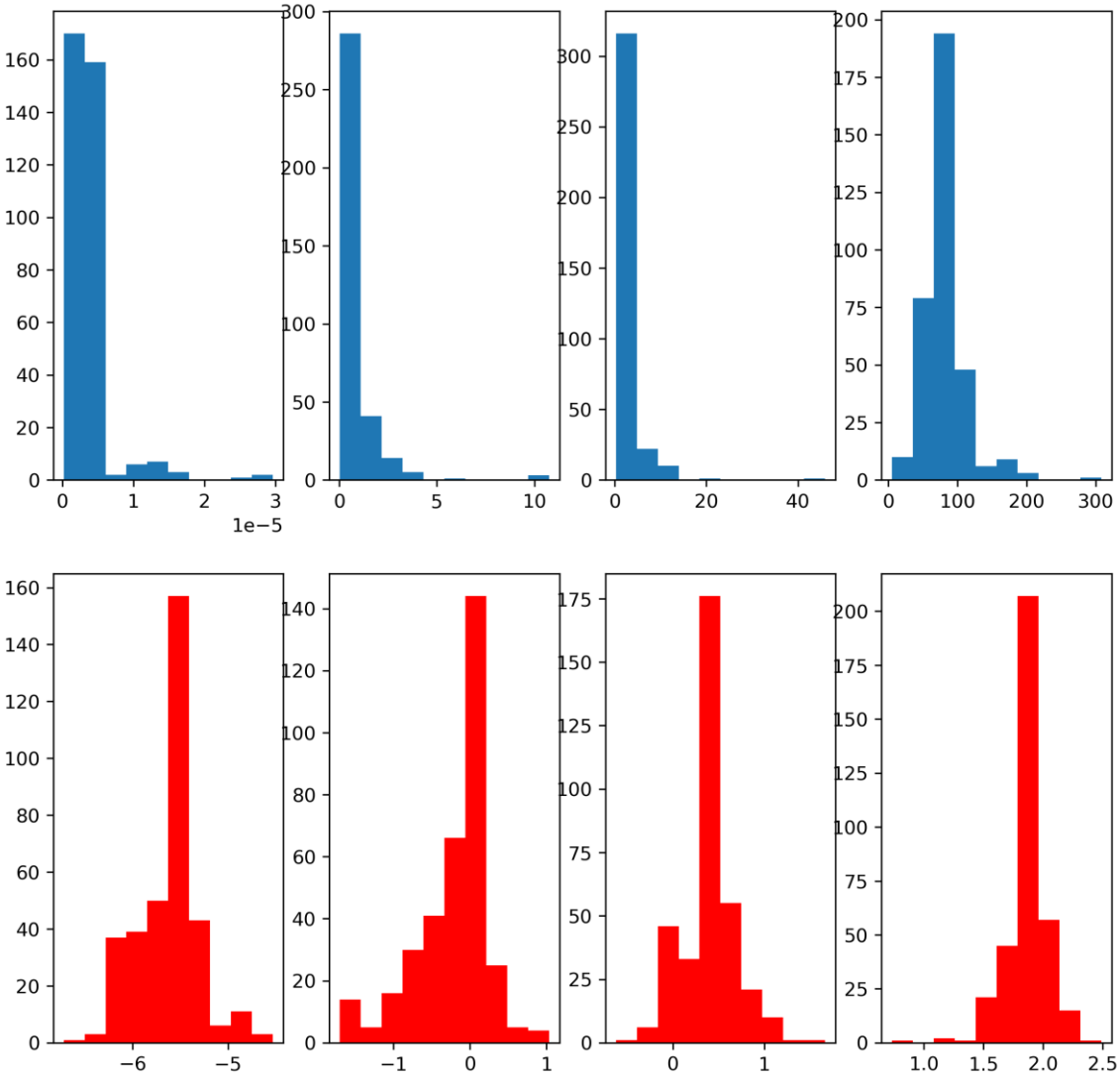


Figure 29. Log transformation of label data. The blue histogram shows the distribution of the actual label dataset while the red histogram shows the log-transform form of the data.

The last step in the data preprocessing is to scale the feature dataset. In algorithms like ANN that calculate the distances between data points, the scale of the features significantly affects the results. If feature scaling is not applied, a feature with a broader

range and higher order of magnitude might disproportionately influence the outcome, regardless of its actual relevance or importance. Additionally, distance-dependent algorithms tend to converge faster when the features are on a similar scale. The features for this work are scaled to be in a range of 0 and 5.

5.5 Model Development

The next crucial step is the model development phase which consists of feature selection, model training, hyperparameter tuning and validation, and model evaluation. The following subsections delve into the various strategies used at each step of the development process.

5.5.1 Feature Selection

Due to the high number of features available, there is a need to select features that make the highest contribution to the model. By focusing on only features that have significant importance to the model, there is reduced computational time, and redundant features are neglected, thus, improving the model accuracy. A total of 10 features are selected with 5 from each thermodynamic and molecular feature set. This is to make sure each feature set has an equal contribution to the model. Furthermore, this makes the model more realistic in terms of its usage by users since only 10 properties are needed to make predictions for the specific chemical in question. Moreover, rather than using the same 10 features to make predictions for the four metrics, each metric has its own distinct feature set, hence, only the features that significantly have an impact on making predictions for that metric is actually used. Additionally, doing so largely reduces redundant calculations and saves computational time. To select the top 5 features from each feature set, a Sequential Backward Feature Selection (SBFS) [56], [185] criterion is used. The idea with

SBFS is to generate all the possible subset of size, *n-1*, from the original feature set, *n*. For each subset, a scoring evaluation is performed based on a defined model. The subset with the best scoring metric is selected for the next iteration. The process is repeated until the number of required features is reached. For this analysis, the predefined model used is a linear regression, with the scoring being MSE.

5.5.2 Model Training and Hyperparameter Tuning

Once the feature set is finalized for each label, the next step is to select the ML model for training. Two ML models are tested in this work, namely, eXtreme Gradient Boosting (XGBoost) and Artificial Neural Network (ANN).

XGBoost [186], [187], which is an ensemble ML model, is an advanced and efficient implementation of the gradient boosting framework designed to optimize large-scale ML problems. In essence, it progressively builds models by correcting the inaccuracies of prior models. The adjustments are guided by the gradient descent method, which identifies and addresses the weaknesses in the current ensemble by adding a new decision tree. This iterative process continues until the error reaches a set limit or once a specified number of trees have been incorporated. For this work, the data is divided into training, validation, and testing sets, as discussed in section 5.3.3. To improve the performance of the model developed, certain hyperparameters of the XGBoost (version 1.7.6) model need to be optimized. Four to six hyperparameters which have the highest impact on the model are chosen and tuned using the ‘hyperopt’ (version, 0.2.7) [188] library, a global optimization package which uses a Bayesian optimization [189] framework. The hyperparameters include the maximum depth of a tree(‘max_depth’), learning rate (‘learning_rate’), the number of trees to include (‘n_estimators’), the

minimum number of instance weight needed in a child node ('min_child_weight'), the fraction of samples that are chosen randomly to grow the trees ('subsample'), and the fraction of features randomly chosen to grow each tree. The validation set is used to determine the optimal hyperparameters by defining an objective function to minimize MSE between the true and predicted values for the validation set, after training the model on the training set. Thus, in all instances, the test set is only used for model evaluation to observe the generalizability of the developed model.

The second model developed is ANN [56], [176]. ANN is a subset of ML algorithms, derived from the way biological neural networks within the human brain operate. Central to these models are elements called 'neurons' or nodes, organized in layers, responsible for processing and conveying data. A neuron accepts several inputs, computes based on these inputs, and generates an output. Each input has an associated weight, adjusted during training to improve the accuracy of predictions. After aggregating the inputs considering their weights, the neuron uses an activation function to determine its final output. The ANN model is constructed using the TensorFlow (version 2.12.0) [176], [190] library. The hyperparameters tuned for this model include the learning rate, the number of hidden layers, number of neurons for input and hidden layers, dropout rate, type of activation function, type of loss function, the batch size, and the number of epochs. The Adam optimizer is used during the training process. Similar to XGBoost, the hyperparameter tuning is done on the validation set. For each model, the dataset is divided into 80% training, 10% validation, and 10% testing.

For model evaluation, the R^2 value and the Root-Mean-Squared-Error (RMSE) are used. Both the coefficient of determination, commonly referred to as R^2, and the RMSE

are crucial metrics that provide a detailed view of the performance of the model. R^2 serves as a measure that indicates the proportion of variance in the outcome variable that can be explained by the predictors in the model. It effectively captures the relative fit to the data of the model. Additionally, due to its intuitive nature and widespread acceptance, R^2 has become an essential tool for communicating the goodness of fit to a diverse audience. Conversely, RMSE provides insight into the accuracy of the model by measuring the average magnitude of errors between the predicted and actual outcomes. It offers a direct, absolute measure of the fit of the model. In essence, while R^2 provides a comparative view of the fit of the model in relation to the variance of the data, RMSE quantifies the average deviation in predictions. Hence, leveraging both metrics simultaneously yields a well-rounded evaluation, highlighting potential concerns such as overfitting. By ensuring predictions are aligned both in terms of relative and absolute fit with the actual values, this dual assessment strategy bolsters the credibility and reliability of the results from the model.

5.6 Model Results and Discussion

Table 10 shows the selected properties for each metric after the implementation of SBFS. For the thermodynamic properties, critical temperature and heat capacity are selected for each metric. XLogP and boiling point are the next properties found in three of the four metrics. For the molecular descriptors, HallKierAlpha, which captures the three dimensionality in terms of shape representation and branching of the molecule, is selected for three out of the four metrics.

Table 10

Selected Features for Each Sustainability Metric

Metric	Selected Features	
	Thermodynamic Properties	Molecular Descriptors
HHI	heat of vaporization, heat capacity, XLogP, acentric factor, critical temperature	Chi0, HallKierAlpha, SMR_VSA7, VSA_EState6, NumValenceElectrons
EQI	heat capacity, standard formation enthalpy (gas), boiling Point, critical temperature, critical volume	Chi2v, BertzCT, HallKierAlpha, qed, fr_halogen
GWP	Heat capacity, boiling point, XLogP, critical temperature, critical molar volume	BertzCT, ExactMolWt, HallKierAlpha, PEOE_VSA6, NOCount
RUI	heat capacity, boiling point, XLogP, critical pressure, critical temperature	ExactMolWt, MaxAbsPartialCharge, MaxPartialCharge, NumRotatableBonds, SMR_VSA2

Figure 30 shows a parity plot for each of the metric from the ANN model. For the HHI (Figure 30 (a)), The test set performs even better from both training and validation set with a value of 0.957 for the R^2 score. It can also be observed that the RMSE for the train-validation-test sets are close enough, indicating acceptable predictions, Furthermore, the predictions vary between 0.606 – 12.138 (10^{-6}) DALY/kg_{chem} for a 95% confidence interval. Despite the large difference in the for the R^2 EQI metric (Figure 30 (b)), the RMSE is within acceptable limits. Hence, prediction from this model is also acceptable, however, more effort must be made to refine the model to improve the R^2 value for a much more

reliable prediction. For a 95% confidence interval, the prediction ranges from 0.202 – 2.341 PDF.m^2.yr, however from Figure 30 (b), it is evident that an appreciable number of predictions fall above the 2.341 upper limit. Therefore, some chemicals with extreme impacts on EQI at the edges of the training data are currently predicted and need further improvements. Results from the GWP model also shows good prediction however, similar to the case of EQI, the model overfits the training data, hence the observed significant difference of the R^2 between the training and testing sets, however, this model is better compared to the EQI predictions. For the GWP model, the prediction ranges from 0.816 to 8.375 $kgCO_2$-eq/kg_{chem} for a 95% confidence interval. Lastly, the RUI metric, which can also be interpreted as the Cumulative Energy Demand (CED) for the production of the chemical gives very good predictions, with predictions ranging from 3.667 to 16.960 ($x10^1$) MJ/kg_{chem} for a 95% confidence interval.

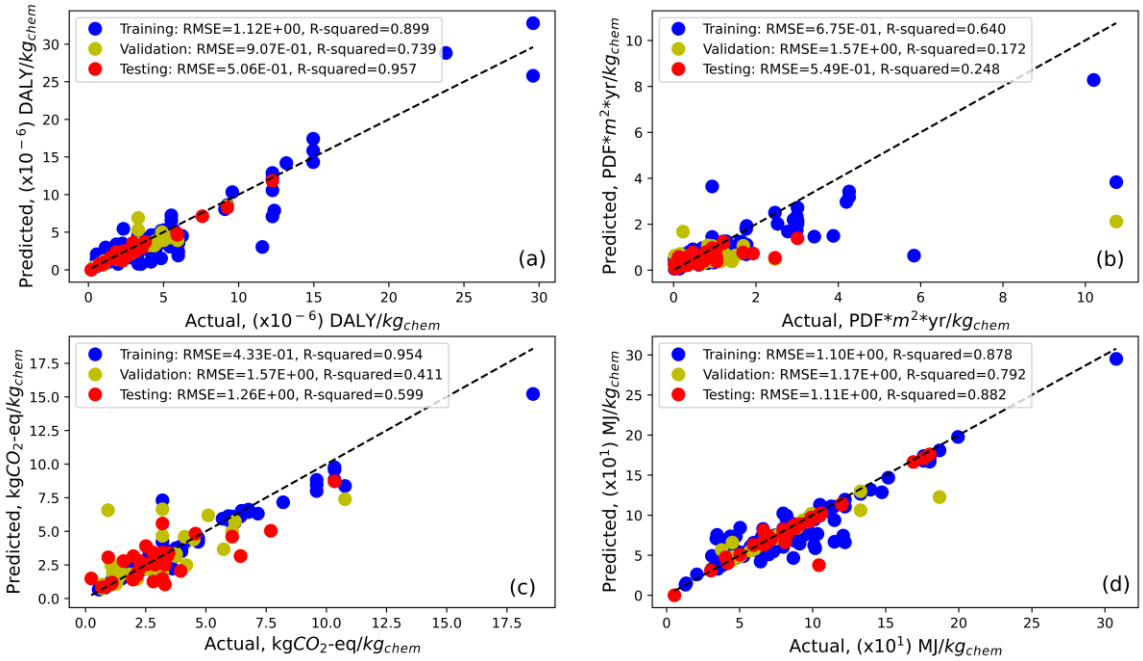


Figure 30. Parity plot for each metric from the ANN model. (a) is for human health impact (HHI), (b) is for ecosystem quality impact (EQI), (c) is for global warming potential (GWP), (d) is for resource utilization impact (RUI)

To understand which features out of the 10 have the most impact to the model performance, permutation importance is implemented. Figure 31 shows the results of each feature on the corresponding metric. It can be observed that for HHI, the molecular descriptors have the highest impact on the model as the top 4 features are all molecular descriptors. Similar observations are made for EQI and RUI. However, for GWP the thermodynamic features have the highest impact on the model. One notable observation about the model for EQI is that the 'boiling point' feature seems to have a negative impact on the model. Hence, one of the ways to improve that model could be to redevelop the model without the inclusion of that feature.

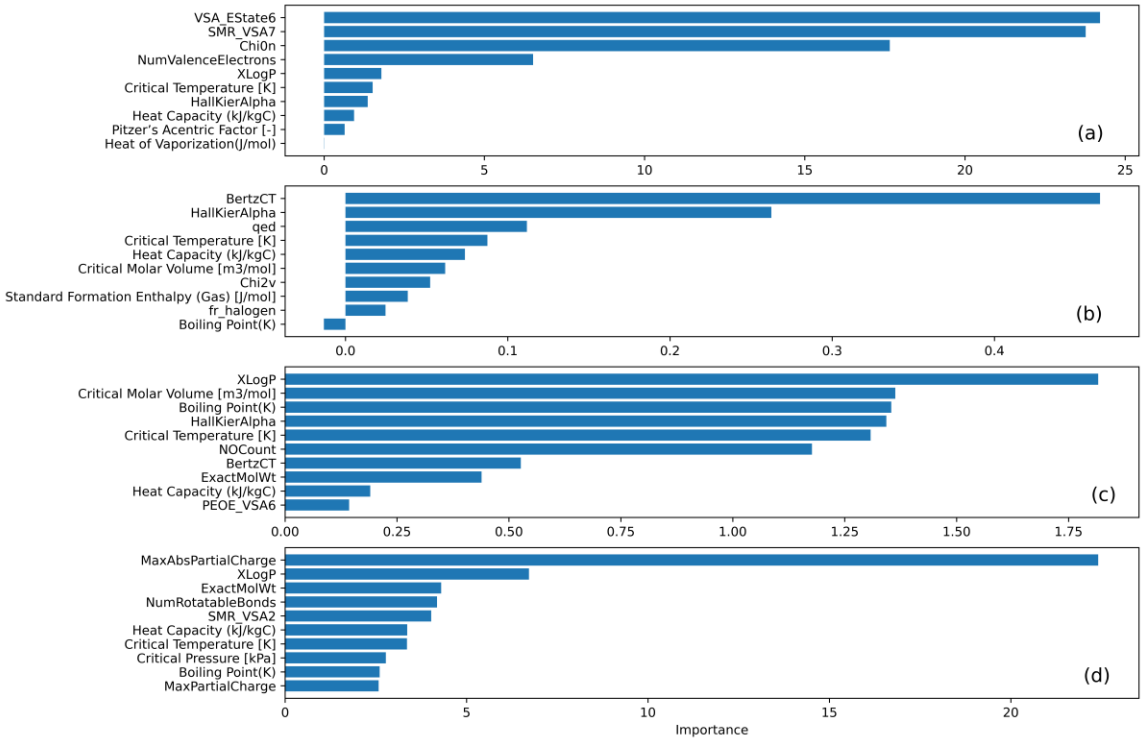


Figure 31. Feature importance for each metric from the ANN model. (a) is for human health impact (HHI), (b) is for ecosystem quality impact (EQI), (c) is for global warming potential (GWP), (d) is for resource utilization impact (RUI)

Sensitivity analysis is implemented to observe the impact of each feature on the corresponding metric. The sensitivity analysis conducted in this context leverages a bootstrap-like method to perturb feature values and observe the resultant variations in the predictions for each model. Bootstrapping [191], in statistical terms, is a resampling technique used to estimate statistics on a population by sampling a dataset with replacement. It is widely acclaimed for its efficacy in approximating the distribution of various statistics without necessitating the assumption of normality. In the context of this sensitivity analysis, the bootstrapping concept is adapted to assess the robustness and behavior of a machine learning model, particularly an artificial neural network.

The implemented method involves systematically altering the values of a specific feature across a defined range while keeping other features constant [191], [192]. This range is determined based on the observed values of the feature in the dataset, typically spanning from the minimum to the maximum observed value. For each perturbed value, the prediction for the model is computed, and the resultant outputs are recorded. This process is akin to “sampling” across the possible values of the feature and observing the corresponding “response” of the model. In this analysis, the bootstrapping is done using the testing set to make sure the analysis is done on unseen data by the respective models.

Figure 32 shows how sensitive each feature from the human health impact metric is for the ANN model. It can be observed that apart from the heat of vaporization and acentric factor, the model is sensitive to the remaining features.

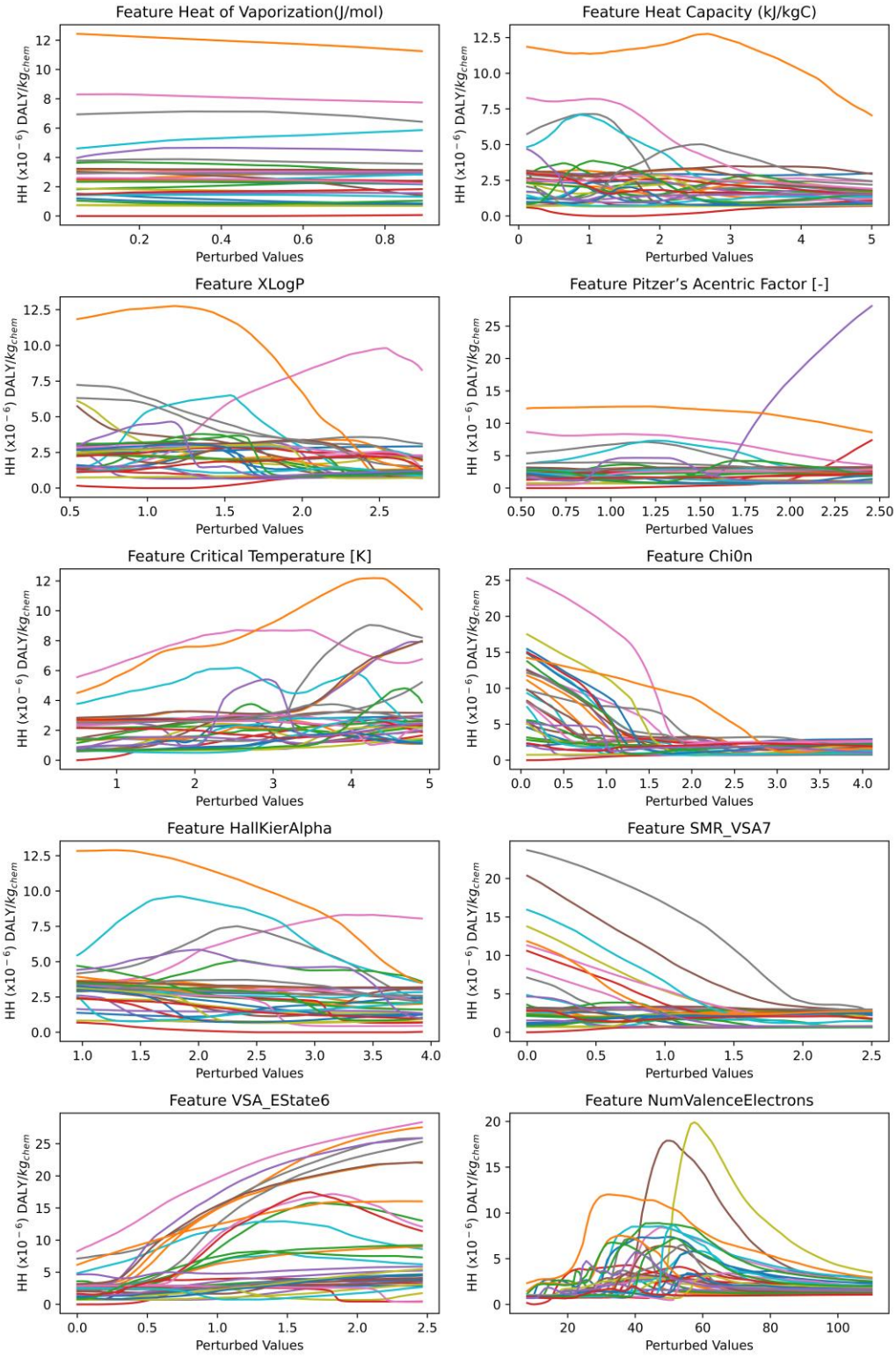


Figure 32. Sensitivity analysis for each feature for human health impact of ANN model

Figure 33 shows the sensitivity of each feature for the ecosystem quality impact metric for the ANN model. It can be observed that each feature fairly impacts the prediction. This is due to the poor predictions for the metric by the ANN model.

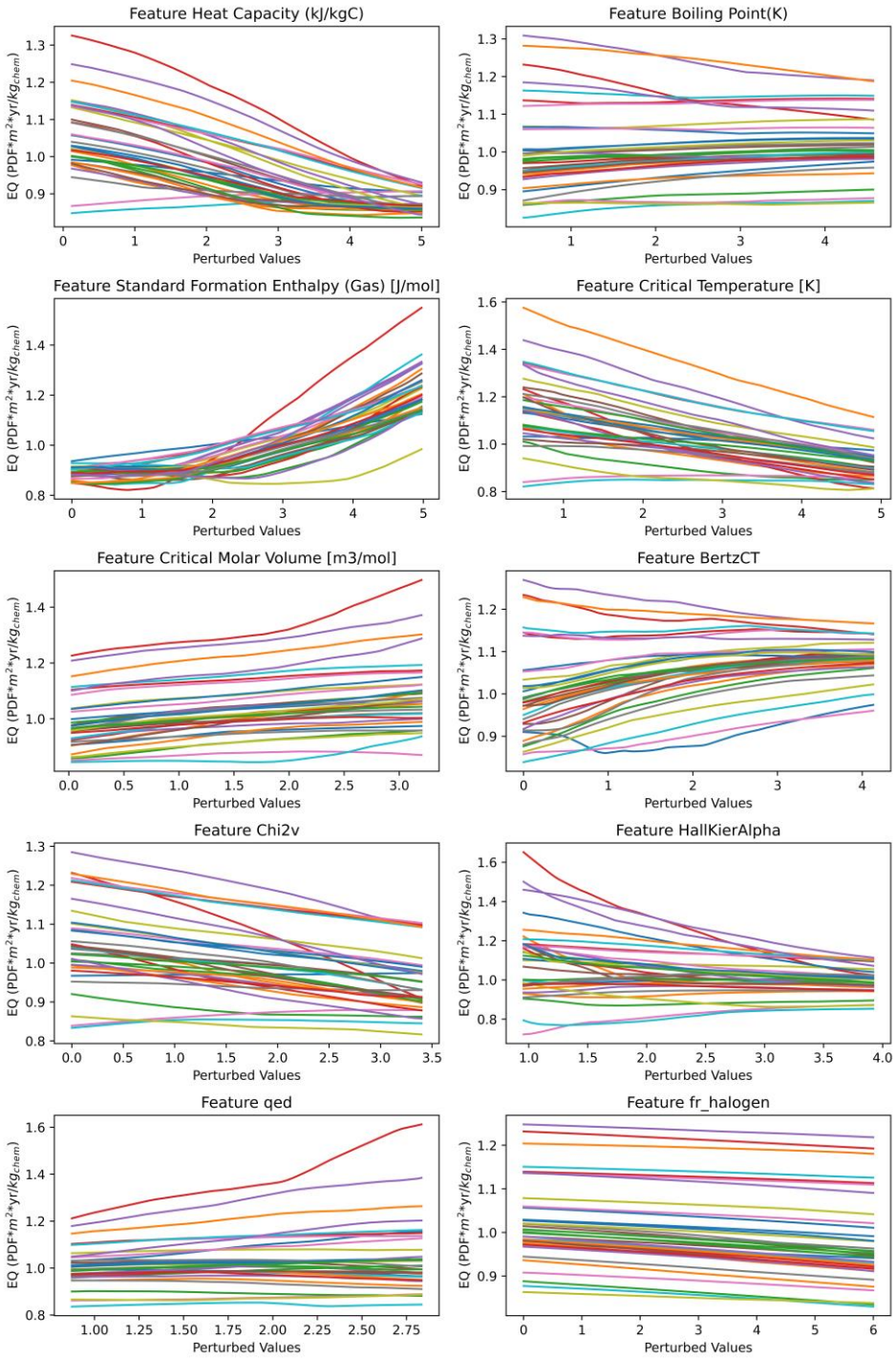


Figure 33. Sensitivity analysis for each feature for ecosystem quality impact of ANN model

Figure 34 shows the sensitivity of each feature for the global warming potential metric for the ANN model. It can be observed that XLogP and critical volume have the highest sensitivity to the model output.

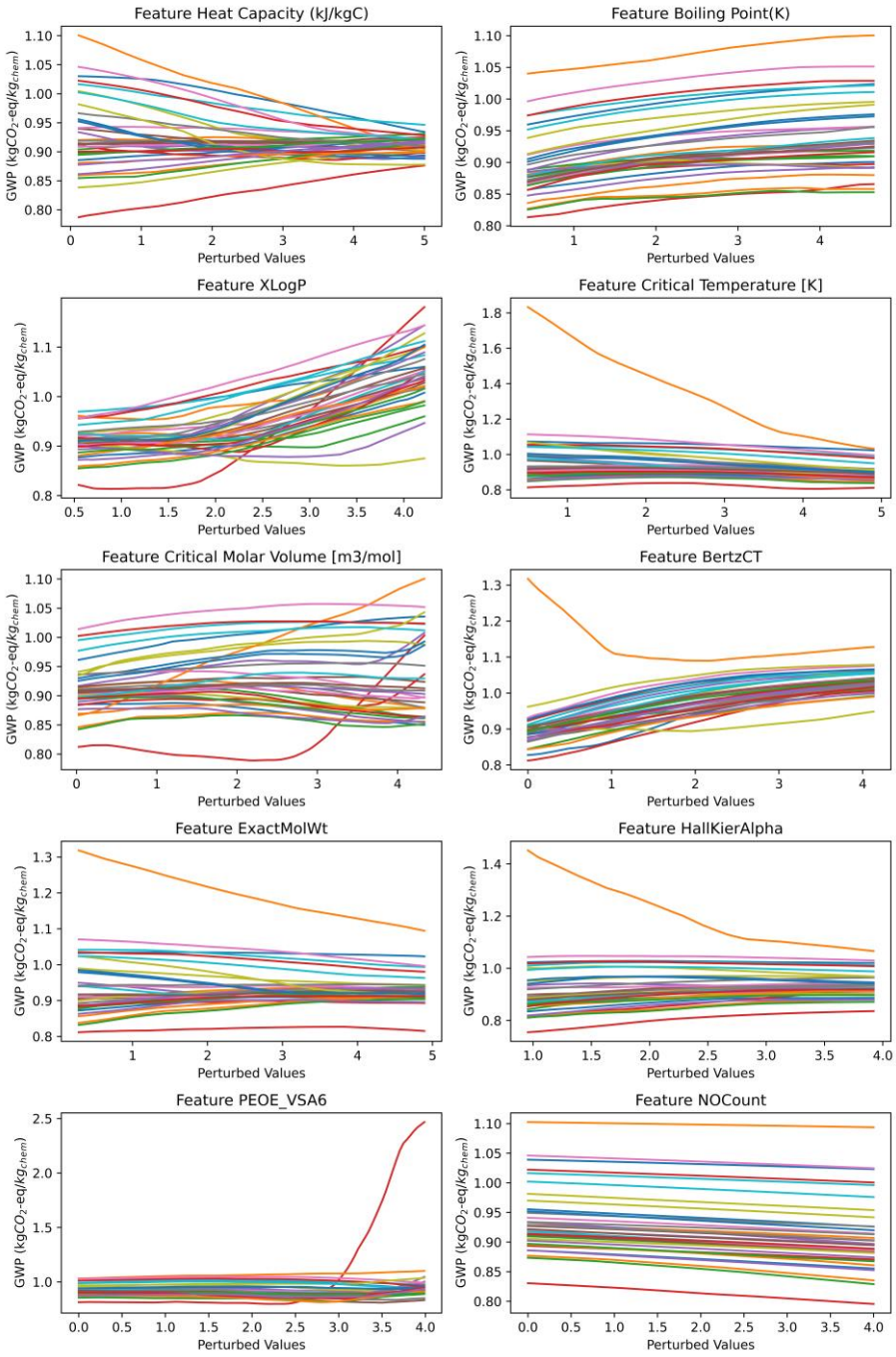


Figure 34. Sensitivity analysis for each feature for global warming potential of ANN model

Figure 35 shows the sensitivity analysis for the resource utilization impact metric. It can be observed that the SMR_VSA2 feature is the most sensitive to predicting this metric.

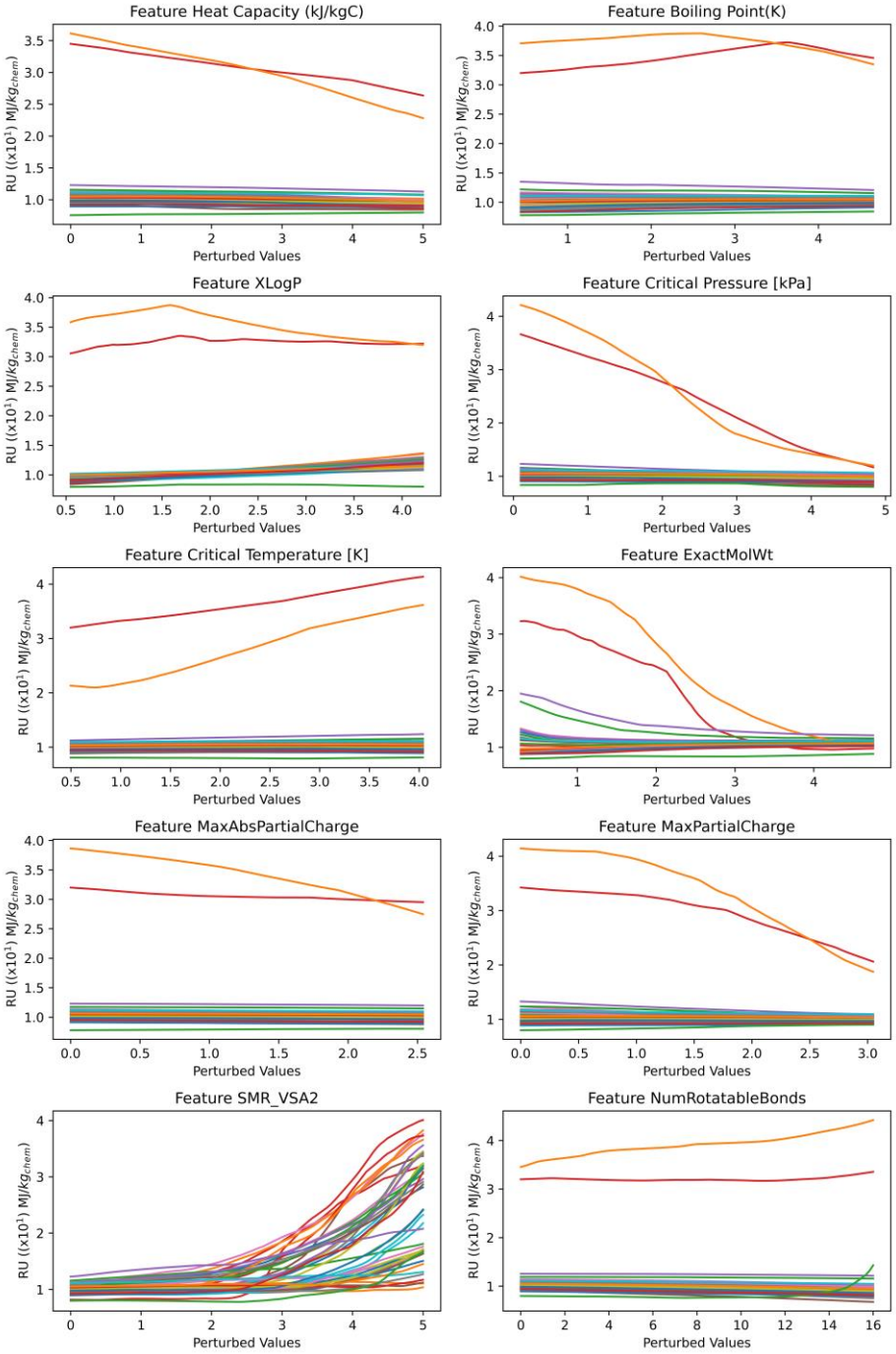


Figure 35. Sensitivity analysis for each feature for resource utilization impact of ANN model

Figure 36 shows results from the XGBoost model. One notable observation is the improvement of the EQI metric as compared to the ANN model. However, it can also be observed that some of the predictions for the validation set in the negatives. This challenge is overcome by changing the type of objective function used or defining a custom objective function. For the RUI metric, ANN performs better at predictions based on the evaluation metrics. Furthermore, there is less overfitting for the GWP metric from the XGBoost model compared to the ANN model.

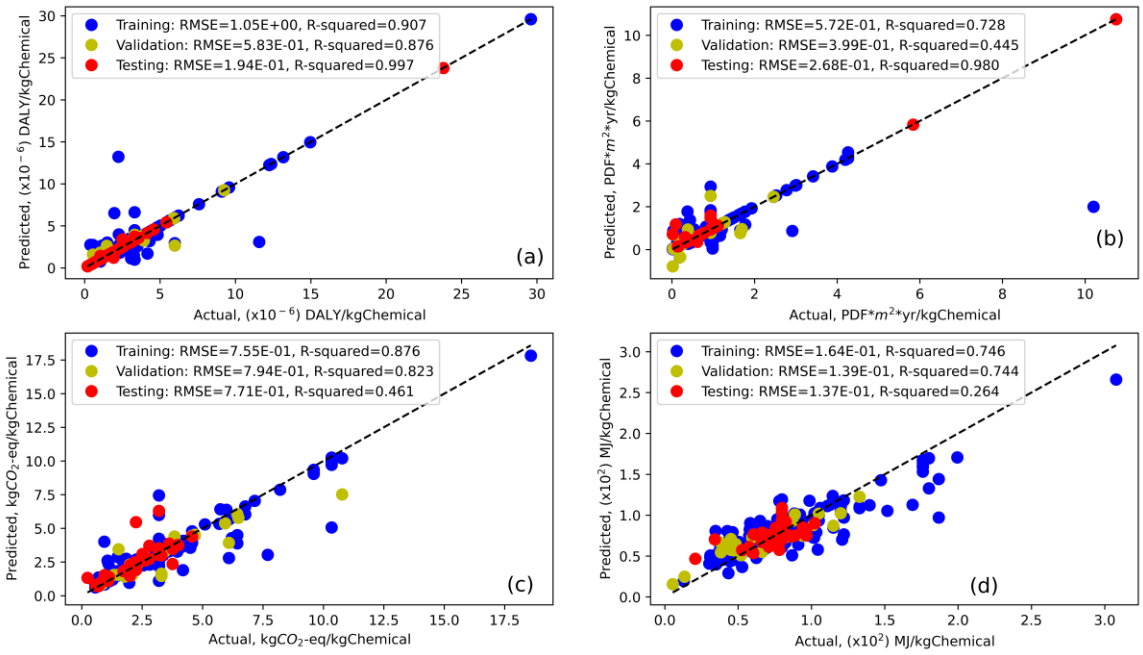


Figure 36. Parity plot for each metric from the XGBoost model. (a) is for human health impact (HHI), (b) is for ecosystem quality impact (EQI), (c) is for global warming potential (GWP), (d) is for resource utilization impact (RUI)

Figure 37 shows how the features contribute to prediction for each environmental impact metric. The heat capacity has the highest impact on the HHI metric as compared to the

second highest which is SMR_VSA7 feature. Similar observation is made for the EQI metric as critical molar volume contributes highest to the prediction. For the GWP metric, both XGBoost and ANN prefer thermodynamic properties as compared to the molecular descriptors. Conversely for the RUI metric both ML models prefer molecular descriptors as opposed to thermodynamic properties.

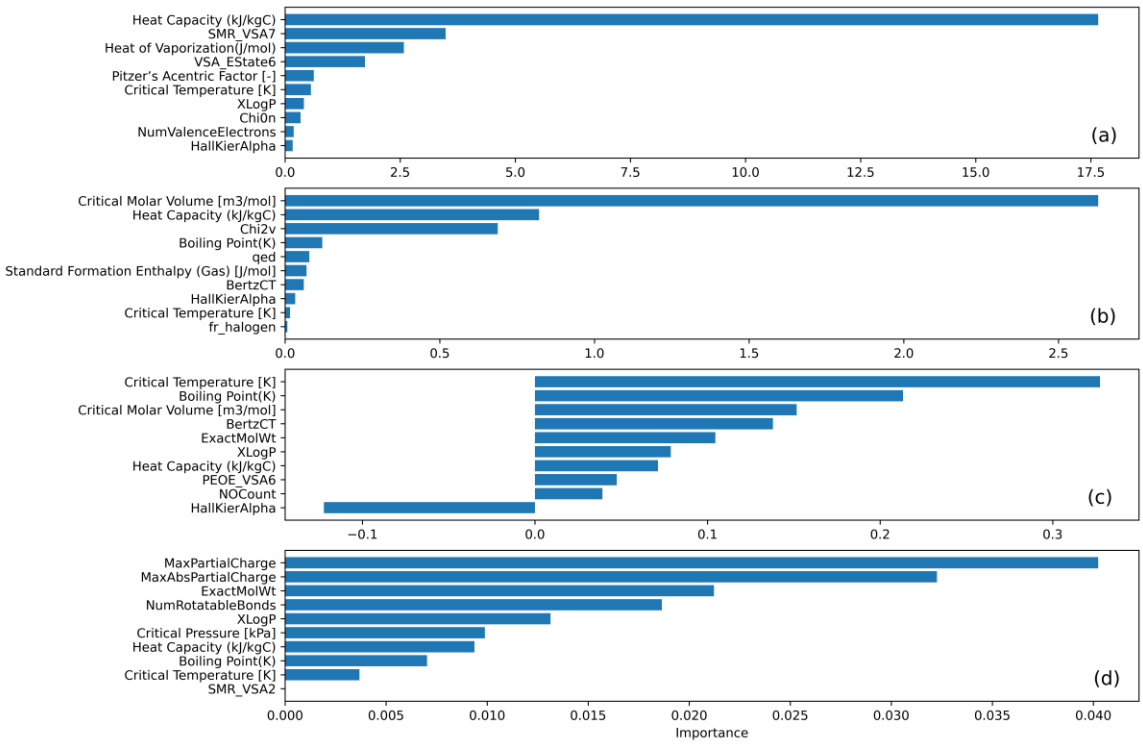


Figure 37. Feature importance for each metric from the XGBoost model. (a) is for human health impact (HHI), (b) is for ecosystem quality impact (EQI), (c) is for global warming potential (GWP), (d) is for resource utilization impact (RUI)

In terms of the sensitivity of each feature for the XGBoost models, Figure 38 shows that heat of vaporization and heat capacity are the most sensitive to the human health model prediction.

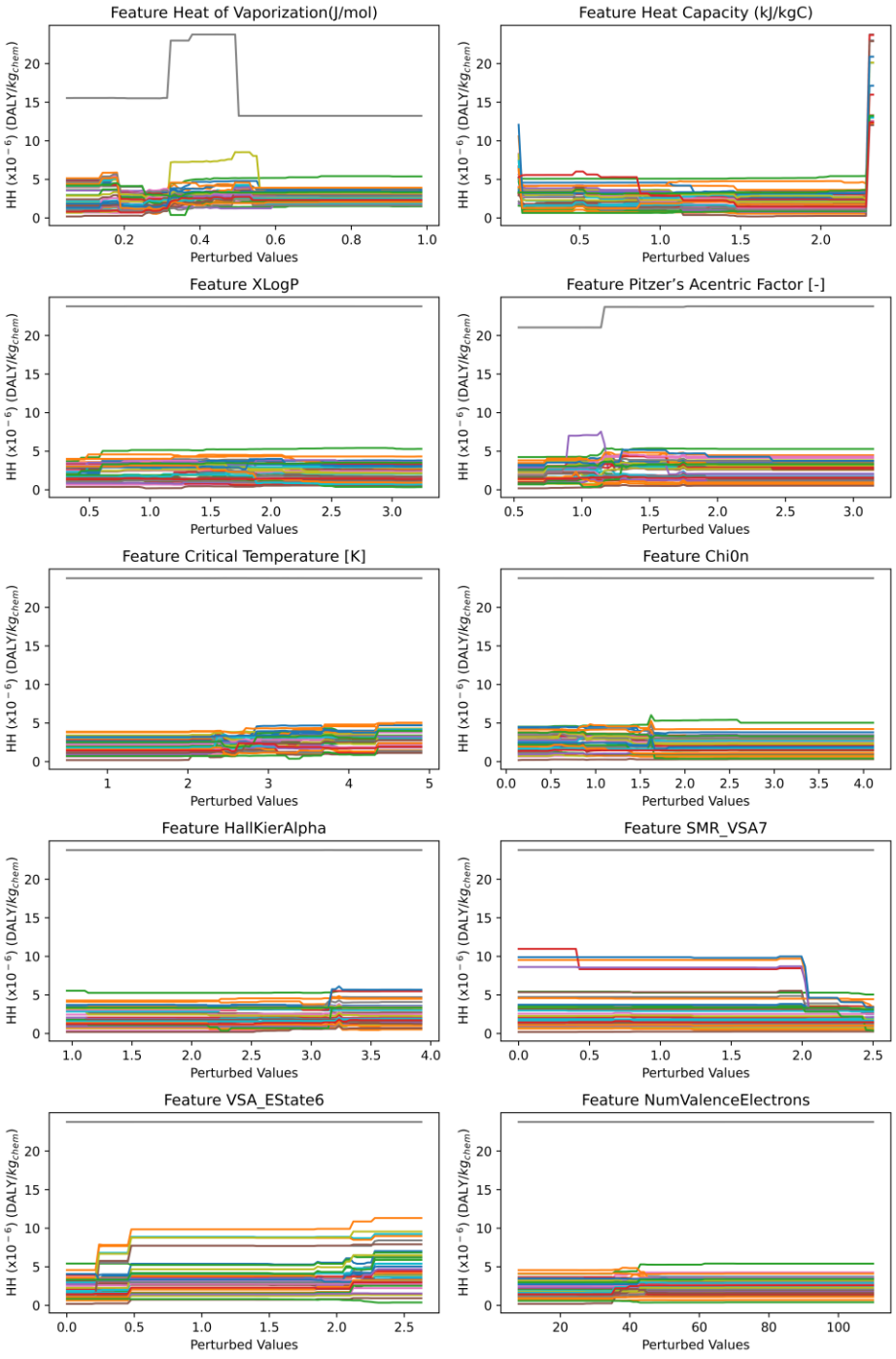


Figure 38. Sensitivity analysis for each feature for human health impact of XGBoost model

Regarding the sensitivity for the ecosystem quality impact metric, it can be observed from Figure 39 that heat capacity, molar volume, and Chi2v features have the highest sensitivity to predicting the metric.

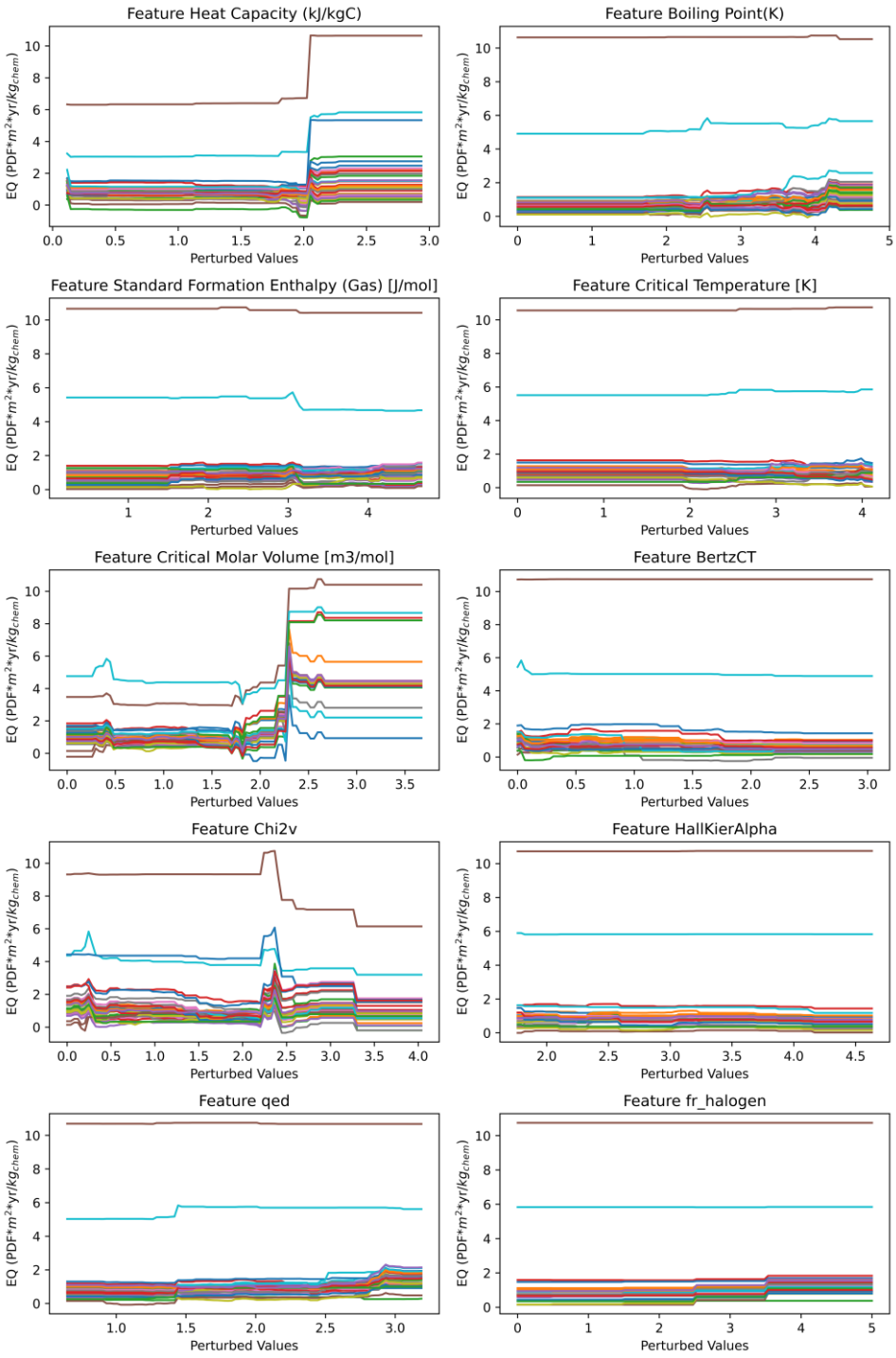


Figure 39. Sensitivity analysis for each feature for ecosystem quality impact of XGBoost model

From Figure 40, it can be observed that all the features have significant sensitivity to the output prediction for the global warming potential metric for the XGBoost model.

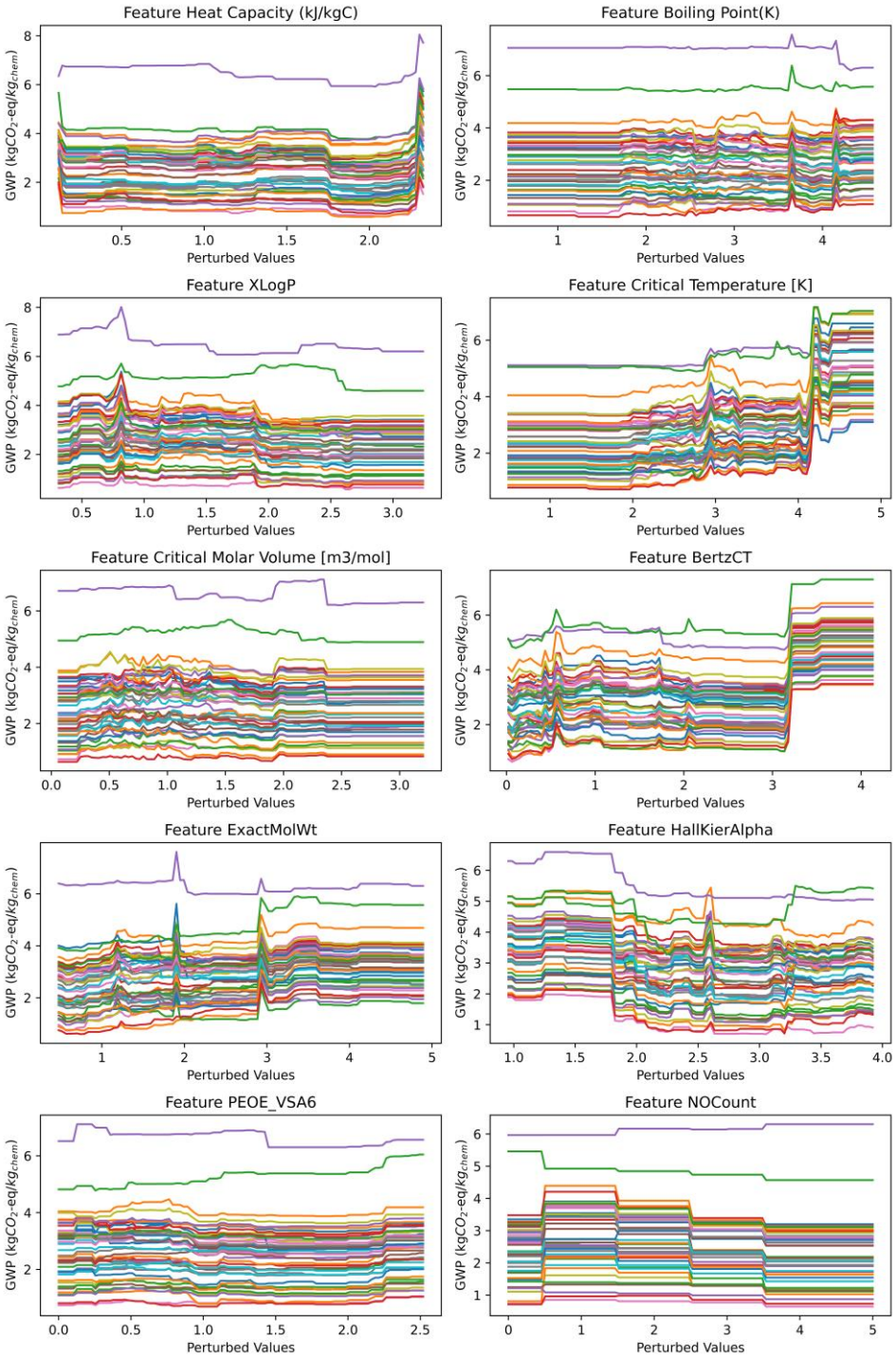


Figure 40. Sensitivity analysis for each feature for global warming potential of XGBoost model

Similar to the global warming potential, it can be observed that all the features for the resource utilization model have significant impact on the prediction of the metric as shown in Figure 41.

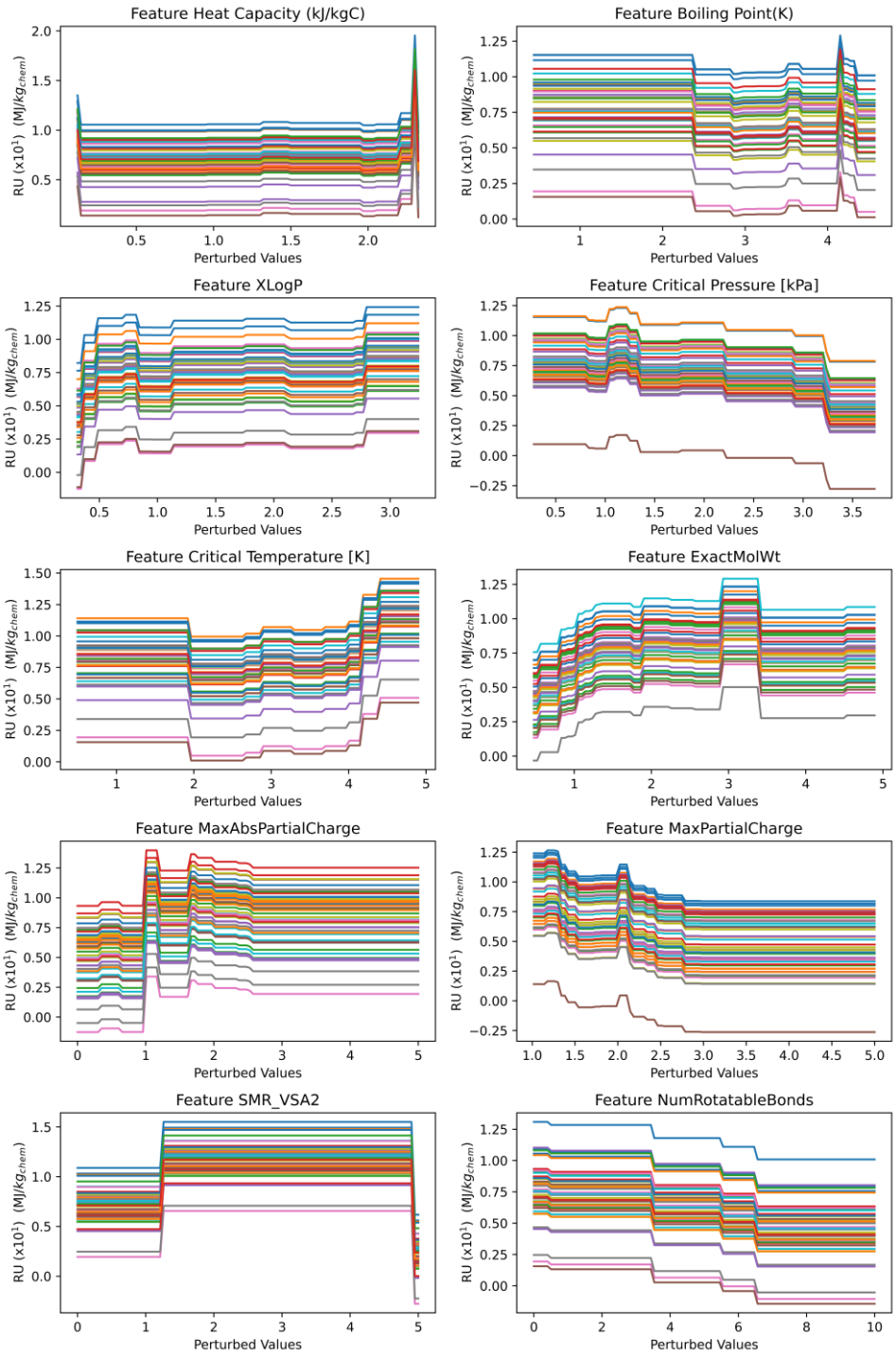


Figure 41. Sensitivity analysis for each feature for resource utilization impact of XGBoost model

Table 11 shows the confidence interval for each matric of both ANN and XGBoost model.

Table 11

Comparison of Confidence Interval for Actual Data and Model

Metric	95% Confidence Interval		
	Actual Data	XGBoost Model	ANN Model
Human Health Impact ($x10^{-6}$)	0.545 – 12.251	0.632 – 12.374	0.606 – 12.138
Ecosystem Quality Impact	0.022 – 3.122	0.022 – 3.010	0.202 – 2.341
Global Warming Potential	0.684 – 9.803	0.810 – 9.048	0.816 – 8.375
Resource Utilization Impact	34.867 – 175.910	44.934 – 153.296	36.67 – 169.60

5.7 Case Study: Cradle-to-Cradle Prediction of Life Cycle Impact Metrics for Chemicals

The ANN model is chosen to demonstrate the capability of the developed ML model using a case study and the novelty of incorporating ML models to help perform a cradle-to-cradle life cycle assessment of chemicals. Additionally, the case study shows how linked this developed model is to the EoL scenarios presented in chapters 3 and 4. In this case study, the ANN model is used to predict the environmental impacts of the chemicals used from cradle-to-gate, an ASPEN model is created for the use phase of the chemicals from a typical chemical process, and solvent recovery is implemented as the EoL option for the waste stream resulting from the process.

5.7.1 Process Description and Solution Strategy

N-Methyl-2-pyrrolidone (NMP) is a polar aprotic solvent characterized by its elevated boiling point. This solvent finds extensive applications in the chemical sector, especially in polymer production. In polymer manufacturing, the application of N-Methyl-2-pyrrolidone (NMP) raises concerns due to its non-consumptive nature in synthesis and processing, leading to its release as waste. Such usage and subsequent waste generation is widespread in the fine and specialty chemical industries. Despite the recognized health and environmental hazards associated with NMP disposal, the absence of viable and less hazardous alternatives to NMP and other dipolar aprotic solvents ensures its continued prominence in the specialty chemical domain. Therefore, it is important to recover the solvent after usage. Pastore et al [193] performed a life cycle assessment of the recovery of NMP from a waste stream. Figure 42 shows the developed flowchart for their analysis. Therefore, this case study is chosen to evaluate the developed ML model.

The impact for the production phase of the chemical is predicted by the developed ANN model. The use-phase impact is captured in the energy demand of the reactor and the washing. The impact for the EoL phase is captured by recovery of NMP.

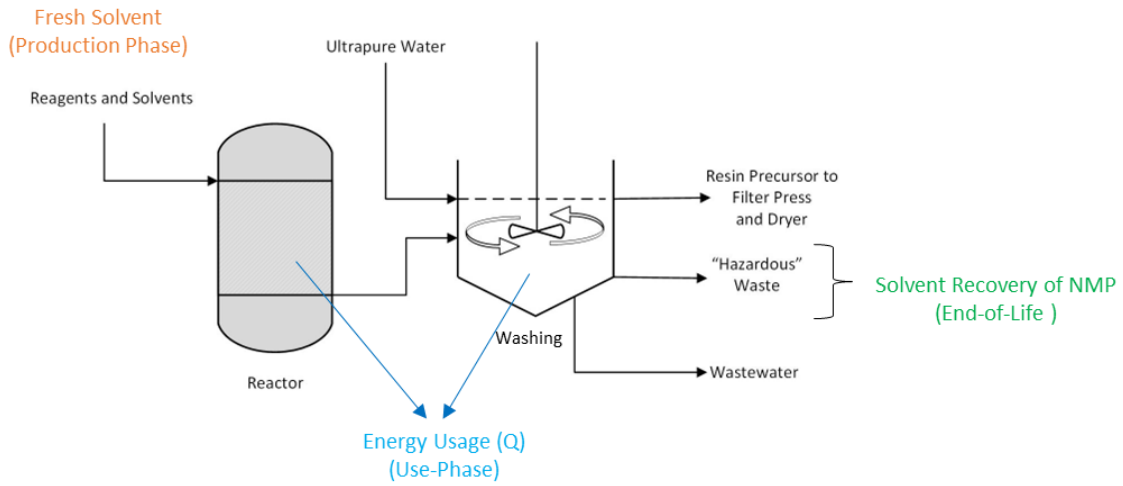


Figure 42. Flowchart for case study showing the different stages of NMP life cycle

Table 12 gives the specifications for the case study. Here ODA is 4,4'-Diaminodiphenyl ether, PMDA is Pyromellitic dianhydride. These are the aromatic dianhydride and aromatic diamine used as monomers for the synthesis of the Polyimide (PI) precursor. Trifluoroacetic acid (TFA), hydroxyethyl methacrylate (HEMA), and hydrochloric acid (HCl) are reagents and catalysts used to speed up the reaction while ethanol (EtOH) is a byproduct from the reaction. There is water as a byproduct from the reaction as well. To simplify the analysis, it is assumed that the ethanol produced, together with HEMA, TFA, and HCl are in small quantities. This assumption is based on the fact that the composition of these components in the waste stream is minute as can be seen in Table 12.

Table 12

NMP Case Study Specifications

Component	Inlet Mass (kg/yr) (Reactor)	Ultrapure water (kg/yr) (Washing)	'Hazardous waste' composition (%wt)
NMP	183416	-	17
ODA	24054.84096	-	-
PMDA	26202.28571	-	-
HEMA	5448	-	0.5
TFA	5448	-	0.5
HCl	5448	-	0.5
H_2O	-	4114148	81.5
EtOH	-	-	-
PI	-	-	-

Equations (18) – (20) gives the environmental impact of each phase of the life cycle assessment.

$$LCA_{i,production} = \sum_{j}^{n} LCA_{i,j,production} \tag{18}$$

$$LCA_{i,use-phase} = \sum_{j}^{m} LCA_{i,k,use-phase} + LCA_{i,water,use-phase} \tag{19}$$

$$LCA_{i,EoL} = \left(1 - R_{rec,NMP}\right) LCA_{i,NMP,EoL} + \sum_{j-1}^{n} LCA_{i,j,EoL} + \sum_{j}^{m} LCA_{i,k,EoL} \qquad (20)$$

Here, $LCA_{i,j,production}$ is the environment impact metric i for the production of chemical j, and n is the total number of chemicals. $LCA_{i,k,use-phase}$ is the environmental metric i for the energy demand of technology k, m is the total number of technologies in the process, $LCA_{i,water,use-phase}$ is the impact metric for the total amount of water used in the process. $R_{rec,NMP}$ is the amount of NMP recovered for reuse, $LCA_{i,NMP,EoL}$ is the environmental impact metric for NMP, $LCA_{i,j,EoL}$ is the environmental impact of the remaining chemicals not being recovered, and is the environmental impact due to the energy demand of the technologies for the solvent recovery process. The total cradle-to-cradle impact assessment per kg of NMP is given by Equation (21).

$$LCA_{i,cradle-to-cradle} = LCA_{i,production} + LCA_{i,use-phase} + LCA_{i,EoL} \qquad (21)$$

Here, i is the environmental impact indicator (HHI, EQI, GWP, RUI), $LCA_{i,production}$ is the life cycle assessment metric for the production phase of the chemical (cradle-to-gate), $LCA_{i,use-phase}$ is the life cycle assessment metric for the use-phase of the chemical (gate-to-gate) and $LCA_{i,EoL}$ is the life cycle assessment metric for the EoL phase (grave-to-cradle) for the chemical. NMP is the functional unit for the analysis hence the impact metric analysis is per kg of NMP basis.

5.7.2 Results and Discussion

Table 13 gives the environmental impacts of the production phase predicted by the ML model, while Table 14 shows the conversion of each phase of the LCA to per kg NMP bases. Due to the higher accuracy of the HHI and RUI metrics from the ANN model, is observed that predictions from these two impacts for each of the chemicals had a deviation ranging from ±2% to ±10%. For example, the SimaPro® value for RUI for NMP is 168.93, while the ML model is 166.40, signifying how well the model predicts these impacts. Similarly, the SimaPro® value for HHI for NMP is 7.58E-6, while the model prediction only deviates from this value by -6.00%. Furthermore, it is observed that for certain chemicals the models perform fairly good for the GWP predictions.

Table 13

Impact Metric Prediction from ANN Model

Component	HHI ($x10^{-6}$, DALY/kgChem)	EQI ($PDF.m^2.yr$ /kgChem)	GWP ($kgCO_2eq$ /kgChem)	RUI (x10, MJ/kgChem)
NMP	7.12	0.73	2.69	16.64
ODA	1.68	0.55	2.54	14.07
PMDA	3.89	0.86	1.97	16.45
HEMA	0.80	0.41	2.47	15.90
TFA	0.77	0.36	1.09	8.23
HCl	0.67	0.17	0.73	1.33

Table 14

Conversion of Each Phase of the LCA to Per kg NMP Basis

Component	HHI ($x10^{-6}$, DALY/kgNMP)	EQI ($PDF.m^2.yr$ /kgNMP)	GWP ($kgCO_2eq$ /kgNMP)	RUI (x10, MJ/kgNM)
Production Phase Analysis Using ANN Model Prediction				
NMP	7.12	0.73	2.69	16.64
ODA	0.22	0.07	0.33	1.85
PMDA	0.56	0.12	0.28	2.35
HEMA	0.02	0.01	0.07	0.47
TFA	0.02	0.01	0.03	0.24
HCl	0.02	0.01	0.02	0.04
Total	*7.96*	*0.95*	*3.43*	*21.59*
Use-phase				
Energy usage	9.61E-04	2.87E-04	3.01E-03	4.58E-01
Ultrapure water	1.16E+00	5.96E-01	9.79E-01	1.55E+00
Total	*1.16E+00*	*5.97E-01*	*9.82E-01*	*2.00E+00*
EoL from Solvent Recovery (assuming 90% NMP recovery)				
NMP	0.71	0.07	0.27	1.66
H_2O	0.24	0.12	0.20	0.31
Total	*0.95*	*0.19*	*0.47*	*1.98*

Then by multiplying the total values for production phase and use-phase in Table 14 by 183416 kg NMP/yr, the cradle-to-gate and gate-to-gate LCA can be estimated. For the grave-to-cradle, the EoL from solvent recovery is multiplied by 18341.6 kg NMP/yr to estimate the yearly impacts. Table 15 shows a summary of this analysis. It can be observed that the grave-to-cradle contributes the least to the overall process impact, followed by gate-to-gate. The solvent production dominates the LCA, hence the importance of this research as the developed ML model can help industries experiment with various alternatives for solvent choice during early-stage process design.

Table 15

Environmental Impact of Each LCA Method

LCA method	HHI ($x10^{-6}$, DALY/yr)	EQI ($PDF.m^2.yr/yr$)	GWP (kgCO2-eq/yr)	RUI (x10, MJ/yr)
Cradle-to-gate	1.46E+06	1.75E+05	6.29E+05	3.96E+06
Gate-to-gate	2.13E+05	1.09E+05	1.80E+05	3.67E+05
Grave-to-cradle	1.74E+04	3.56E+03	8.58E+03	3.63E+04

Figure 43 shows a comparison between solvent recovery and incineration as EoL scenarios. In the case of incineration, there is no recovery of the solvent, which makes the analysis a cradle-to-grave. However, there is the option of heat recovery with incineration, which is not considered in this analysis. Since the production and use-phases are the same

for the evaluated LCA, the figure shows the impact of just the EoL scenarios. It can be observed that solvent recovery performs better in all categories. However, the human health impact of the recovery process is close to that of incineration due to the high value of the life cycle inventory of NMP.

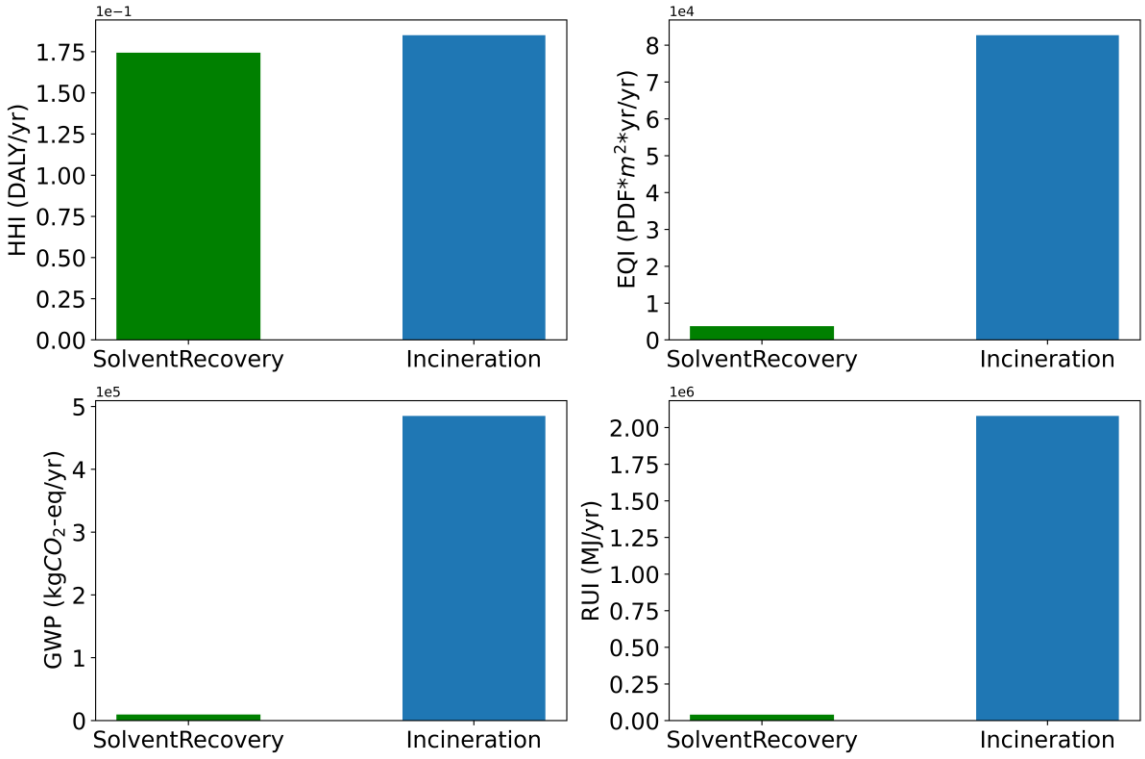


Figure 43. Consequential LCA for various EoL options

5.8 Conclusions

Developed here are two ML models for the prediction of four environmental impact metrics. Both models predict the human health and resource utilization impacts of the chemicals with higher accuracy. However, more effort is required to improve the model

for ecosystem quality predictions for ANN. Furthermore, other objective functions should be considered to make the XGBoost predictions from for the ecosystem impacts more reliable. Additionally, the GWP model needs further improvements to increase the accuracy of predictions for both models. One major highlight from this work is how different ML models have better predictions depending on the sustainability metric being investigated. Combining the developed ANN model with the solvent recovery developed in chapter 4, a novel way of performing a cradle-to-cradle LCA is introduced for processes that use high volumes of solvents through case study.

Chapter 6

Summary, Conclusions, and Future Work

6.1 Summary and Conclusions

This dissertation unfolds a new perspective for interfacing both PSE and ML in the sustainable design of chemical processes. This work presents a practical importance which aims at not only understanding the powerful advantage of incorporating sustainability assessment in the design of chemical systems, but the ability to perform an entire LCA of the process and product even at the early-stage process design.

The use of graph theory coupled with optimization for wastewater treatment networks reveals that rather than performing an exhaustive search on the synthesized maximal structure, it is far advantageous to narrow the search space to only networks that are combinatorially feasible and focus on optimization of those structures. This provides a guarantee that a global optimum is available within the feasible structure search space. Regarding the ecological impact of the wastewater treatment networks it is imperative to find ways to reduce the area needed to embed the water emissions, since this is the highest contributor to SPI in most cases. As more stringent legislations are made by governmental bodies on the allowable concentrations of contaminants within the various compartments, designing systems that can meet the anticipated future dynamic nature of effluent specifications will be crucial to improving the sustainability of industrial processes. This will help to prevent future retrofitting of the treatment plants, hence, preventing additional land area usage. Furthermore, treating wastewater for reuse presents an ecological advantage since the area needed to provide new process water for the process is prevented.

Addressing the persistent issue of solvent waste generation necessitates sustainable solutions. One promising approach is the recovery and reuse of solvents, which can yield significant benefits. However, it is essential to adopt a comprehensive perspective during the design process to ensure that solvent recovery does not inadvertently introduce other environmental concerns. This dissertation introduced a multi-objective approach that facilitates a balanced consideration of both economic and environmental implications of various treatment pathways. It advocates for the integration of solvent recovery as an integral objective during the process synthesis phase, rather than just a reactive measure to manage waste. By emphasizing proactive planning in the early stages of process design, companies have the opportunity to substantially reduce their carbon footprint, alleviate ecological pressures, and minimize energy consumption associated with their operations.

Finally, the design of sustainable industrial systems means there should be a way to quantify the environmental impacts of the proposed design right at the onset at the synthesis phase. Therefore, a comprehensive method should be implemented to aid this assessment. However, trying to perform a comprehensive assessment at early-stage where the optimal design is still unknown seems to be a daunting task. However, through ML, as demonstrated by this dissertation, a cradle-to-cradle assessment is possible. By considering solvent recovery and wastewater treatment as the EoL phase scenarios of the process and using ML to make predictions for the cradle-to-gate, a cradle-to-cradle LCA can be performed for different synthesis routes and the best option can be selected.

6.2 Future Work

This work addresses the application of mathematical modeling, optimization, process synthesis, sustainability assessment, and ML in wastewater treatment networks,

solvent recovery, and LCA metrics prediction for early-stage process synthesis. This section first discusses future work in the wastewater treatment, and then looks at possible ways of leveraging other aspects of ML for improved sustainability metrics prediction.

6.2.1 Future Work in Wastewater Treatment

The work demonstrated in the wastewater treatment aims at using P-graph framework to synthesize the networks that are structurally feasible, and then incorporating SPI, aside cost, in the optimization of each feasible structure. For the tannery case study presented in section 3.6, over 30,000 structures were optimized using an MILP approach. Thus, each of these structures have a cost and SPI calculated. One way to leverage all the data from this case study is to develop regression model using Graph Neural Network (GNN) for both cost and SPI prediction, which can be used as a guiding system for future estimation of the economics and ecological sustainability of new structures. GNN [194], [195] is a type of ML architecture where the input to the neural network is a graph. Graph Neural Networks (GNNs) represent a notable advancement in the ML domain, specifically tailored to handle the intricacies of data structured as graphs. Central to GNNs is a method that cyclically updates the features of a node by accumulating features from its neighboring nodes. This approach diligently captures the inherent relationships and localized configurations within the graph. GNNs exhibit adaptability, as evidenced by their diverse applications. They play a pivotal role in analyzing social networks, facilitating a deeper understanding of user patterns and improving recommendation algorithms. Moreover, in molecular chemistry, GNNs have been instrumental in forecasting molecular attributes and potential drug effects [196], [197]. Beyond these, GNNs are also being harnessed for refining transportation strategies and augmenting the richness of knowledge graphs. Hence,

by leveraging the power of GNNs, based on the graphs generated, future wastewater treatment network synthesis can be done rapidly and efficiently.

Furthermore, from the coffee wastewater treatment problem discussed in section 3.7, out of the 151,848 networks that were structurally feasible, only 2,779 structures were numerically feasible after optimization. Thus, GNNs can be leveraged to develop a classification model to make predictions as to whether a feasible structure will converge numerically to a solution or not. This can also be used as a guiding system to make well-informed decisions even before venturing into optimization of the structure.

Another aspect of the design of wastewater treatment networks that needs urgent attention is how resilient these systems are when faced with unforeseen circumstances. Climate change poses significant challenges to wastewater treatment networks, including increased frequency of extreme weather events such as floods and droughts. These events can severely disrupt wastewater treatment processes, leading to system overloads, damage to infrastructure, and potential environmental contamination. A resilient design incorporates adaptive measures to withstand these challenges, such as elevated structures to prevent flood damage, expanded capacity to handle increased stormwater runoff, and advanced treatment processes to ensure consistent water quality under varying conditions [198]–[200]. Furthermore, as urban populations grow, the demand for wastewater treatment systems increases. Resilient design considers not only current demands but also future growth and changing demographics. This foresight involves scalable and flexible system designs that can accommodate increased wastewater volumes and evolving treatment needs without significant overhauls.

Additionally, the resilience of wastewater treatment networks is not just an environmental or technical issue but also a public health imperative. Ineffective or interrupted wastewater treatment can lead to the spread of waterborne diseases and contamination of natural water bodies, impacting public health and ecosystem health. A resilient wastewater treatment system ensures continuous protection of public health, even in the face of disruptions, by maintaining consistent and effective treatment of wastewater.

With most wastewater treatment systems more than 75 years old and approaching their end-of-life in the Unites States, and climate change issues on the rise, it is imperative to find ways to make these systems more resilient. Thus, resilience assessment of wastewater treatment networks should also be a consideration in the synthesis and design of these systems.

6.2.2 Future Work in Machine Learning Approaches to Sustainability Assessment

The current work discussed in chapter 5 considers the feature dataset used in the model development as numerical data. Data representation is very key to developing good ML models. Representing the data as graphs and training a GNN model might also be a better way at capturing the interactions between the molecules for better predictions [201], [202]. Thus, this work can be extended to using a GNN for prediction and a comparative assessment and trade-offs can be made.

Physics-informed neural networks (PINNs) [203]–[205] have also gained tremendous attraction in the chemical engineering space. PINNs innovatively incorporate fundamental principles of physics into the structure of neural networks, establishing a harmony between data-driven approaches and specialized domain knowledge. Although data from related processes can offer valuable insights, the unique characteristics of new

processes can lead to potential inaccuracies when relying solely on data-driven predictions. However, with the utilization of PINNs, established laws of thermodynamics and chemical kinetics can be embedded directly into the predictive framework.

Currently, certain domains within machine learning, such as Large Language Models (LLMs), remain underutilized within the realm of sustainability assessment. These advanced models can be synergistically paired with conventional regression and classification methodologies to simulate a variety of design situations. By inputting distinct scenarios into the LLM, one can obtain generated textual outputs that elucidate potential sustainability outcomes. These outcomes can encompass aspects such as carbon footprint, energy efficacy, implications for human health, and impacts on ecosystem quality related to the given scenario.

Another tool that can help in acquiring data for LCIA assessment is EPI Suite tool [206] from the US EPA. EPI Suite is a collection of physical and chemical property and environmental fate estimation programs developed by the US EPA's Office of Pollution Prevention and Toxics and Syracuse Research Corporation (SRC). It serves as a user-friendly tool for estimating key environmental parameters of organic chemicals based on their molecular structure. The suite includes a variety of estimation models that predict properties such as biodegradation, soil sorption, aquatic toxicity, and air-water partitioning. One of the key features of EPI Suite is its ability to estimate the environmental fate of chemicals. It can predict how chemicals will distribute in the environment, whether they will accumulate in water, soil, or air, and their potential for long-range transport. This information is crucial for assessing the potential exposure of ecosystems and human populations to these chemicals, and hence can be integrated into the ML model.

References

[1] E. A. Aboagye, J. D. Chea, and K. M. Yenkie, "Systems level roadmap for solvent recovery and reuse in industries," *iScience*, vol. 24, no. 10, p. 103114, Oct. 2021, doi: 10.1016/j.isci.2021.103114.

[2] W. Zhang, C. Wang, Y. Li, P. Wang, Q. Wang, and D. Wang, "Seeking Sustainability: Multiobjective Evolutionary Optimization for Urban Wastewater Reuse in China," *Environ. Sci. Technol.*, vol. 48, no. 2, pp. 1094–1102, Jan. 2014, doi: 10.1021/es404082f.

[3] K. V. Gernaey, A. E. Cervera-Padrell, and J. M. Woodley, "A perspective on PSE in pharmaceutical process development and innovation," *Computers & Chemical Engineering*, vol. 42, pp. 15–29, Jul. 2012, doi: 10.1016/j.compchemeng.2012.02.022.

[4] S. Ebnesajjad, "Introduction to Plastics," in *Chemical Resistance of Commodity Thermoplastics*, Elsevier, 2016, pp. xiii–xxv. doi: 10.1016/B978-0-323-47358-3.00017-X.

[5] M. Aresta, A. Dibenedetto, and F. Dumeignil, *Biorefinery: From Biomass to Chemicals and Fuels*. Walter de Gruyter, 2012.

[6] U. T. Bornscheuer and A. T. Nielsen, "Editorial overview: Chemical biotechnology: Interdisciplinary concepts for modern biotechnological production of biochemicals and biofuels," *Current Opinion in Biotechnology*, vol. 35, pp. 133–134, 2015, doi: 10.1016/j.copbio.2015.07.002.

[7] J. H. Clark, T. J. Farmer, A. J. Hunt, and J. Sherwood, "Opportunities for Bio-Based Solvents Created as Petrochemical and Fuel Products Transition towards Renewable Resources," *Int J Mol Sci*, vol. 16, no. 8, Art. no. 8, Jul. 2015, doi: 10.3390/ijms160817101.

[8] C. Robert McElroy, A. Constantinou, L. C. Jones, L. Summerton, and J. H. Clark, "Towards a holistic approach to metrics for the 21st century pharmaceutical industry," *Green Chemistry*, vol. 17, no. 5, pp. 3111–3121, 2015, doi: 10.1039/C5GC00340G.

[9] G. Jonker and J. Harmsen, *Engineering for sustainability: a practical guide for sustainable design*, 1st ed. Amsterdam ; Boston: Elsevier, 2012.

[10] B. R. Bakshi, *Bhavik R Bakshi - Sustainable Engineering_ Principles and Practice-Cambridge University Press (2019).pdf*, 1st ed., 1 vols. Cambridge, United Kingdom: Cambridge University Press, 2019.

[11] A. Argoti, A. Orjuela, and P. C. Narváez, "Challenges and opportunities in assessing sustainability during chemical process design," *Current Opinion in Chemical Engineering*, vol. 26, pp. 96–103, Dec. 2019, doi: 10.1016/j.coche.2019.09.003.

[12] P. Barnes, "Towards sustainability: Implementing the european community's strategy for sustainable development," *Eur. Env.*, vol. 2, no. 6, pp. 5–9, Jul. 2007, doi: 10.1002/eet.3320020604.

[13] H. Cabezas and U. Diwekar, *Sustainability: Multi-Disciplinary Perspectives*, 1st ed. Bentham Science Publishers, 2012.

[14] ISO 14040, "Environmental management - Life cycle assessment - Principles and framework." ISO 14040:1997(E), 1997.

[15] Brundtland Commission (UN), “The World Commission on Environment and Development: Our Common Future,” Oxford University Press, Oxford. Accessed: Oct. 17, 2022. [Online]. Available: https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf
[16] UNEP, “Emissions Gap Report 2022: The Closing Window - Climate crisis calls for rapid transformation of societies.,” United Nations Environment Programme, Nairobi, ISBN: 978-92-807-3979-4, 2022. [Online]. Available: https://www.unep.org/emissions-gap-report-2022
[17] U. Nations, “The Paris Agreement,” United Nations. Accessed: Sep. 01, 2023. [Online]. Available: https://www.un.org/en/climatechange/paris-agreement
[18] ISO 14041, “Environmental management - Life cycle assessment - Goal and scope definition and inventory analysis.” ISO 14041:1998(E), 1998.
[19] ISO 14042, “Environmental management - Life cycle assessment - Life cycle impact assessment.” ISO 14042:2000(E), 2000.
[20] ISO 14043, “Environmental management - Life cycle assessment - Life cycle interpretation.” ISO 14043:2000(E), 2000.
[21] ISO 14044, “Environmental management - Life cycle assessment - Requirements and guidelines.” ISO 14044:2006(E), 2006.
[22] B. Corona, L. Shen, D. Reike, J. Rosales Carreón, and E. Worrell, “Towards sustainable development through the circular economy—A review and critical assessment on current circularity metrics,” *Resources, Conservation and Recycling*, vol. 151, p. 104498, Dec. 2019, doi: 10.1016/j.resconrec.2019.104498.
[23] “Circular economy | UNIDO.” Accessed: Jul. 28, 2023. [Online]. Available: https://www.unido.org/our-focus-cross-cutting-services/circular-economy
[24] O. US EPA, “Reduce, Reuse, Recycle.” Accessed: Jul. 28, 2023. [Online]. Available: https://www.epa.gov/recycle
[25] G. Bressanelli, M. Perona, and N. Saccani, “Challenges in supply chain redesign for the Circular Economy: a literature review and a multiple case study,” *International Journal of Production Research*, vol. 57, no. 23, pp. 7395–7422, Dec. 2019, doi: 10.1080/00207543.2018.1542176.
[26] P. Hopkinson, M. Zils, P. Hawkins, and S. Roper, “Managing a Complex Global Circular Economy Business Model: Opportunities and Challenges,” *California Management Review*, vol. 60, no. 3, pp. 71–94, May 2018, doi: 10.1177/0008125618764692.
[27] J. D. Hernandez-Betancur, M. Martin, and G. J. Ruiz-Mercado, “A data engineering approach for sustainable chemical end-of-life management,” *Resources, Conservation and Recycling*, vol. 178, p. 106040, Mar. 2022, doi: 10.1016/j.resconrec.2021.106040.
[28] M. Davis, *Water and Wastewater Engineering*. McGraw-Hill Education, 2010.
[29] F. P. Shariati, M. R. Mehrnia, B. M. Salmasi, M. Heran, C. Wisniewski, and M. H. Sarrafzadeh, “Membrane bioreactor for treatment of pharmaceutical wastewater containing acetaminophen,” *Desalination*, vol. 250, no. 2, pp. 798–800, Jan. 2010, doi: 10.1016/j.desal.2008.11.044.

[30] E. A. Aboagye *et al.*, “Systematic Design, Optimization, and Sustainability Assessment for Generation of Efficient Wastewater Treatment Networks,” *Water*, vol. 13, no. 9, p. 1326, May 2021, doi: 10.3390/w13091326.
[31] K. M. Yenkie, “Integrating the three E’s in wastewater treatment: efficient design, economic viability, and environmental sustainability,” *Current Opinion in Chemical Engineering*, vol. 26, pp. 131–138, Dec. 2019, doi: 10.1016/j.coche.2019.09.002.
[32] H. Bozkurt, A. Quaglia, K. V. Gernaey, and G. Sin, “A mathematical programming framework for early stage design of wastewater treatment plants,” *Environmental Modelling & Software*, vol. 64, pp. 164–176, Feb. 2015, doi: 10.1016/j.envsoft.2014.11.023.
[33] G. Crini and E. Lichtfouse, “Advantages and disadvantages of techniques used for wastewater treatment,” *Environ Chem Lett*, vol. 17, no. 1, pp. 145–155, Mar. 2019, doi: 10.1007/s10311-018-0785-9.
[34] F. Friedler, Á. Orosz, and J. Pimentel Losada, *P-graphs for Process Systems Engineering: Mathematical Models and Algorithms*. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-92216-0.
[35] B. Lu, S. Huang, and I. E. Grossmann, “Optimal Synthesis and Operation of Wastewater Treatment Process with Dynamic Influent,” *Ind. Eng. Chem. Res.*, vol. 56, no. 30, pp. 8663–8676, Aug. 2017, doi: 10.1021/acs.iecr.7b01805.
[36] C. Jimenez-Gonzalez, C. S. Ponder, Q. B. Broxterman, and J. B. Manley, “Using the Right Green Yardstick: Why Process Mass Intensity Is Used in the Pharmaceutical Industry To Drive More Sustainable Processes,” *Org. Process Res. Dev.*, vol. 15, no. 4, pp. 912–917, Jul. 2011, doi: 10.1021/op200097d.
[37] S. Sharma and S. S. Kanwar, “Organic Solvent Tolerant Lipases and Applications,” *The Scientific World Journal*, vol. 2014, pp. 1–15, 2014, doi: 10.1155/2014/625258.
[38] K. Zhang, “Organic solvent pretreatment of lignocellulosic biomass for biofuels and biochemicals: A review,” *Bioresource Technology*, p. 13, 2016.
[39] D. J. C. Constable, C. Jimenez-Gonzalez, and R. K. Henderson, “Perspective on Solvent Use in the Pharmaceutical Industry,” *Org. Process Res. Dev.*, vol. 11, no. 1, pp. 133–137, Jan. 2007, doi: 10.1021/op060170h.
[40] Dunn Peter J., Wells Andrew S., and Williams Michael T., “Future Trends for Green Chemistry in the Pharmaceutical Industry,” *Green Chemistry in the Pharmaceutical Industry*, Mar. 2010, doi: 10.1002/9783527629688.ch16.
[41] R. A. Sheldon, “The E factor 25 years on: the rise of green chemistry and sustainability,” *Green Chem.*, vol. 19, no. 1, pp. 18–43, Jan. 2017, doi: 10.1039/C6GC02157C.
[42] A. Das, A. K. Gupta, and T. N. Mazumder, “A comprehensive risk assessment framework for offsite transportation of inflammable hazardous waste,” *Journal of Hazardous Materials*, vol. 227–228, pp. 88–96, Aug. 2012, doi: 10.1016/j.jhazmat.2012.05.014.
[43] M. Goedkoop and R. Spriensma, “The Eco-indicator 99 A damage oriented method for Life Cycle Impact Assessment,” p. 87, 2001.
[44] O. Jolliet *et al.*, “IMPACT 2002+: A new life cycle impact assessment methodology,” *Int J LCA*, vol. 8, no. 6, pp. 324–330, Nov. 2003, doi: 10.1007/BF02978505.

[45] J. N. Sangshetti, M. Deshpande, Z. Zaheer, D. B. Shinde, and R. Arote, “Quality by design approach: Regulatory need,” *Arabian Journal of Chemistry*, vol. 10, pp. S3412–S3425, May 2017, doi: 10.1016/j.arabjc.2014.01.025.
[46] Q. Chen and I. E. Grossmann, “Recent Developments and Challenges in Optimization-Based Process Synthesis,” *Annu. Rev. Chem. Biomol. Eng.*, vol. 8, no. 1, pp. 249–283, Jun. 2017, doi: 10.1146/annurev-chembioeng-080615-033546.
[47] U. Diwekar, *Introduction to Applied Optimization*. Springer Science & Business Media, 2013.
[48] S. Papadokonstantakis, P. Karka, Y. Kikuchi, and A. Kokossis, “Challenges for Model-Based Life Cycle Inventories and Impact Assessment in Early to Basic Process Design Stages,” in *Sustainability in the Design, Synthesis and Analysis of Chemical Engineering Processes*, Elsevier, 2016, pp. 295–326. doi: 10.1016/B978-0-12-802032-6.00013-X.
[49] P. Karka, S. Papadokonstantakis, and A. Kokossis, “Predictive LCA - a systems approach to integrate LCA decisions ahead of design,” in *Computer Aided Chemical Engineering*, vol. 46, Elsevier, 2019, pp. 97–102. doi: 10.1016/B978-0-12-818634-3.50017-5.
[50] S. D. Barnicki and J. J. Siirola, “Process synthesis prospective,” *Computers & Chemical Engineering*, vol. 28, no. 4, pp. 441–446, Apr. 2004, doi: 10.1016/j.compchemeng.2003.09.030.
[51] S. Cremaschi, “A perspective on process synthesis: Challenges and prospects,” *Computers & Chemical Engineering*, vol. 81, pp. 130–137, Oct. 2015, doi: 10.1016/j.compchemeng.2015.05.007.
[52] G. Tchobanoglous, F. L. Burton, and H. D. Stensel, *Wastewater Engineering: Treatment and Reuse*, 4th edition. Boston: McGraw Hill Higher Education, 2002.
[53] B. L. Blaney, “Treatment Technologies for Hazardous Wastes: Part II Alternative Techniques for Managing Solvent Wastes,” *Journal of the Air Pollution Control Association*, vol. 36, no. 3, pp. 275–285, Mar. 1986, doi: 10.1080/00022470.1986.10466070.
[54] T. Wuest, *Identifying Product and Process State Drivers in Manufacturing Systems Using Supervised Machine Learning*. in Springer Theses. Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-17611-6.
[55] G. M. Abdella, M. Kucukvar, N. C. Onat, H. M. Al-Yafay, and M. E. Bulak, “Sustainability assessment and modeling based on supervised machine learning techniques: The case for food consumption,” *Journal of Cleaner Production*, vol. 251, p. 119661, Apr. 2020, doi: 10.1016/j.jclepro.2019.119661.
[56] A. Kumar and J. Flores-Cerrillo, “Machine Learning in Python for Process Systems Engineering,” p. 352.
[57] H. Cabezas, A. Argoti, F. Friedler, P. Mizsey, and J. Pimentel, “Design and engineering of sustainable process systems and supply chains by the P-graph framework,” *Environ. Prog. Sustainable Energy*, vol. 37, no. 2, pp. 624–636, Mar. 2018, doi: 10.1002/ep.12887.
[58] U. Diwekar, “Green process design, industrial ecology, and sustainability: A systems analysis perspective,” *Resources, Conservation and Recycling*, vol. 44, no. 3, pp. 215–235, Jun. 2005, doi: 10.1016/j.resconrec.2005.01.007.

[59] J. Gong and F. You, “Sustainable design and synthesis of energy systems,” *Current Opinion in Chemical Engineering*, vol. 10, pp. 77–86, Nov. 2015, doi: 10.1016/j.coche.2015.09.001.
[60] H. Karlsson, S. Ahlgren, M. Sandgren, V. Passoth, O. Wallberg, and P.-A. Hansson, “A systems analysis of biodiesel production from wheat straw using oleaginous yeast: process design, mass and energy balances,” *Biotechnology for Biofuels*, vol. 9, p. 229, 2016, doi: 10.1186/s13068-016-0640-9.
[61] A. Szlama, I. Heckl, and H. Cabezas, “Optimal design of renewable energy systems with flexible inputs and outputs using the P-graph framework,” *AIChE Journal*, vol. 62, no. 4, pp. 1143–1153, 2016, doi: 10.1002/aic.15137.
[62] M. Narodoslawsky and C. Krotscheck, “The sustainable process index (SPI): evaluating processes according to environmental compatibility,” *Journal of Hazardous Materials*, vol. 41, no. 2, pp. 383–397, May 1995, doi: 10.1016/0304-3894(94)00114-V.
[63] M. Narodoslawsky, “Chapter 3 - Sustainable process index,” in *Assessing and Measuring Environmental Impact and Sustainability*, J. J. Klemeš, Ed., Oxford: Butterworth-Heinemann, 2015, pp. 73–86. doi: 10.1016/B978-0-12-799968-5.00003-8.
[64] M. Narodoslawsky and Ch. Krotscheck, “What can we learn from ecological valuation of processes with the sustainable process index (SPI) — the case study of energy production systems,” *Journal of Cleaner Production*, vol. 12, no. 2, pp. 111–115, Mar. 2004, doi: 10.1016/S0959-6526(02)00184-1.
[65] H. T. Odum, “Self-Organization, Transformity, and Information,” *Science*, vol. 242, no. 4882, pp. 1132–1139, Nov. 1988, doi: 10.1126/science.242.4882.1132.
[66] H. T. Odum and E. C. Odum, “Modeling for All Scales,” p. 479, 2000.
[67] S. Ulgiati, H. T. Odum, and S. Bastianoni, “Emergy use, environmental loading and sustainability an emergy analysis of Italy,” *Ecological Modelling*, vol. 73, no. 3–4, pp. 215–268, Jun. 1994, doi: 10.1016/0304-3800(94)90064-7.
[68] S. E. Jørgensen, S. N. Nielsen, and H. Mejer, “Emergy, environ, exergy and ecological modelling,” *Ecological Modelling*, vol. 77, no. 2–3, pp. 99–109, Feb. 1995, doi: 10.1016/0304-3800(93)E0080-M.
[69] M. T. Brown and R. A. Herendeen, “Embodied energy analysis and EMERGY analysis: a comparative view,” *Ecological Economics*, vol. 19, no. 3, pp. 219–235, Dec. 1996, doi: 10.1016/S0921-8009(96)00046-8.
[70] J. Ren, S. Tan, L. Yang, M. E. Goodsite, C. Pang, and L. Dong, “Optimization of emergy sustainability index for biodiesel supply network design,” *Energy Conversion and Management*, vol. 92, pp. 312–321, Mar. 2015, doi: 10.1016/j.enconman.2014.12.066.
[71] M. T. Brown, M. Raugei, and S. Ulgiati, “On boundaries and ‘investments’ in Emergy Synthesis and LCA: A case study on thermal vs. photovoltaic electricity,” *Ecological Indicators*, vol. 15, no. 1, pp. 227–235, Apr. 2012, doi: 10.1016/j.ecolind.2011.09.021.

[72] M. Kiani-Moghaddam, M. Shivaie, and P. D. Weinsier, “Introduction to Multi-objective Optimization and Decision-Making Analysis,” in *Modern Music-Inspired Optimization Algorithms for Electric Power Systems*, in Power Systems. , Cham: Springer International Publishing, 2019, pp. 21–45. doi: 10.1007/978-3-030-12044-3_2.
[73] D. A. L. Silva, A. O. Nunes, V. A. da S. Moris, and C. M. Piekarski, “How important is the LCA software tool you choose? Comparative results from GaBi, openLCA, SimaPro and Umberto,” presented at the International Life Cycle Conference in Latin America, Antioquia, Colombia, 2017.
[74] A. Ciroth, C. D. Noi, M. Lohse, and M. Srocka, “openLCA Comprehensive user manual.” GreenDelta, Berlin, Germany, 2020. [Online]. Available: https://www.openlca.org/wp-content/uploads/2020/02/openLCA_1.10_User-Manual.pdf
[75] M. Goedkoop, M. Oele, J. Leijting, T. Ponsioen, and E. Meijer, “Introduction to LCA with SimaPro.” PRe Sustainability, 2016.
[76] P. Karka, S. Papadokonstantakis, and A. Kokossis, “Environmental impact assessment of biomass process chains at early design stages using decision trees,” *Int J Life Cycle Assess*, vol. 24, no. 9, pp. 1675–1700, Sep. 2019, doi: 10.1007/s11367-019-01591-0.
[77] C. Huyen, *Designing machine learning systems: an iterative process for production-ready applications*, First edition. Sebastopol, CA: O’Reilly Media, Inc, 2022.
[78] J. Wei *et al.*, “Machine learning in materials science,” *InfoMat*, vol. 1, no. 3, pp. 338–358, 2019, doi: 10.1002/inf2.12028.
[79] *Machine Learning in Chemistry*. 2020. doi: 10.1039/9781839160233.
[80] C. S. Khor, B. Chachuat, and N. Shah, “A superstructure optimization approach for water network synthesis with membrane separation-based regenerators,” *Computers & Chemical Engineering*, vol. 42, pp. 48–63, Jul. 2012, doi: 10.1016/j.compchemeng.2012.02.020.
[81] C. S. Khor, D. C. Y. Foo, M. M. El-Halwagi, R. R. Tan, and N. Shah, “A Superstructure Optimization Approach for Membrane Separation-Based Water Regeneration Network Synthesis with Detailed Nonlinear Mechanistic Reverse Osmosis Model,” *Ind. Eng. Chem. Res.*, vol. 50, no. 23, pp. 13444–13456, Dec. 2011, doi: 10.1021/ie200665g.
[82] R. B. P. Marcelino, M. M. D. Leão, R. M. Lago, and C. C. Amorim, “Multistage ozone and biological treatment system for real wastewater containing antibiotics,” *Journal of Environmental Management*, vol. 195, pp. 110–116, Jun. 2017, doi: 10.1016/j.jenvman.2016.04.041.
[83] P. Arroyo and M. Molinos-Senante, “Selecting appropriate wastewater treatment technologies using a choosing-by-advantages approach,” *Science of The Total Environment*, vol. 625, pp. 819–827, Jun. 2018, doi: 10.1016/j.scitotenv.2017.12.331.
[84] Y. H. Yang, H. H. Lou, and Y. L. Huang, “Synthesis of an optimal wastewater reuse network,” *Waste Management*, vol. 20, no. 4, pp. 311–319, Jul. 2000, doi: 10.1016/S0956-053X(99)00298-6.
[85] “Mitigation of eutrophication caused by wastewater discharge: A simulation-based approach | Ambio.” Accessed: Nov. 27, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s13280-020-01346-4

[86] “Eutrophication and Wastewater Purification - Gleisberg - 1976 - Angewandte Chemie International Edition in English - Wiley Online Library.” Accessed: Nov. 27, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.197603541
[87] P. S. Goh and A. F. Ismail, “A review on inorganic membranes for desalination and wastewater treatment,” *Desalination*, vol. 434, pp. 60–80, May 2018, doi: 10.1016/j.desal.2017.07.023.
[88] K. Wang, A. A. Abdalla, M. A. Khaleel, N. Hilal, and M. K. Khraisheh, “Mechanical properties of water desalination and wastewater treatment membranes,” *Desalination*, vol. 401, pp. 190–205, Jan. 2017, doi: 10.1016/j.desal.2016.06.032.
[89] E. T. Sayed, N. Shehata, M. A. Abdelkareem, and M. A. Atieh, “Recent progress in environmentally friendly bio-electrochemical devices for simultaneous water desalination and wastewater treatment,” *Science of The Total Environment*, vol. 748, p. 141046, Dec. 2020, doi: 10.1016/j.scitotenv.2020.141046.
[90] B. Galan and I. E. Grossmann, “Optimal Design of Distributed Wastewater Treatment Networks,” *Ind. Eng. Chem. Res.*, vol. 37, no. 10, pp. 4036–4048, Oct. 1998, doi: 10.1021/ie980133h.
[91] S. Rigopoulos and P. Linke, “Systematic development of optimal activated sludge process designs,” *Computers & Chemical Engineering*, vol. 26, no. 4–5, pp. 585–597, May 2002, doi: 10.1016/S0098-1354(01)00782-7.
[92] N. Vidal, R. Bañares-Alcántara, I. Rodríguez-Roda, and M. Poch, “Design of Wastewater Treatment Plants Using a Conceptual Design Methodology,” *Ind. Eng. Chem. Res.*, vol. 41, no. 20, pp. 4993–5005, Oct. 2002, doi: 10.1021/ie010652b.
[93] G. Zeng, R. Jiang, G. Huang, M. Xu, and J. Li, “Optimization of wastewater treatment alternative selection by hierarchy grey relational analysis,” *Journal of Environmental Management*, vol. 82, no. 2, pp. 250–259, Jan. 2007, doi: 10.1016/j.jenvman.2005.12.024.
[94] V. Ku-Pineda and R. R. Tan, “Environmental performance optimization using process water integration and Sustainable Process Index,” *Journal of Cleaner Production*, vol. 14, no. 18, pp. 1586–1592, Jan. 2006, doi: 10.1016/j.jclepro.2005.04.018.
[95] C. S. Lee, J. Robinson, and M. F. Chong, “A review on application of flocculants in wastewater treatment,” *Process Safety and Environmental Protection*, vol. 92, no. 6, pp. 489–508, Nov. 2014, doi: 10.1016/j.psep.2014.04.010.
[96] H. Ødegaard, “Optimization of flocculation/flotation in chemical wastewater treatment,” *Water Science and Technology*, vol. 31, no. 3–4, pp. 73–82, 1995, doi: 10.1016/0273-1223(95)99878-8.
[97] N. K. Shammas, “Water Engineering,” p. 829.
[98] R. Singh, “Water and Membrane Treatment,” in *Membrane Technology and Engineering for Water Purification*, Elsevier, 2015, pp. 81–178. doi: 10.1016/B978-0-444-63362-0.00002-1.
[99] A. Adin, “Slow granular filtration for water reuse,” *Water Science and Technology: Water Supply*, vol. 3, no. 4, pp. 123–130, Aug. 2003, doi: 10.2166/ws.2003.0053.
[100] L. Ho *et al.*, “Assessing granular media filtration for the removal of chemical contaminants from wastewater,” *Water Research*, vol. 45, no. 11, pp. 3461–3472, May 2011, doi: 10.1016/j.watres.2011.04.005.

[101] L. Boutilier, R. Jamieson, R. Gordon, C. Lake, and W. Hart, “Adsorption, sedimentation, and inactivation of E. coli within wastewater treatment wetlands,” *Water Research*, vol. 43, no. 17, pp. 4370–4380, Sep. 2009, doi: 10.1016/j.watres.2009.06.039.
[102] A. Acher, E. Fischer, R. Turnheim, and Y. Manor, “Ecologically friendly wastewater disinfection techniques,” *Water Research*, vol. 31, no. 6, pp. 1398–1404, Jun. 1997, doi: 10.1016/S0043-1354(96)00000-0.
[103] M. De Sanctis, G. Del Moro, C. Levantesi, M. L. Luprano, and C. Di Iaconi, “Integration of an innovative biological treatment with physical or chemical disinfection for wastewater reuse,” *Science of The Total Environment*, vol. 543, pp. 206–213, Feb. 2016, doi: 10.1016/j.scitotenv.2015.11.006.
[104] T. Blume, “Improved wastewater disinfection by ultrasonic pre-treatment*1,” *Ultrasonics Sonochemistry*, vol. 11, no. 5, pp. 333–336, Jul. 2004, doi: 10.1016/S1350-4177(03)00156-1.
[105] M. C. Dodd, “Potential impacts of disinfection processes on elimination and deactivation of antibiotic resistance genes during water and wastewater treatment,” *J Environ Monit*, vol. 14, no. 7, pp. 1754–1771, Jul. 2012, doi: 10.1039/c2em00006g.
[106] M. Kitis, “Disinfection of wastewater with peracetic acid: a review,” *Environ Int*, vol. 30, no. 1, pp. 47–55, Mar. 2004, doi: 10.1016/S0160-4120(03)00147-8.
[107] F. Hassard, J. Biddle, E. Cartmell, B. Jefferson, S. Tyrrel, and T. Stephenson, “Rotating biological contactors for wastewater treatment – A review,” *Process Safety and Environmental Protection*, vol. 94, pp. 285–306, Mar. 2015, doi: 10.1016/j.psep.2014.07.003.
[108] B. G. Lipták and D. H. F. Liu, *Wastewater treatment*. Boca Raton, Fla. ; London : Lewis Pub, 2000. Accessed: Oct. 16, 2017. [Online]. Available: http://trove.nla.gov.au/version/46381639
[109] M. Pankhania, T. Stephenson, and M. J. Semmens, “Hollow fibre bioreactor for wastewater treatment using bubbleless membrane aeration,” *Water Research*, vol. 28, no. 10, pp. 2233–2236, Oct. 1994, doi: 10.1016/0043-1354(94)90037-X.
[110] J. Radjenović, M. Matošić, I. Mijatović, M. Petrović, and D. Barceló, “Membrane Bioreactor (MBR) as an Advanced Wastewater Treatment Technology,” in *Emerging Contaminants from Industrial and Municipal Waste: Removal Technologies*, D. Barceló and M. Petrovic, Eds., in The Handbook of Environmental Chemistry. , Berlin, Heidelberg: Springer, 2008, pp. 37–101. doi: 10.1007/698_5_093.
[111] T. Melin *et al.*, “Membrane bioreactor technology for wastewater treatment and reuse,” *Desalination*, vol. 187, no. 1–3, pp. 271–282, Feb. 2006, doi: 10.1016/j.desal.2005.04.086.
[112] A. K. Kivaisi, “The potential for constructed wetlands for wastewater treatment and reuse in developing countries: a review,” *Ecological Engineering*, vol. 16, no. 4, pp. 545–560, Feb. 2001, doi: 10.1016/S0925-8574(00)00113-0.
[113] G. A. Moshiri, *Constructed Wetlands for Water Quality Improvement*. Boca Raton: CRC Press, 2020. doi: 10.1201/9781003069997.
[114] M. M. Saggaï, A. Ainouche, M. Nelson, F. Cattin, and A. El Amrani, “Long-term investigation of constructed wetland wastewater treatment and reuse: Selection of adapted plant species for metaremediation,” *Journal of Environmental Management*, vol. 201, pp. 120–128, Oct. 2017, doi: 10.1016/j.jenvman.2017.06.040.

[115] L. Zhang *et al.*, "Effects of constructed wetland design on ibuprofen removal – A mesocosm scale study," *Science of The Total Environment*, vol. 609, pp. 38–45, Dec. 2017, doi: 10.1016/j.scitotenv.2017.07.130.
[116] J. T. A. Verhoeven and A. F. M. Meuleman, "Wetlands for wastewater treatment: Opportunities and limitations," *Ecological Engineering*, vol. 12, no. 1–2, pp. 5–12, Jan. 1999, doi: 10.1016/S0925-8574(98)00050-0.
[117] Z. Z. Ismail and A. A. Habeeb, "Experimental and modeling study of simultaneous power generation and pharmaceutical wastewater treatment in microbial fuel cell based on mobilized biofilm bearers," *Renewable Energy*, vol. 101, pp. 1256–1265, Feb. 2017, doi: 10.1016/j.renene.2016.10.008.
[118] V. M. Ortiz-Martínez, M. J. Salar-García, A. P. de los Ríos, F. J. Hernández-Fernández, J. A. Egea, and L. J. Lozano, "Developments in microbial fuel cell modeling," *Chemical Engineering Journal*, vol. 271, pp. 50–60, Jul. 2015, doi: 10.1016/j.cej.2015.02.076.
[119] D. Bose, M. Gopinath, and P. Vijay, "Sustainable power generation from wastewater sources using Microbial Fuel Cell: Sustainable Power generation from Wastewater sources using Microbial fuel cell," *Biofuels, Bioprod. Bioref.*, vol. 12, no. 4, pp. 559–576, Jul. 2018, doi: 10.1002/bbb.1892.
[120] P. Rajasulochana and V. Preethy, "Comparison on efficiency of various techniques in treatment of waste and sewage water – A comprehensive review," *Resource-Efficient Technologies*, vol. 2, no. 4, pp. 175–184, Dec. 2016, doi: 10.1016/j.reffit.2016.09.004.
[121] G. Durai and M. Rajasimman, "Biological Treatment of Tannery Wastewater - A Review," *J. of Environmental Science and Technology*, vol. 4, no. 1, pp. 1–17, Dec. 2010, doi: 10.3923/jest.2011.1.17.
[122] F. Ju and T. Zhang, "Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant," *ISME J*, vol. 9, no. 3, pp. 683–695, Mar. 2015, doi: 10.1038/ismej.2014.162.
[123] X. Shi, K. Y. Leong, and H. Y. Ng, "Anaerobic treatment of pharmaceutical wastewater: A critical review," *Bioresource Technology*, vol. 245, pp. 1238–1244, Dec. 2017, doi: 10.1016/j.biortech.2017.08.150.
[124] B. D. Deshpande, P. S. Agrawal, and M. K. N. Yenkie, "Advanced oxidative degradation of benzoic acid and 4-nitro benzoic acid–A comparative study," presented at the ADVANCES IN BASIC SCIENCE (ICABS 2019), Bahal, India, 2019, p. 210003. doi: 10.1063/1.5122650.
[125] Y. Deng and R. Zhao, "Advanced Oxidation Processes (AOPs) in Wastewater Treatment," *Curr Pollution Rep*, vol. 1, no. 3, pp. 167–176, Sep. 2015, doi: 10.1007/s40726-015-0015-z.
[126] A. Kale and M. K. N. Yenkie, "A Comparative Study of Efficiency of Different AOPs for Degradation of 4-nitrophenol in Batch and Continuous Systems," p. 9, 2016.
[127] D. M. Follansbee, J. D. Paccione, and L. L. Martin, "Globally Optimal Design and Operation of a Continuous Photocatalytic Advanced Oxidation Process Featuring Moving Bed Adsorption and Draft-Tube Transport," *Ind. Eng. Chem. Res.*, vol. 47, no. 10, pp. 3591–3600, May 2008, doi: 10.1021/ie071275r.

[128] R. Andreozzi, “Advanced oxidation processes (AOP) for water purification and recovery,” *Catalysis Today*, vol. 53, no. 1, pp. 51–59, Oct. 1999, doi: 10.1016/S0920-5861(99)00102-9.
[129] C. S. Slater, C. G. Uchrin, and R. C. Ahlert, “Ultrafiltration processes for the characterization and separation of landfill leachates,” *Journal of Environmental Science and Health . Part A: Environmental Science and Engineering*, vol. 20, no. 1, pp. 97–111, Jan. 1985, doi: 10.1080/10934528509375212.
[130] W. L. Ang, A. W. Mohammad, N. Hilal, and C. P. Leo, “A review on the applicability of integrated/hybrid membrane processes in water treatment and desalination plants,” *Desalination*, vol. 363, pp. 2–18, May 2015, doi: 10.1016/j.desal.2014.03.008.
[131] L.-C. Juang, D.-H. Tseng, and H.-Y. Lin, “Membrane processes for water reuse from the effluent of industrial park wastewater treatment plant: a study on flux and fouling of membrane,” *Desalination*, vol. 202, no. 1–3, pp. 302–309, Jan. 2007, doi: 10.1016/j.desal.2005.12.068.
[132] H. A. Abdulgader, V. Kochkodan, and N. Hilal, “Hybrid ion exchange – Pressure driven membrane processes in water treatment: A review,” *Separation and Purification Technology*, vol. 116, pp. 253–264, Sep. 2013, doi: 10.1016/j.seppur.2013.05.052.
[133] O. Ashrafi, L. Yerushalmi, and F. Haghighat, “Wastewater treatment in the pulp-and-paper industry: A review of treatment processes and the associated greenhouse gas emission,” *Journal of Environmental Management*, vol. 158, pp. 146–157, Aug. 2015, doi: 10.1016/j.jenvman.2015.05.010.
[134] A. R. Shawwa, D. W. Smith, and D. C. Sego, “Color and chlorinated organics removal from pulp mills wastewater using activated petroleum coke,” *Water Research*, vol. 35, no. 3, pp. 745–749, Feb. 2001, doi: 10.1016/S0043-1354(00)00322-5.
[135] S. Bhojwani, K. Topolski, R. Mukherjee, D. Sengupta, and M. M. El-Halwagi, “Technology review and data analysis for cost assessment of water treatment systems,” *Science of The Total Environment*, vol. 651, pp. 2749–2761, Feb. 2019, doi: 10.1016/j.scitotenv.2018.09.363.
[136] R. Kollmann *et al.*, “Renewable energy from wastewater - Practical aspects of integrating a wastewater treatment plant into local energy supply concepts,” *Journal of Cleaner Production*, vol. 155, pp. 119–129, Jul. 2017, doi: 10.1016/j.jclepro.2016.08.168.
[137] NPDES, USEPA, “Primer for Municipal Wastewater Treatment Systems,” 2004.
[138] N. Kulik, M. Trapido, A. Goi, Y. Veressinina, and R. Munter, “Combined chemical treatment of pharmaceutical effluents from medical ointment production,” *Chemosphere*, vol. 70, no. 8, pp. 1525–1531, Feb. 2008, doi: 10.1016/j.chemosphere.2007.08.026.
[139] K. M. Yenkie, W. Wu, R. L. Clark, B. F. Pfleger, T. W. Root, and C. T. Maravelias, “A roadmap for the synthesis of separation networks for the recovery of bio-based chemicals: Matching biological and process feasibility,” *Biotechnology Advances*, vol. 34, no. 8, pp. 1362–1383, Dec. 2016, doi: 10.1016/j.biotechadv.2016.10.003.
[140] R. Herna, “Superstructure Decomposition and Parametric Optimization Approach for the Synthesis of Distributed Wastewater Treatment Networks,” p. 17.

[141] E. Ahmetović and I. E. Grossmann, “Global superstructure optimization for the design of integrated process water networks,” *AIChE J.*, vol. 57, no. 2, pp. 434–457, Feb. 2011, doi: 10.1002/aic.12276.
[142] J. Garibay-Rodriguez, V. Rico-Ramirez, and J. M. Ponce-Ortega, “Mixed Integer Nonlinear Programming Model for Sustainable Water Management in Macroscopic Systems: Integrating Optimal Resource Management to the Synthesis of Distributed Treatment Systems,” *ACS Sustainable Chem. Eng.*, vol. 5, no. 3, pp. 2129–2145, Mar. 2017, doi: 10.1021/acssuschemeng.6b02128.
[143] F. Friedler, K. Tarján, Y. W. Huang, and L. T. Fan, “Graph-theoretic approach to process synthesis: axioms and theorems,” *Chemical Engineering Science*, vol. 47, no. 8, pp. 1973–1988, Jun. 1992, doi: 10.1016/0009-2509(92)80315-4.
[144] F. Friedler, L. T. Fan, and B. Imreh, “Process network synthesis: Problem definition,” *Networks*, vol. 31, no. 2, pp. 119–124, 1998.
[145] C. Gadipelly *et al.*, “Pharmaceutical Industry Wastewater: Review of the Technologies for Water Treatment and Reuse,” *Ind. Eng. Chem. Res.*, vol. 53, no. 29, pp. 11571–11592, Jul. 2014, doi: 10.1021/ie501210j.
[146] T. Deblonde, C. Cossu-Leguille, and P. Hartemann, “Emerging pollutants in wastewater: A review of the literature,” *International Journal of Hygiene and Environmental Health*, vol. 214, no. 6, pp. 442–448, Nov. 2011, doi: 10.1016/j.ijheh.2011.08.002.
[147] M. R. Kılınç and N. V. Sahinidis, “Exploiting integrality in the global optimization of mixed-integer nonlinear programming problems with BARON,” *Optimization Methods and Software*, vol. 33, no. 3, pp. 540–562, May 2018, doi: 10.1080/10556788.2017.1350178.
[148] B. Tribe, “U.S. Hide, Skin & Leather Fact Sheet,” The U.S. Sustainability Alliance. Accessed: Jun. 09, 2021. [Online]. Available: https://thesustainabilityalliance.us/u-s-hide-skin-leather-fact-sheet/
[149] G. Saxena and R. N. Bharagava, “Organic Pollutants in Tannery Wastewater and Bioremediation Approaches for Environmental Safety,” p. 34.
[150] M. A.-S. Elsheikh, “Tannery wastewater pre-treatment,” *Water Science and Technology*, vol. 60, no. 2, pp. 433–440, Jul. 2009, doi: 10.2166/wst.2009.351.
[151] E. M. Ijanu, M. A. Kamaruddin, and F. A. Norashiddin, “Coffee processing wastewater treatment: a critical review on current treatment technologies with a proposed alternative,” *Appl Water Sci*, vol. 10, no. 1, p. 11, Nov. 2019, doi: 10.1007/s13201-019-1091-9.
[152] S. Rattan, A. K. Parande, V. D. Nagaraju, and G. K. Ghiwari, “A comprehensive review on utilization of wastewater from coffee processing,” *Environ Sci Pollut Res*, vol. 22, no. 9, pp. 6461–6472, May 2015, doi: 10.1007/s11356-015-4079-5.
[153] C. M. Wisniewski, C. S. Slater, and M. J. Savelski, “Dynamic vibratory membrane processing for use in water recovery from soluble coffee product manufacturing wastewater,” *Clean Techn Environ Policy*, vol. 20, no. 8, pp. 1791–1803, Oct. 2018, doi: 10.1007/s10098-018-1569-4.
[154] M. V. O. Laurio, K. M. Yenkie, and C. S. Slater, “Optimization of vibratory nanofiltration for sustainable coffee extract concentration via response surface methodology,” *Separation Science and Technology*, pp. 1–19, Feb. 2021, doi: 10.1080/01496395.2021.1879858.

[155] C. Yu, Y. Liu, G. Chen, X. Gu, and W. Xing, “Pretreatment of Isopropanol Solution from Pharmaceutical Industry and Pervaporation Dehydration by NaA Zeolite Membranes,” *Chinese Journal of Chemical Engineering*, vol. 19, no. 6, pp. 904–910, Dec. 2011, doi: 10.1016/S1004-9541(11)60071-2.

[156] C. S. Slater, M. Savelski, G. Hounsell, D. Pilipauskas, and F. Urbanski, “Green design alternatives for isopropanol recovery in the celecoxib process,” *Clean Technologies and Environmental Policy*, vol. 14, no. 4, pp. 687–698, 2012.

[157] A. Jouyban, S. Nozohouri, and F. Martinez, “Solubility of celecoxib in {2-propanol (1) + water (2)} mixtures at various temperatures: Experimental data and thermodynamic analysis,” *Journal of Molecular Liquids*, vol. 254, pp. 1–7, Mar. 2018, doi: 10.1016/j.molliq.2018.01.033.

[158] K. D. Dahm and D. P. Visco, “Fundamentals of Chemical Engineering Thermodynamics,” 2015.

[159] S. I. Sandler, *Chemical, biochemical and engineering thermodynamics*, Fifth edition. Hoboken, NJ: Wiley, 2017.

[160] R. Parthasarathi and A. Dhawan, “Chapter 5 - In Silico Approaches for Predictive Toxicology,” in *In Vitro Toxicology*, A. Dhawan and S. Kwon, Eds., Academic Press, 2018, pp. 91–109. doi: 10.1016/B978-0-12-804667-8.00005-5.

[161] S. Hongmao, “Chapter 6 - Quantitative Structure–Property Relationships Models for Lipophilicity and Aqueous Solubility,” in *A Practical Guide to Rational Drug Design*, S. Hongmao, Ed., Woodhead Publishing, 2016, pp. 193–223. doi: 10.1016/B978-0-08-100098-4.00006-5.

[162] T. Stuyver, F. De Proft, P. Geerlings, and S. Shaik, “How Do Local Reactivity Descriptors Shape the Potential Energy Surface Associated with Chemical Reactions? The Valence Bond Delocalization Perspective,” *J. Am. Chem. Soc.*, vol. 142, no. 22, pp. 10102–10113, Jun. 2020, doi: 10.1021/jacs.0c02390.

[163] R. Srivastava, “Theoretical Studies on the Molecular Properties, Toxicity, and Biological Efficacy of 21 New Chemical Entities,” *ACS Omega*, vol. 6, no. 38, pp. 24891–24901, Sep. 2021, doi: 10.1021/acsomega.1c03736.

[164] E. A. Aboagye *et al.*, “Systematic Design of Solvent Recovery Pathways: Integrating Economics and Environmental Metrics,” *ACS Sustainable Chem. Eng.*, vol. 10, no. 33, pp. 10879–10887, Aug. 2022, doi: 10.1021/acssuschemeng.2c02497.

[165] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, “Mathematics for Machine Learning”.

[166] P. Wilmott, *Machine learning: an applied mathematics introduction*, First edition. Oxford, United Kingdom? Panda Ohana Publishing, 2019.

[167] T. Kolosova and S. Berestizhevsky, *Supervised machine learning: optimization framework and applications with SAS and R*, First edition. Boca Raton, FL: CRC Press, 2021.

[168] T. Jo, *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-65900-4.

[169] C. Aldrich and L. Auret, *Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods*. in Advances in Computer Vision and Pattern Recognition. London: Springer London, 2013. doi: 10.1007/978-1-4471-5185-2.

[170] A. A. Patel, “Hands-On Unsupervised Learning Using Python”.

[171] R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction”.
[172] P. Kulkarni, *Reinforcement and Systemic Machine Learning for Decision Making*, 1st ed. Wiley, 2012. doi: 10.1002/9781118266502.
[173] Y. Sun, X. Wang, N. Ren, Y. Liu, and S. You, “Improved Machine Learning Models by Data Processing for Predicting Life-Cycle Environmental Impacts of Chemicals,” *Environ. Sci. Technol.*, vol. 57, no. 8, pp. 3434–3444, Feb. 2023, doi: 10.1021/acs.est.2c04945.
[174] I. Belghazi, S. Rajeswar, A. Baratin, R. D. Hjelm, and A. Courville, “MINE: Mutual Information Neural Estimation,” Jan. 2018.
[175] A. Avadí, L. Nitschelm, M. Corson, and F. Vertès, “Data strategy for environmental assessment of agricultural regions via LCA: case study of a French catchment,” *Int J Life Cycle Assess*, vol. 21, no. 4, pp. 476–491, Apr. 2016, doi: 10.1007/s11367-016-1036-6.
[176] A. Géron, “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow”.
[177] “CIRpy — CIRpy 1.0.2 documentation.” Accessed: Aug. 25, 2023. [Online]. Available: https://cirpy.readthedocs.io/en/latest/
[178] “chemicals: Chemical properties component of Chemical Engineering Design Library (ChEDL) — Chemicals 1.1.4 documentation.” Accessed: Aug. 25, 2023. [Online]. Available: https://chemicals.readthedocs.io/
[179] N. O. of D. and Informatics, “NIST Chemistry WebBook.” Accessed: Aug. 25, 2023. [Online]. Available: https://webbook.nist.gov/chemistry/
[180] G. H. Thomson, “The DIPPR® databases,” *Int J Thermophys*, vol. 17, no. 1, pp. 223–232, Jan. 1996, doi: 10.1007/BF01448224.
[181] “The PubChem Project.” Accessed: Apr. 21, 2018. [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/
[182] “RDKit.” Accessed: Aug. 25, 2023. [Online]. Available: https://www.rdkit.org/
[183] “sklearn.impute.KNNImputer,” scikit-learn. Accessed: Aug. 31, 2023. [Online]. Available: https://scikit-learn/stable/modules/generated/sklearn.impute.KNNImputer.html
[184] C. C. Aggarwal, *Outlier Analysis*, 2nd ed. 2017. Cham: Springer International Publishing : Imprint: Springer, 2017. doi: 10.1007/978-3-319-47578-3.
[185] “sklearn.feature_selection.SequentialFeatureSelector,” scikit-learn. Accessed: Aug. 27, 2023. [Online]. Available: https://scikit-learn/stable/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html
[186] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
[187] M. Chen, Q. Liu, S. Chen, Y. Liu, C.-H. Zhang, and R. Liu, “XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System,” *IEEE Access*, vol. 7, pp. 13149–13158, 2019, doi: 10.1109/ACCESS.2019.2893448.

[188] J. Bergstra, D. Yamins, and D. Cox, "Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms," presented at the Python in Science Conference, Austin, Texas, 2013, pp. 13–19. doi: 10.25080/Majora-8b375195-003.

[189] P. I. Frazier, "Bayesian Optimization," in *Recent Advances in Optimization and Modeling of Contemporary Problems*, in INFORMS TutORials in Operations Research. , INFORMS, 2018, pp. 255–278. doi: 10.1287/educ.2018.0188.

[190] "TensorFlow," TensorFlow. Accessed: Aug. 27, 2023. [Online]. Available: https://www.tensorflow.org/

[191] M. La Rocca and C. Perna, "Opening the Black Box: Bootstrapping Sensitivity Measures in Neural Networks for Interpretable Machine Learning," *Stats*, vol. 5, no. 2, Art. no. 2, Jun. 2022, doi: 10.3390/stats5020026.

[192] S. Razavi *et al.*, "The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support," *Environmental Modelling & Software*, vol. 137, p. 104954, Mar. 2021, doi: 10.1016/j.envsoft.2020.104954.

[193] B. M. Pastore, M. J. Savelski, C. S. Slater, and F. A. Richetti, "Life cycle assessment of N-methyl-2-pyrrolidone reduction strategies in the manufacture of resin precursors," *Clean Techn Environ Policy*, vol. 18, no. 8, pp. 2635–2647, Dec. 2016, doi: 10.1007/s10098-016-1180-5.

[194] S. Qin, S. Jiang, J. Li, P. Balaprakash, R. C. V. Lehn, and V. M. Zavala, "Capturing Molecular Interactions in Graph Neural Networks: A Case Study in Multi-Component Phase Equilibrium".

[195] J. Zhou *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.

[196] J. G. Rittig, Q. Gao, M. Dahmen, A. Mitsos, and A. M. Schweidtmann, "Graph neural networks for the prediction of molecular structure-property relationships." arXiv, Jul. 25, 2022. Accessed: Aug. 28, 2023. [Online]. Available: http://arxiv.org/abs/2208.04852

[197] J. G. Rittig, K. C. Felton, A. A. Lapkin, and A. Mitsos, "Gibbs-Duhem-Informed Neural Networks for Binary Activity Coefficient Prediction." arXiv, May 31, 2023. Accessed: Aug. 28, 2023. [Online]. Available: http://arxiv.org/abs/2306.07937

[198] P. Juan-García *et al.*, "Resilience theory incorporated into urban wastewater systems management. State of the art," *Water Research*, vol. 115, pp. 149–161, May 2017, doi: 10.1016/j.watres.2017.02.047.

[199] D. L. Alderson, G. G. Brown, and W. M. Carlyle, "Operational Models of Infrastructure Resilience: Operational Models of Infrastructure Resilience," *Risk Analysis*, vol. 35, no. 4, pp. 562–586, Apr. 2015, doi: 10.1111/risa.12333.

[200] O. US EPA, "Creating Resilient Water Utilities (CRWU)." Accessed: Apr. 06, 2023. [Online]. Available: https://www.epa.gov/crwu

[201] "Flowsheet generation through hierarchical reinforcement learning and graph neural networks - Stops - 2023 - AIChE Journal - Wiley Online Library." Accessed: Aug. 28, 2023. [Online]. Available: https://aiche.onlinelibrary.wiley.com/doi/full/10.1002/aic.17938

[202] A. Thebelt, J. Wiebe, J. Kronqvist, C. Tsay, and R. Misener, “Maximizing information from chemical engineering data sets: Applications to machine learning,” *Chemical Engineering Science*, vol. 252, p. 117469, Apr. 2022, doi: 10.1016/j.ces.2022.117469.

[203] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis, “Physics-Informed Neural Networks for Heat Transfer Problems,” *Journal of Heat Transfer*, vol. 143, no. 6, Apr. 2021, doi: 10.1115/1.4050542.

[204] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nat Rev Phys*, vol. 3, no. 6, Art. no. 6, Jun. 2021, doi: 10.1038/s42254-021-00314-5.

[205] M. S. Alhajeri, F. Abdullah, Z. Wu, and P. D. Christofides, “Physics-informed machine learning modeling for predictive control using noisy data,” *Chemical Engineering Research and Design*, vol. 186, pp. 34–49, Oct. 2022, doi: 10.1016/j.cherd.2022.07.035.

[206] O. US EPA, “EPI Suite™-Estimation Program Interface.” Accessed: Nov. 24, 2023. [Online]. Available: https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface

Appendix A

Support Information for Chapter 3

A.1 Information for Municipal Case Study

A.1.1 Model equations and details

$i \in \boldsymbol{I}$ – technologies (used as subscript to variables)

{*flc* - flocculation,
sdm - sedimentation,
ftt - filtration,
ads - adsorption,
asl - activated sludge,
rbc - rotating biological contactors,
dis - disinfection,
mbrt - membrane bioreactor,
aop - advanced oxidation process,
blc - bleaching,
mbr - membrane processes
splt# - splitter and # = {1, 2, 3, 4}
mxr# - mixer and # = {1, 2, 3, 4}
byp# - bypass and # = {1, 2, 3, 4}}

$j \in \boldsymbol{J}$ – stream (used as subscript to variables)

{1, 2, 3, 4, ……, 49}

$k \in \boldsymbol{K}$ – components (used as subscript to variables)

{*Wtr* - water,
Ssld - solids,
Mtls - metals
Chem - chemicals
Flcnt - flocculants,
Oz - ozone,
NaClO - sodium hypochlorite,

L_Chlrn - liquid chlorine}

$s \in \boldsymbol{S}$ - stages {*s1, s2, s3, s4*}

A.1.2 Subsets

Subsets for technologies

$\boldsymbol{I}^{CST}$ – technologies with costs

{*flc, sdm, ftt, ads, asl, rbc, mbrt, dis, mbr, blc* }

$\boldsymbol{I}^{CF}$ – technologies with concentration factor

{*ftt, mbrt, sdm, mbr*}

$\boldsymbol{I}^{CONS}$ – technologies with consumables

{*ftt, ads, mbrt, mbr*}

$\boldsymbol{I}^{EAC}$ – technologies with externally added components

{*flc, aop, dis, blc*}

$\boldsymbol{I}^{BV}$ – technologies with binary variables

{*flc, sdm, ftt, ads, asl, rbc, mbrt, dis, mbr, blc, byp1, byp2, byp3, byp4* }

$\boldsymbol{I}^{S1}$ – technologies in stage 1

{*flc, byp1*}

$\boldsymbol{I}^{S2}$ – technologies in stage 2

{*ftt, sdm, byp2*}

$\boldsymbol{I}^{S3}$ – technologies in stage 3

{*ads, asl, rbc, dis, mbrt, byp3*}

$\boldsymbol{I}^{S4}$ – technologies in stage 4

{*aop, blc, mbr, byp4*}

Subsets for streams

$\boldsymbol{J}^{flc}$ – streams for flocculation

{2, 4, 5}

$\boldsymbol{J}^{byp1}$ – streams for bypass 1

{3, 6}

$\boldsymbol{J}^{sdm}$ – streams for sedimentation

{9, 13, 14}

$\boldsymbol{J}^{ftt}$ – streams for filtration

{8, 11, 12}

$\boldsymbol{J}^{byp2}$ – streams for bypass 2

{10, 15}

$\boldsymbol{J}^{ads}$ – streams for adsorption

{17, 23, 24}

$\boldsymbol{J}^{asl}$ – streams for activated sludge

{18, 25, 26}

$\boldsymbol{J}^{rbc}$ – streams for rotating biological containers

{19, 27, 28}

$\boldsymbol{J}^{dis}$ – streams for disinfection

{20, 29, 30, 31}

$\boldsymbol{J}^{mbrt}$ – streams for membrane bioreactor

{21, 32, 33}

$\boldsymbol{J}^{byp3}$ – streams for bypass 3

{22, 34}

$\boldsymbol{J}^{aop}$ – streams for advanced oxidation process

{36, 40, 41, 42}

$\boldsymbol{J}^{mbr}$ – streams for membrane processes

{38, 46, 47}

$\boldsymbol{J}^{blc}$ – streams for bleaching

{37, 43, 44, 45}

$\boldsymbol{J}^{byp4}$ – streams for bypass 4

{39, 48}

Subsets for components

$\boldsymbol{K}^{S}$ – components in streams

{*Wtr, Ssld, Mtls, Chem, Flcnt, Oz, NaClO, L_Chlrn*}

$\boldsymbol{K}^{SP}$ – components in initial wastewater stream

{*Wtr, Ssld, Mtls, Chem*}

$\boldsymbol{K}^{CONT}$ – components that are contaminants

{*Ssld, Mtls, Chem*}

$\boldsymbol{K}^{EAC}$ – components that are externally added

{*Flcnt, Oz, NaClO, L_Chlrn*}

A.1.3 Dynamic sets for connectivity

$\boldsymbol{J_i}$ - streams of technology *i*

$\boldsymbol{Jin_i}$ - inlet streams of technology *i*

$\boldsymbol{Jout_i}$ - outlet streams of technology *i*

$\boldsymbol{K_i}$ - components *k* in technology *i*

$\boldsymbol{K_j}$ - components *k* in stream *j*

A.1.4 Model Parameters

General Parameters

ρ_k (kg/m^3) = Density of component *k*

π_{WW} (m^3 WW/h) = Entering volumetric flowrate of wastewater (WW)

$\pi^{Rep}{}_i$ ($/unit) = Replacement cost of consumables per unit capacity in technology *i*

μ (N-s/m^2) = viscosity of fluid

η_i (%) = efficiency of technology *i*

θ_i^R (hr) – residence time in technology *i*

θ_i^{Rep} (h/year) = Replacement time for consumables in technology *i*

τ_{ann} (h/annum) = annual operation in hours (330 day x 24 h/day = 7920 hrs)

$C0_i$ ($/capacity) = standard capacity cost in technology *i*

g (m/s^2) = gravitational constant

nc = cost scaling index (2/3 rule)

$Nlabr_i$ (#/h) = standard # of laborers required for technology *i* per hour

$Q0_i$ (m^3 or m^2 or m^3/h) = standard capacity of technology *i*

MW_k (kg/kmol) = molecular weight of component *k*

Min_k (kg/m3) = initial mass concentration of component *k*

$Cpur_k$ ($/kg) = purchase cost of added component *k*

Wsp_i (kW/h) = standard power required for technology *i* per hour

MM_k (--) = Big-M constant for component *k*

Φ_k (kg/kg WW) = amount of externally added component *k*

dp (m) = diameter of particle

$\xi_{j,i}$ (--) = retention factor of component *k* for technology *i* {*ftt*, *mbrt*, and *mbr* technologies}

ς_{RF} (--) = capital recovery/ charge factor (0.11)

ζ_i (m^3/m^2h) = flux of technology i { *ftt*, *mbrt*, and *mbr* technologies}

κd_i (h^{-1}) = decay of biomass coefficient of technology i {*asl* technology}

γ_i (kg/kg) = biomass yield of technology i { asl technology}

χ_i (m^3/m^2h) = hydraulic loading of technology i {rbc technology}

BMC_{mult} (--) = bare module cost multiplier (5.4)

C_{Lab} ($/h) = labor cost – operator basis (30)

C_{Elec} ($/kW) = cost of electricity per hour (0.1)

Evaluated Parameters

SOR_i (m/s) = surface overflow rate id sedimentation

U_i (m/s) = settling velocity of technology i

A.1.5 Model Variables

General Variables

$Cc_{,i}$ ($) = Purchase cost of technology $i \in \boldsymbol{I}^{\mathbf{CST}}$

CF_i (m^3/m^3) = Concentration factor for technologies $i \in \boldsymbol{I}^{\boldsymbol{CF}}$

Cpr_k ($/h) = Purchase cost of added components $k \in \boldsymbol{K}^{\boldsymbol{EAC}}$

$M_{j,k}$ (kg/h) = Mass flowrate of component k in stream j

Qc_{i} (m^3 or m^2 or m^3/h) = capacity cost of technologies $i \in \boldsymbol{I}^{\mathbf{CST}}$

PW_i (kW/h) = power requirements for technologies $i \in \boldsymbol{I}^{\mathrm{CST}}$

$Nlbr_i$ (#/h) = number of laborers required for technology $i \in \boldsymbol{I}^{\mathrm{CST}}$

Yo_i (kg/kg) = observed bacteria yield of technology i (*asl* technology)

Sr_i (m/h) = settling rate of unit i (*asl* technology)

X_i (kg/h) = biomass produced in technology i (*asl* technology)

Srt_i (h) = solids residence time in technology i (*asl* technology)

D_i (m) = diameter of technology i (*mbrt* technology)

Consi ($/annum) = consumable cost of technology $i \in \boldsymbol{I}^{\mathrm{CONS}}$

Binary Variables

y_i (--) = binary variables for technologies to selected $i \in \boldsymbol{I}^{\mathrm{BV}}$

Stage-wise Costing Variables

$CCAC_{Nstg}$ = annualized capital (fixed) cost in n^{th} stage

$CCRM_{Nstg}$ = material cost in n^{th} stage

$CCCS_{Nstg}$ = consumable cost in n^{th} stage

$CCLB_{Nstg}$ = labor cost in n^{th} stage

$CCUT_{Nstg}$ = utility cost in n^{th} stage

$CCOT_{Nstg}$ = other cost in n^{th} stage (plant overhead and supervision costs)

$CCTC_{Nstg}$ = total cost in n^{th} stage (all cost added in that particular stage)

CCTPC = total cost for process (summation of total cost in each stage)

A.1.6 Model Equations:

Initial wastewater flowrate equations:

$$M_{1,k} = (\sum_k Min_k)\pi_{WW} \; ; \; \forall \; k \in \boldsymbol{K^{SP}}$$

Component balances in all technologies:

$$\sum_{j \in \boldsymbol{Jin_i}} M_{j,k} = \sum_{j \in \boldsymbol{Jout_i}} M_{j,k} \; ; \; \forall \; k \in \boldsymbol{K^S}$$

Cost of technologies:

$$\left(\frac{C_{c_i}}{C0_i}\right) = \left(\frac{Q_{c_i}}{Q0_i}\right)^{nc} \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Labor requirements in technologies:

$$Nlbr_i Q0_i = Nlabr_i Qc_i \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Consumable costs in technologies:

$$Cons_i = \frac{\tau_{ann}}{\theta_i^{Rep}} \pi_i^{Rep} Qc_i \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Logical equations:

$$M_{i,j} - M1_k y_i \leq 0 \; ; \; \forall \; i \in \boldsymbol{I^{BV}} \; , \; j \in \boldsymbol{J}, \; k \in \boldsymbol{K_i} \; and \; \boldsymbol{K_j}$$

Selection of technologies in each stage:

Preliminary (Pretreatment) stage:

$$y_{flc} + y_{byp,1} = 1$$

Primary Treatment Stage:

$$y_{ftt} + y_{sdm} + y_{byp,2} = 1$$

Secondary Treatment Stage:

$$y_{ads} + y_{asl} + y_{rbc} + y_{dis} + y_{mbrt} + y_{byp,3} = 1$$

Tertiary Treatment Stage:

$$y_{aop} + y_{mbr} + y_{blc} = 1$$

Preliminary (Pretreatment) Stage Model Equations for Technologies

Flocculation (flc):

Flocculent added:

$$M_{5,Flcnt} = \Phi_{Flcnt} \sum_{k \in \boldsymbol{K^{CONT}}} M_{2,k}$$

Flocculent cost:

$$Cpur_k = \pi_{Flcnt} M_{5,Flcnt}$$

Volume of flocculation unit:

$$Qc_{flc} = \theta_{flc}^{R} \left[\sum_{k \in \boldsymbol{K^{S}}} \left(\frac{M_{2,k}}{\rho_k}\right)\right]$$

Power required in flocculation unit:

$$PW_{flc} = Wsp_{flc} Qc_{flc}$$

Primary Stage Model Equations for Technologies

Sedimentation (sdm):

Efficiency equation:

$$\eta_{sdm} = \frac{M_{13,k}}{M_{9,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Concentration factor (CF_{sdm}):

$$CF_{sdm} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{sdm} \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$2 \leq CF_{sdm} \leq 15$

Written as: $CF_{sdm} \leq 15 y_{sdm}$ *and* $CF_{sdm} \geq 2 y_{sdm}$

Area of sedimentation unit:

$$Qc_{sdm} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]}{SOR_{sdm}}$$

Power required in sedimentation unit:

$$PW_{sdm} = Wsp_{sdm} Qc_{sdm}$$

Filtration (ftt):

Retention factor equation

$$\xi_{ftt_k} = \frac{M_{11,k}}{M_{8,k}} \; ; \; k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{ftt_k} M_{8,k} = M_{11,k}$

Concentration factor equation (ftt):

$$CF_{ftt} = \frac{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right]}{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{11,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{ftt}\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$1 \leq CF_{ftt} \leq 30$

Written as: $CF_{ftt} \leq 30y_{ftt}\ and\ CF_{ftt} \geq 1y_{ftt}$

Area of filtration unit (flux balance):

$$Qc_{ftt} = A_{ftt} = \frac{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)}{\zeta_{ftt}CF_{ftt}}$$

Written as: $Qc_{ftt}\zeta_{ftt}CF_{ftt} = \left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)$

Power requirements for filtration unit:

$$PW_{ftt} = Wsp_{ftt}Qc_{ftt}$$

Power required in sedimentation unit:

$$PW_{sdm} = Wsp_{sdm}Qc_{sdm}$$

Filtration (ftt):

Retention factor equation

$$\xi_{ftt_k} = \frac{M_{11,k}}{M_{8,k}}\ ;\ k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{ftt_k}M_{8,k} = M_{11,k}$

Concentration factor equation (ftt):

$$CF_{ftt} = \frac{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right]}{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{11,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{ftt}\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$1 \leq CF_{ftt} \leq 30$

Written as: $CF_{ftt} \leq 30y_{ftt}\ and\ CF_{ftt} \geq 1y_{ftt}$

Area of filtration unit (flux balance):

$$Qc_{ftt} = A_{ftt} = \frac{\left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)}{\zeta_{ftt}CF_{ftt}}$$

Written as: $Qc_{ftt}\zeta_{ftt}CF_{ftt} = \left[\sum_{k\in \boldsymbol{K^{SP}}}\left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)$

Power requirements for filtration unit:

$$PW_{ftt} = Wsp_{ftt} Qc_{ftt}$$

Power required in sedimentation unit:

$$PW_{sdm} = Wsp_{sdm} Qc_{sdm}$$

Filtration (ftt):

Retention factor equation

$$\xi_{ftt_k} = \frac{M_{11,k}}{M_{8,k}} \; ; \; k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{ftt_k} M_{8,k} = M_{11,k}$

Concentration factor equation (ftt):

$$CF_{ftt} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{11,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{ftt} \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$1 \leq CF_{ftt} \leq 30$

Written as: $CF_{ftt} \leq 30 y_{ftt}$ *and* $CF_{ftt} \geq 1 y_{ftt}$

Area of filtration unit (flux balance):

$$Qc_{ftt} = A_{ftt} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)}{\zeta_{ftt} CF_{ftt}}$$

Written as: $Qc_{ftt} \zeta_{ftt} CF_{ftt} = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt} - 1)$

Power requirements for filtration unit:

$$PW_{ftt} = Wsp_{ftt} Qc_{ftt}$$

Tertiary Stage Model Equations for Technologies

Advanced Oxidation Processes:

Mass of ozone needed for advanced oxidation processes unit

$$M_{42,Oz} = \Phi_{Oz} \sum_{k \in \boldsymbol{K^{CONT}}} M_{31,k}$$

Efficiency equation:

$$\eta_{aop} = \frac{M_{40,k}}{M_{36,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Volume of advanced oxidation processes unit:

$$Qc_{aop} = V_{aop} = \theta^{R}_{aop} \left[\sum_{k \in \boldsymbol{K^{S}}} \left(\frac{M_{36,k}}{\rho_k}\right)\right]$$

Power required for advanced oxidation processes unit:

$PW_{aop} = Wsp_{aop} Qc_{aop}$

Membrane Processes:

Retention factor equation

$$\xi_{mbr_k} = \frac{M_{46,k}}{M_{38,k}} \; ; \; k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{mbr_k} M_{38,k} = M_{46,k}$

Concentration factor equation (mbr):

$$CF_{mbr} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{46,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{mbr} \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{46,k}}{\rho_k}\right)\right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right]$

$1 \leq CF_{mbr} \leq 35$

Written as: $CF_{mbr} \leq 35 y_{mbr}$ *and* $CF_{mbr} \geq 1 y_{mbr}$

Area of membrane processes unit (flux balance):

$$Qc_{mbr} = A_{mbr} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right](CF_{mbr} - 1)}{\zeta_{mbr} CF_{mbr}}$$

Written as: $Qc_{mbr} \zeta_{mbr} CF_{mbr} = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right](CF_{mbr} - 1)$

Power requirements for membrane unit:

$$PW_{mbr} = Wsp_{mbr} Qc_{mbr}$$

Bleaching:

Efficiency equation:

$$\eta_{blc} = \frac{M_{43,k}}{M_{37,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Mass of disinfectant required for disinfection unit

$$M_{45,NaClO} = \Phi_{NaClO} \sum_{k \in \boldsymbol{K^{CONT}}} M_{37,k}$$

Volume of disinfection unit:

$$Qc_{dis} = V_{dis} = \theta_{dis}^{R} \left[\sum_{k \in \boldsymbol{K^{S}}} \left(\frac{M_{37,k}}{\rho_k}\right)\right]$$

Power required for disinfection unit:

$$PW_{dis} = Wsp_{dis} Qc_{dis}$$

Cost Model Equations

Stagewise Cost Equations

Annualized capital cost in each stage:

$$CCAC_{Nstg} = 1.66\, \varsigma_{RF} BMC_{mult} \sum_{i \in istg\{1,2,3,4\}} C_{ei}$$

Material Cost:

$$CCRM_{s1} = [\tau_{ann}(Cpur_{Flcnt})]$$

$$CCRM_{s2} = 0$$

$$CCRM_{s3} = [\tau_{ann}(Cpur_{L_{Chlrn}})]$$

$$CCRM_{s4} = [\tau_{ann}(\, Cpur_{Oz} + Cpur_{NaClO})]$$

Consumable Cost:

$$CCCS_{s1} = 0$$

$$CCCS_{s2} = Cons_{ftt}$$

$$CCCS_{s3} = Cons_{ads} + Cons_{mbrt}$$

$$CCCS_{s4} = Cons_{mbr}$$

Labor Cost

$$CCLC_{Nstg} = \tau_{ann} C_{Lab} \sum_{i \in istg\{1,2,3,4\}} Nlb_i$$

Utility Cost

$$CCUC_{Nstg} = \tau_{ann} C_{elec} \sum_{i \in istg\{1,2,3,4\}} PW_i$$

Other Cost

$$CCOC_{Nstg} = 2.78\, \tau_{ann} C_{Lab} \sum_{i \in istg\{1,2,3,4\}} Nlb_i$$

Total Cost in each Stage

$$CCTC_{Nstg} = CCAC_{Nstg} + CCRM_{Nstg} + CCCS_{Nstg} + CCLC_{Nstg} + CCUC_{Nstg} CCOC_{Nstg}$$

Total Category Cost:

$$CCTAC = \sum_{n \in Nstg\{1,2,3,4\}} CCAC_n$$

$$CCTRM = \sum_{n \in Nstg\{1,2,3,4\}} CCRM_n$$

$$CCTCS = \sum_{n \in Nstg\{1,2,3,4\}} CCCS_n$$

$$CCTLC = \sum_{n \in Nstg\{1,2,3,4\}} CCLC_n$$

$$CCTUC = \sum_{n \in Nstg\{1,2,3,4\}} CCUC_n$$

$$CCTOC = \sum_{n \in Nstg\{1,2,3,4\}} CCOC_n$$

$$CCTPC = CCTAC + CCTRM + CCTCS + CCTLC + CCTUC + CCTOC$$

Objective Function 1:

$$Obj = Min\ CCTPC$$

A.1.7 Model parameters and input data

Table A1

Density and Molecular Weight of Components

Component	**Value (**kg/m^3**)**	**Value (**kg/kmol**)**
Water	1000	18
Solid particles	1800	102
Metals	2500	98
Chemicals	1100	48
Ozone	2.14	48
Sodium Hypochlorite	1110	74.44
Flocculant	2200	2200
Liquid chlorine	1470	70.91

Table A2

Purchase Cost of Added Components

Component	**Value ($/kg)**
Ozone	3.53
Sodium Hypochlorite	0.35
Flocculant	0.75
Liquid chlorine	2

Table A3

Standard Capacity, Cost, Laborers, and Power for Technologies

Unit Operation (costing capacity)	**Standard Capacity (Units)**	**Standard Capacity Cost (million $)**	**Standard Laborers Required (#/h)**	**Standard Power Required (kW/h)**
Flocculation (Volume)	2000 (m^3)	0.538	0.1	0.0002
Filtration (Area)	80 (m^2)	0.039	0.5	0.1
Sedimentation (Volume)	2500 (m^3)	1.128	0.1	0.0002
Adsorption (Volume)	100 (m^3)	0.12	0.2	0.3
Membrane Bioreactor (Area)	40 (m^2)	1.194	0.1	0.2
Rotational Biological Container (Area)	185 (m^2)	0.045	0.3	0.01
Activated Sludge (Volume)	250 (m^3)	0.569	0.4	0.2
Disinfection (Volume)	540 (m^3)	0.627	0.7	0.5
Membrane Processes (Area)	80 (m^2)	0.938	0.5	0.2
Advanced Oxidation Processes (Volume)	1000 (m^3)	0.735	0.1	0.5
Bleaching (Volume)	500 (m^3)	0.100	0.5	0.33

Table A4

Replacement Time for Technologies with Consumables

Technology	Value (h/yr)
Filtration	2000
Adsorption	720
Membrane Processes	1000
Membrane Bioreactor	1000

Table A5

Replacement Cost for Technologies with Consumables

Technology	Value ($/Unit)
Filtration	100
Adsorption	74.16
Membrane Processes	400
Membrane Bioreactor	400

A.1.7.1 Flocculation (flc):

Flocculent added (kg/kg) – 0.005

Residence time (h) – 0.5

A.1.7.2 Sedimentation (sdm):

Efficiency – 80%

Depth – 3m

A.1.7.3 Filtration (ftt):

Retention factor: Water – 0.05, Solids – 0.80, Metals – 0.10, Chemicals – 0.05
Flux ($m^3/m^2.h$): 0.2

A.1.7.4 Membrane Bioreactor (mbrt):

Retention factor: Water – 0.005, Solids – 0.95, Metals – 0.85, Chemicals – 0.50
Flux ($m^3/m^2.h$): 0.025

A.1.7.5 Adsorption (ads):

Empty bed contact time (h): 0.25
Density of granular activated carbon (GAC) (kg/m^3): 1030
Efficiency: 90%
Void fraction of GAC: 50%

A.1.7.6 Activated Sludge (asl):

Decay of biomass coefficient (h^{-1}): 0.0021
Biomass yield (kg/kg): 0.5
Hydraulic retention time (h): 2
Efficiency: 80%

A.1.7.7 Rotating Biological Contactors (rbc):

Efficiency: 80%
Hydraulic loading (m^3/m^2h): 20

A.1.7.8 Disinfection (dis):

Efficiency: 95%

Ratio of liquid chlorine to contaminant (kg/kg): 0.00173

Residence time (h): 2

A.1.7.9 Advanced Oxidation Processes (aop):

Ratio of ozone to contaminant (kg/kg): 0.000173

Efficiency: 98%

Residence time (h): 0.21

A.1.7.10 Membrane Processes (mbr):

Retention factor: Water – 0.05, Solids – 0.90, Metals – 0.90, Chemicals – 0.95

Flux ($m^3/m^2.h$): 0.0856

A.1.7.11 Bleaching (blc):

Efficiency: 98%

Ratio of sodium hypochlorite to contaminant (kg/kg): 0.00173

Residence time (h): 2

A.1.8 Integer-cuts for determining 1st, 2nd, and 3rd best configuration

$$\sum_{y_{ibv}=1} y_{ibv} - \sum_{y_{ibv}=0} y_{ibv} \leq (\#\ of\ 1'\ \text{sin}\ solution) - 1$$

First integer-cut to determine first best alternative:

$$[y_{blc} + y_{sdm} + y_{flc} + y_{ads}] - [y_{ftt} + y_{asl} + y_{rbc} + y_{mbrt} + y_{aop} + y_{mbr} + y_{dis}] \leq 4 - 1$$

Second integer-cut to determine the second best alternative

$$[y_{blc} + y_{sdm} + y_{flc} + y_{dis}] - [y_{ftt} + y_{asl} + y_{rbc} + y_{mbrt} + y_{aop} + y_{mbr} + y_{ads}] \leq 4 - 1$$

Third integer-cut to determine third-best alternative:

$$[y_{blc} + y_{ftt} + y_{flc} + y_{dis}] - [y_{sdm} + y_{asl} + y_{rbc} + y_{mbrt} + y_{aop} + y_{mbr} + y_{ads}] \leq 4 - 1$$

A.1.9 Cost Distribution

Table A6

Breakdown of Stage-wise Cost of Purification

Cost Category	**Stage-wise Cost Distribution**			
	Pretreatment	**Primary**	**Secondary**	**Tertiary**
Annualized Capital Cost($/y)	4.81E+04	6.42E+04	4.59E+04	5.16E+04
Material Cost($/y)	5.97E+05	0.00E+00	0.00E+00	4.58E+05
Consumable Cost($/y)	0.00E+00	0.00E+00	1.98E+04	0.00E+00
Labor Cost ($/y)	6.61E+02	3.37E+02	1.16E+04	4.53E+04
Utilities Cost($/y)	8.81E+00	5.61E+00	5.78E+03	4.98E+04
Other Cost($/y)	1.18E+03	2.04E+04	2.06E+04	8.05E+04
Total($/y)	6.47E+05	8.50E+04	1.04E+05	6.85E+05
Percentage Distribution				
Annualized capital cost	7.43%	75.56%	44.28%	7.54%
Raw material cost	92.28%	0.00%	0.00%	66.82%
Consumable cost	0.00%	0.00%	19.14%	0.00%
Labor cost	0.10%	0.40%	11.15%	6.61%
Utilities Cost	0.00%	0.01%	5.58%	7.27%
Other Cost	0.18%	24.04%	19.85%	11.76%
Total	100.00%	100.00%	100.00%	100.00%

A.1.10 Sustainable Process Index (SPI) Calculations

A.1.10.1 SPI for Municipal Wastewater Using Values from GAMS Codes

Parameters

F_{RR} (kg/yr) - feed of a processed resource (752400000)

y_R (kg/m^2.yr) - specific yield (243.1542)

$y_{EI,RN}$ (kWh/m^2.yr) - mean industrial energy yield or mean industrial energy supply density (7)

F_{RN} (kg/yr) - feed of a processed resource (2172028)

C_N (\$/kg) - price of the material (world market price, taxes excluded) (7)

C_E (\$/kWh) - price of one kilowatt-hour of energy (industrial price, taxes excluded) (0.1)

F_E (kWh/yr) - energy used in the process (555659.28)

y_E (kWh/m^2.yr) - energy yield (43)

C_I (\$) - total cost of energy for indirect installation (122720)

LS (yr) - depreciated area over the life-span years (30)

$y_{EI,II}$ (kWh/m^2.yr) - industrial energy supply density or yield (43)

N_S (cap/yr) - total number of workers in the treatment plant (80.19)

y_S (cap/m^2.yr) - yield factor due to staff (3.59E-05)

$c_{c,m}$ (kg$_m$/kg) - allowable concentration of substance, *m* [Solids, Chemical, Metals, Water] in the compartment, *c* [air, water, soil]

R_c (kg/m^2.yr) - rate of renewal of the environmental compartments, c [air, water, soil]

Estimated Parameters

$E_{D,RN}$ - energy demand to supply one kilogram of the material in question for non-renewable energy (kWh/kg)

$$E_{D,RN} = \frac{0.95\ C_N}{C_E} = \frac{0.95\ \times 7}{0.1} = 66.5$$

$E_{D,II}$ - energy demand to supply one kilogram of the material in question for indirect land energy (kWh/yr)

$$E_{D,II} = \frac{0.54\ C_I}{C_E . LS} = \frac{0.54\ \times 122720}{0.1 \times 30} = 22089.6$$

$S_{c,m}$ - dissipation to potential sink (kg_m/m^2yr)

$S_{c,m} = R_c . c_{c,m}$

Analysis for Rsoil ($kg/m^2.yr$)

Rate of soil renewal (RSN) in the US is 2.2E-04 m/yr

Assuming the soil is loamy with a 50% pore space, then the bulk density is 1330 kg/m3

$R_{soil} = RSN \times Den_{bulk} = 0.00022 \times 1330 = 0.2926$

Since contaminants are categorized into solids, chemicals, and metals, we used contaminants that had the smallest allowable concentration for each category in the compartments to estimate S. For solids contaminant, we used lead (Pb), for chemical we used Chromium (Cr), and for metals we used lead (Pb).

Analysis for Ssoil,m ($kgm/m^2.yr$)

$S_{soil,m} = R_{soil} . c_{soil,m}$

Table A7

Data on soil yield for municipal wastewater contaminants

Component (m)	**$C_{(soil,m)}$ (kgm/kg)**	**$S_{(soil,m)}$ ($kgm/m^2.y$)**
Solids (Sslds)	1.00E+00	2.93E-01
Chemical (Chem)	1.00E-04	2.93E-05
Metals (Mtls)	2.00E-06	5.85E-07
Water (Wtr)	1.00E+00	2.93E-01

Analysis for Rwater (kg/m^2.yr)

Average rate of precipitation (RP) form Jan, 2009 to Dec, 2019 in the US is 31.91 in/yr (0.810514 m/yr)

Seeping ratio (SR) of water is 0.30

$R_{water} = RP \times SR \times Den_{water} = 0.810514 \times 0.30 \times 1000 = 243.1542$

Analysis for Swater,m (kgm/m^2.yr)

$S_{water,m} = R_{water} \cdot c_{water,m}$

Table A8

Data on water yield for municipal wastewater contaminants

Component (m)	**$C_{(water,m)}$ (kgm/kg)**	**$S_{(water,m)}$ (kgm/m^2.y)**
Solids (Sslds)	1.67E-06	4.05E-04
Chemical (Chem)	1.00E-07	2.43E-05
Metals (Mtls)	1.50E-08	3.65E-06
Water (Wtr)	1.00E+00	2.43E+02

$S_{air,chem}$ (kgm/m^2.yr) = 6.50E-03

Table A9

$F_{P,c,m}$ product flow rate to compartment c, by substance, m (kg_m/yr) [from GAMS code]

Component (m)	Fraction of m into compartment c			$F_{P,c,m}$		
	Air	Water	Soil	Air	Water	Soil
Solids (Sslds)	0	0.05	0.95	0.00E+00	3.17E+03	6.02E+04
Chemical (Chem)	0.03	0.9	0.07	9.50E+00	2.85E+02	2.22E+01
Metals (Mtls)	0	0.7	0.3	0.00E+00	2.22E+01	9.50E+00
Water (Wtr)	0	0.95	0.05	0.00E+00	7.15E+08	3.76E+07

Variables

A_R - area for raw material production (m^2/yr)

A_{RR} - area for renewable raw material production (m^2/yr)

A_E - area for energy production (m^2/yr)

A_I - area for installation for equipment and other infrastructure (m^2/yr)

A_{ID} - area for direct installation (m^2/yr)

A_{II} - area for indirect installation (m^2/yr)

A_S - area for staff (m^2/yr)

$A_{P,c,m}$ - area for dissipating a single component of particular product flow to a given compartment (m^2/yr)

$A_{PS,c}$ - area assigned to the dissipation of a particular product stream, S (m^2/yr)

A_P - area for product dissipation (m^2/yr)

A_{tot} - total area (m^2/m^3WW-yr)

Area for Raw Material Production (A_R)

Area for Renewable Raw Material Production (A_{RR})

$$A_{RR} = \frac{F_{RR}}{y_R} = \frac{7524000000}{243.1542} = 3.09E06$$

Area for Non-Renewable Raw Material Production (A_{RN})

$$A_{RN} = \frac{F_{RN}.E_{D,RN}}{y_{EI,RN}} = \frac{2172028 \times 66.5}{7} = 2.06E07$$

$$A_R = A_{RR} + A_{RN} = 2.37\text{E}07$$

Area for Energy Production (A_E)

$$A_E = \frac{F_E}{y_E} = \frac{555659.28}{43} = 1.29E4$$

Area for Installations (A_I)

$$A_I = A_{ID} + A_{II}$$

Area for Direct Installation (A_{ID})[from GAMS Code]

Flocculation technology = 18.54

Sedimentation technology = 35.43

Adsorption technology = 4.86

Bleaching technology = 38.09

$$A_{ID} = 96.91$$

Area for Indirect Installation (A_{II})

$$A_{II} = \frac{E_{D,II}}{y_{EI,II}} = \frac{22089.6}{43} = 5.14E2$$

Area for Staff (A_S)

$$A_S = N_S.a_{in} = \frac{N_S}{y_S} = \frac{80.19}{0.0000359} = 2.23E6$$

Area for Product Dissipation into Various Environmental Compartment (A_P)

$$A_{P,c,m} = \frac{F_{P,c,m}}{S_{c,m}}$$

Table A10

Area needed for emission dissipation for municipal wastewater treatment

Component (m)	$A_{P,c,m}$		
	Air	Water	Soil
Solids (Sslds)	--	7.82E+06	2.06E+05
Chemical (Chem)	1.46E+03	1.17E+07	7.58E+05
Metals (Mtls)	--	6.08E+06	1.62E+07
Water (Wtr)	--	2.94E+06	1.29E+08

$A_{PS,c} = max_m(A_{P,c,m})$

Highlighted are the maximum values for each component

Table A11

Highest area from emission dissipation

Component (m)	AP,c,m		
	Air	Water	Soil
Solids (Sslds)	--	7.82E+06	2.06E+05
Chemical (Chem)	1.46E+03	1.17E+07	7.58E+05
Metals (Mtls)	--	6.08E+06	1.62E+07
Water (Wtr)	--	2.94E+06	1.29E+08

$$A_P = \sum_c A_{PS,c} = 1.40E8$$

$A_{tot} = A_R + A_E + A_I + A_S + A_P$ = 2.10E+02 (m^2/m^3WW-yr)

A.1.10.2 SPI for Direct Disposal of Municipal Wastewater

We considered on the area needed to sustainably embed the wastewater into the ecosystem, i.e. A_P

Area for Product Dissipation into Various Environmental Compartment (A_P)

Table A12

$F_{P,c,m}$ product flow rate to compartment c, by substance, m (kg_m/yr) [from GAMS code]

Component (m)	**Fraction of**		**$F_{P,c,m}$**	
	Water	Soil	Water	Soil
Solids (Sslds)	0.05	0.95	8.76E+06	1.66E+08
Chemical (Chem)	0.95	0.05	8.32E+05	4.38E+04
Metals (Mtls)	0.7	0.3	6.13E+04	2.63E+04
Water (Wtr)	0.95	0.05	8.32E+08	4.38E+07

$$A_{P,c,m} = \frac{F_{P,c,m}}{S_{c,m}}$$

Table A13

Area needed for emission dissipation for direct disposal of wastewater

Component (m)	**AP,c,m**	
	Water	Soil
Solids (Sslds)	7.82E+06	2.06E+05
Chemical (Chem)	1.17E+07	7.58E+05
Metals (Mtls)	6.08E+06	1.62E+07
Water (Wtr)	2.94E+06	1.29E+08

$A_{PS,c} = max_m(A_{P,c,m})$

Highlighted are the maximum values for each component

Table A14

Highest area from emission dissipation for direct disposal

Component (m)	**$A_{P,c,m}$**	
	Water	Soil
Solids (Sslds)	2.16E+10	5.69E+08
Chemical (Chem)	3.42E+10	1.50E+09
Metals (Mtls)	1.68E+10	4.49E+10
Water (Wtr)	3.42E+06	1.50E+08

$$A_P = \sum_c A_{PS,c} = 7.91E10$$

$A_{tot} = A_P$ = 9.03E+04 (m^2/m^3WW-yr)

A.2 Information for Pharmaceutical Case Study

A.2.1 Model equations and details

$i \in \boldsymbol{I}$ – technologies (used as subscript to variables)

{*flc* - flocculation,
sdm - sedimentation,
ftt - filtration,
ads - adsorption,
asl - activated sludge,
rbc - rotating biological contactors,
dis - disinfection,
mbrt - membrane bioreactor,
aop - advanced oxidation process,
blc - bleaching,
mbr - membrane processes
splt# - splitter and # = {1, 2, 3, 4}
mxr# - mixer and # = {1, 2, 3, 4}
byp# - bypass and # = {1, 2, 3, 4}}

$j \in \boldsymbol{J}$ – stream (used as subscript to variables)

{1, 2, 3, 4, ……, 49}

$k \in \boldsymbol{K}$ – components (used as subscript to variables)

{*Wtr* - water,
Ssld - solids,
Mtls - metals
Chem - chemicals
API - *active pharmaceutical ingredients*
Flcnt - flocculants,
Oz - ozone,
NaClO - sodium hypochlorite,

L_Chlrn - liquid chlorine}

$s \in \boldsymbol{S}$ - stages {*s1, s2, s3, s4*}

A.2.2 Subsets

Subsets for technologies

$\boldsymbol{I}^{CST}$ – technologies with costs

{*flc, sdm, ftt, ads, asl, rbc, mbrt, dis, mbr, blc* }

$\boldsymbol{I}^{CF}$ – technologies with concentration factor

{*ftt, mbrt, sdm, mbr*}

$\boldsymbol{I}^{CONS}$ – technologies with consumables

{*ftt, ads, mbrt, mbr*}

$\boldsymbol{I}^{EAC}$ – technologies with externally added components

{*flc, aop, dis, blc*}

$\boldsymbol{I}^{BV}$ – technologies with binary variables

{*flc, sdm, ftt, ads, asl, rbc, mbrt, dis, mbr, blc, byp1, byp2, byp3, byp4* }

$\boldsymbol{I}^{S1}$ – technologies in stage 1

{*flc, byp1*}

$\boldsymbol{I}^{S2}$ – technologies in stage 2

{*ftt, sdm, byp2*}

$\boldsymbol{I}^{S3}$ – technologies in stage 3

{*ads, asl, rbc, dis, mbrt, byp3*}

$\boldsymbol{I}^{S4}$ – technologies in stage 4

{*aop, blc, mbr, byp4*}

Subsets for streams

$\boldsymbol{J}^{flc}$ – streams for flocculation

{2, 4, 5}

$\boldsymbol{J}^{byp1}$ – streams for bypass 1

{3, 6}

$\boldsymbol{J}^{sdm}$ – streams for sedimentation

{9, 13, 14}

$\boldsymbol{J}^{ftt}$ – streams for filtration

{8, 11, 12}

$\boldsymbol{J}^{byp2}$ – streams for bypass 2

{10, 15}

$\boldsymbol{J}^{ads}$ – streams for adsorption

{17, 23, 24}

$\boldsymbol{J}^{asl}$ – streams for activated sludge

{18, 25, 26}

$\boldsymbol{J}^{rbc}$ – streams for rotating biological containers

{19, 27, 28}

$\boldsymbol{J}^{dis}$ – streams for disinfection

{20, 29, 30, 31}

$\boldsymbol{J}^{mbrt}$ – streams for membrane bioreactor

{21, 32, 33}

$\boldsymbol{J}^{byp3}$ – streams for bypass 3

{22, 34}

$\boldsymbol{J}^{aop}$ – streams for advanced oxidation process

{36, 40, 41, 42}

$\boldsymbol{J}^{mbr}$ – streams for membrane processes

{38, 46, 47}

$\boldsymbol{J}^{blc}$ – streams for bleaching

{37, 43, 44, 45}

$\boldsymbol{J}^{byp4}$ – streams for bypass 4

{39, 48}

Subsets for components

$\boldsymbol{K}^{S}$ – components in streams

{*Wtr, Ssld, Mtls, Chem, API, Flcnt, Oz, NaClO, L_Chlrn*}

$\boldsymbol{K}^{SP}$ – components in initial wastewater stream

{*Wtr, Ssld, Mtls, Chem, API*}

$\boldsymbol{K}^{CONT}$ – components that are contaminants

{*Ssld, Mtls, Chem, API*}

$\boldsymbol{K}^{EAC}$ – components that are externally added

{*Flcnt, Oz, NaClO, L_Chlrn*}

A.2.3 Dynamic sets for connectivity

$\boldsymbol{J_i}$ - streams of technology *i*

$\boldsymbol{Jin_i}$ - inlet streams of technology *i*

$\boldsymbol{Jout_i}$ - outlet streams of technology *i*

$\boldsymbol{K_i}$ - components *k* in technology *i*

$\boldsymbol{K_j}$ - components *k* in stream *j*

A.2.4 Model Parameters

General Parameters

ρ_k (kg/m^3) = Density of component *k*

π_{WW} (m^3 WW/h) = Entering volumetric flowrate of wastewater (WW)

π^{Rep}_i ($/unit) = Replacement cost of consumables per unit capacity in technology *i*

μ (N-s/m^2) = viscosity of fluid

η_i (%) = efficiency of technology *i*

θ_i^R (hr) – residence time in technology *i*

θ_i^{Rep} (h/year) = Replacement time for consumables in technology *i*

τ_{ann} (h/annum) = annual operation in hours (330 day x 24 h/day = 7920 hrs)

$C0_i$ ($/capacity) = standard capacity cost in technology *i*

g (m/s^2) = gravitational constant

nc = cost scaling index (2/3 rule)

$Nlabr_i$ (#/h) = standard # of laborers required for technology *i* per hour

$Q0_i$ (m^3 or m^2 or m^3/h) = standard capacity of technology *i*

MW_k (kg/kmol) = molecular weight of component *k*

Min_k (kg/m3) = initial mass concentration of component *k*

$Cpur_k$ ($/kg) = purchase cost of added component *k*

Wsp_i (kW/h) = standard power required for technology *i* per hour

MM_k (--) = Big-M constant for component *k*

Φ_k (kg/kg WW) = amount of externally added component *k*

dp (m) = diameter of particle

$\xi_{j,i}$ (--) = retention factor of component *k* for technology *i* {*ftt*, *mbrt*, and *mbr* technologies}

ς_{RF} (--) = capital recovery/ charge factor (0.11)

ζ_i (m^3/m^2h) = flux of technology *i* { *ftt*, *mbrt*, and *mbr* technologies}

κd_i (h^{-1}) = decay of biomass coefficient of technology *i* {*asl* technology}

γ_i (kg/kg) = biomass yield of technology *i* { asl technology}

χ_i (m^3/m^2h) = hydraulic loading of technology *i* {rbc technology}

BMC_{mult} (--) = bare module cost multiplier (5.4)

C_{Lab} ($/h) = labor cost – operator basis (30)

C_{Elec} ($/kW) = cost of electricity per hour (0.1)

Evaluated Parameters

SOR_i (m/s) = surface overflow rate id sedimentation

U_i (m/s) = settling velocity of technology *i*

A.2.5 Model Variables

General Variables

$Cc_{,i}$ ($) = Purchase cost of technology $i \in \boldsymbol{I^{CST}}$

CF_i (m^3/m^3) = Concentration factor for technologies $i \in \boldsymbol{I^{CF}}$

Cpr_k ($/h) = Purchase cost of added components $k \in \boldsymbol{K^{EAC}}$

$M_{j,k}$ (kg/h) = Mass flowrate of component *k* in stream *j*

$Qc_{i'}$ (m^3 or m^2 or m^3/h) = capacity cost of technologies $i \in \boldsymbol{I^{CST}}$

PW_i (kW/h) = power requirements for technologies $i \in \boldsymbol{I}^{CST}$

$Nlbr_i$ (#/h) = number of laborers required for technology $i \in \boldsymbol{I}^{CST}$

Yo_i (kg/kg) = observed bacteria yield of technology *i* (*asl* technology)

Sr_i (m/h) = settling rate of unit *i* (*asl* technology)

X_i (kg/h) = biomass produced in technology *i* (*asl* technology)

Srt_i (h) = solids residence time in technology *i* (*asl* technology)

D_i (m) = diameter of technology *i* (*mbrt* technology)

$Cons_i$ ($/annum) = consumable cost of technology $i \in \boldsymbol{I}^{CONS}$

Binary Variables

y_i (--) = binary variables for technologies to selected $i \in \boldsymbol{I}^{BV}$

Stage-wise Costing Variables

$CCAC_{Nstg}$ = annualized capital (fixed) cost in nth stage

$CCRM_{Nstg}$ = material cost in nth stage

$CCCS_{Nstg}$ = consumable cost in nth stage

$CCLB_{Nstg}$ = labor cost in nth stage

$CCUT_{Nstg}$ = utility cost in nth stage

$CCOT_{Nstg}$ = other cost in nth stage (plant overhead and supervision costs)

$CCTC_{Nstg}$ = total cost in nth stage (all cost added in that particular stage)

CCTPC = total cost for process (summation of total cost in each stage)

A.2.6 Model Equations:

Initial wastewater flowrate equations:

$$M_{1,k} = (\sum_k Min_k)\pi_{WW} \; ; \; \forall \; k \in \boldsymbol{K^{SP}}$$

Component balances in all technologies:

$$\sum_{j \in \boldsymbol{Jin_i}} M_{j,k} = \sum_{j \in \boldsymbol{Jout_i}} M_{j,k} \; ; \; \forall \; k \in \boldsymbol{K^{S}}$$

Cost of technologies:

$$\left(\frac{C_{c_i}}{C0_i}\right) = \left(\frac{Q_{c_i}}{Q0_i}\right)^{nc} \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Labor requirements in technologies:

$$Nlbr_i Q0_i = Nlabr_i Qc_i \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Consumable costs in technologies:

$$Cons_i = \frac{\tau_{ann}}{\theta_i^{Rep}} \pi_i^{Rep} Qc_i \; ; \; \forall \; i \in \boldsymbol{I^{CST}}$$

Logical equations:

$$M_{i,j} - M1_k y_i \leq 0 \; ; \forall \; i \in \boldsymbol{I^{BV}} \; , \; j \in \boldsymbol{J}, \; k \in \boldsymbol{K_i} \; and \; \boldsymbol{K_j}$$

Selection of technologies in each stage:

Preliminary (Pretreatment) stage:

$$y_{flc} + y_{byp,1} = 1$$

Primary Treatment Stage:

$$y_{ftt} + y_{sdm} + y_{byp,2} = 1$$

Secondary Treatment Stage:

$$y_{ads} + y_{asl} + y_{rbc} + y_{dis} + y_{mbrt} + y_{byp,3} = 1$$

Tertiary Treatment Stage:

$$y_{aop} + y_{mbr} + y_{blc} = 1$$

Preliminary (Pretreatment) Stage Model Equations for Technologies

Flocculation (flc):

Flocculent added:

$$M_{5,Flcnt} = \Phi_{Flcnt} \sum_{k \in \boldsymbol{K^{CONT}}} M_{2,k}$$

Flocculent cost:

$$Cpur_k = \pi_{Flcnt} M_{5,Flcnt}$$

Volume of flocculation unit:

$$Qc_{flc} = \theta_{flc}^{R} \left[\sum_{k \in \boldsymbol{K^S}} \left(\frac{M_{2,k}}{\rho_k} \right) \right]$$

Power required in flocculation unit:

$$PW_{flc} = Wsp_{flc} Qc_{flc}$$

Primary Stage Model Equations for Technologies

Sedimentation (sdm):

Efficiency equation:

$$\eta_{sdm} = \frac{M_{13,k}}{M_{9,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Concentration factor (CF_{sdm}):

$$CF_{sdm} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{sdm}\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$2 \leq CF_{sdm} \leq 15$

Written as: $CF_{sdm} \leq 15 y_{sdm}$ *and* $CF_{sdm} \geq 2 y_{sdm}$

Area of sedimentation unit:

$$Qc_{sdm} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]}{SOR_{sdm}}$$

Power required in sedimentation unit:

$$PW_{sdm} = Wsp_{sdm} Qc_{sdm}$$

Filtration (ftt):

Retention factor equation

$$\xi_{ftt_k} = \frac{M_{11,k}}{M_{8,k}} \; ; \; k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{ftt_k} M_{8,k} = M_{11,k}$

Concentration factor equation (ftt):

$$CF_{ftt} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{11,k}}{\rho_k}\right)\right]}$$

Written as: $CF_{ftt}\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{13,k}}{\rho_k}\right)\right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{9,k}}{\rho_k}\right)\right]$

$1 \leq CF_{ftt} \leq 30$

Written as: $CF_{ftt} \leq 30 y_{ftt}$ *and* $CF_{ftt} \geq 1 y_{ftt}$

Area of filtration unit (flux balance):

$$Qc_{ftt} = A_{ftt} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)}{\zeta_{ftt} CF_{ftt}}$$

Written as: $Qc_{ftt}\zeta_{ftt}CF_{ftt} = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{8,k}}{\rho_k}\right)\right](CF_{ftt}-1)$

Power requirements for filtration unit:

$$PW_{ftt} = Wsp_{ftt} Qc_{ftt}$$

Power required in sedimentation unit:

$$PW_{sdm} = Wsp_{sdm} Qc_{sdm}$$

Tertiary Stage Model Equations for Technologies

Advanced Oxidation Processes:

Mass of ozone needed for advanced oxidation processes unit

$$M_{42,Oz} = \Phi_{Oz} \sum_{k \in \boldsymbol{K^{CONT}}} M_{31,k}$$

Efficiency equation:

$$\eta_{aop} = \frac{M_{40,k}}{M_{36,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Volume of advanced oxidation processes unit:

$$Qc_{aop} = V_{aop} = \theta_{aop}^{R} \left[\sum_{k \in \boldsymbol{K^{S}}} \left(\frac{M_{36,k}}{\rho_k} \right) \right]$$

Power required for advanced oxidation processes unit:

$$PW_{aop} = Wsp_{aop} Qc_{aop}$$

Membrane Processes:

Retention factor equation

$$\xi_{mbr_k} = \frac{M_{46,k}}{M_{38,k}} \; ; \; k \in \boldsymbol{K^{SP}}$$

Written as: $\xi_{mbr_k} M_{38,k} = M_{46,k}$

Concentration factor equation (mbr):

$$CF_{mbr} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k} \right) \right]}{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{46,k}}{\rho_k} \right) \right]}$$

Written as: $CF_{mbr} \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{46,k}}{\rho_k} \right) \right] = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k} \right) \right]$

$1 \leq CF_{mbr} \leq 35$

Written as: $CF_{mbr} \leq 35 y_{mbr} \; and \; CF_{mbr} \geq 1 y_{mbr}$

Area of membrane processes unit (flux balance):

$$Qc_{mbr} = A_{mbr} = \frac{\left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right](CF_{mbr}-1)}{\zeta_{mbr} CF_{mbr}}$$

Written as: $Qc_{mbr}\zeta_{mbr}CF_{mbr} = \left[\sum_{k \in \boldsymbol{K^{SP}}} \left(\frac{M_{38,k}}{\rho_k}\right)\right](CF_{mbr}-1)$

Power requirements for membrane unit:

$$PW_{mbr} = Wsp_{mbr} Qc_{mbr}$$

<u>*Bleaching:*</u>

Efficiency equation:

$$\eta_{blc} = \frac{M_{43,k}}{M_{37,k}} \; ; \; k \in \boldsymbol{K^{CONT}}$$

Mass of disinfectant required for disinfection unit

$$M_{45,NaClO} = \Phi_{NaClO} \sum_{k \in \boldsymbol{K^{CONT}}} M_{37,k}$$

Volume of disinfection unit:

$$Qc_{dis} = V_{dis} = \theta_{dis}^{R} \left[\sum_{k \in \boldsymbol{K^{S}}} \left(\frac{M_{37,k}}{\rho_k}\right)\right]$$

Power required for disinfection unit:

$$PW_{dis} = Wsp_{dis} Qc_{dis}$$

Cost Model Equations

Stagewise Cost Equations

Annualized capital cost in each stage:

$$CCAC_{Nstg} = 1.66 \, \varsigma_{RF} BMC_{mult} \sum_{i \in istg\{1,2,3,4\}} C_{ei}$$

Material Cost:

$$CCRM_{s1} = [\tau_{ann}(Cpur_{Flcnt})]$$

$$CCRM_{s2} = 0$$

$$CCRM_{s3} = [\tau_{ann}(Cpur_{L_{Chlrn}})]$$

$$CCRM_{s4} = [\tau_{ann}(Cpur_{Oz} + Cpur_{NaClO})]$$

Consumable Cost:

$$CCCS_{s1} = 0$$

$$CCCS_{s2} = Cons_{ftt}$$

$$CCCS_{s3} = Cons_{ads} + Cons_{mbrt}$$

$$CCCS_{s4} = Cons_{mbr}$$

Labor Cost

$$CCLC_{Nstg} = \tau_{ann} C_{Lab} \sum_{i \in istg\{1,2,3,4\}} Nlb_i$$

Utility Cost

$$CCUC_{Nstg} = \tau_{ann} C_{elec} \sum_{i \in istg\{1,2,3,4\}} PW_i$$

Other Cost

$$CCOC_{Nstg} = 2.78\, \tau_{ann} C_{Lab} \sum_{i \in istg\{1,2,3,4\}} Nlb_i$$

Total Cost in each Stage

$$CCTC_{Nstg} = CCAC_{Nstg} + CCRM_{Nstg} + CCCS_{Nstg} + CCLC_{Nstg} + CCUC_{Nstg} CCOC_{Nstg}$$

Total Category Cost:

$$CCTAC = \sum_{n \in Nstg\{1,2,3,4\}} CCAC_n$$

$$CCTRM = \sum_{n \in Nstg\{1,2,3,4\}} CCRM_n$$

$$CCTCS = \sum_{n \in Nstg\{1,2,3,4\}} CCCS_n$$

$$CCTLC = \sum_{n \in Nstg\{1,2,3,4\}} CCLC_n$$

$$CCTUC = \sum_{n \in Nstg\{1,2,3,4\}} CCUC_n$$

$$CCTOC = \sum_{n \in Nstg\{1,2,3,4\}} CCOC_n$$

$$CCTPC = CCTAC + CCTRM + CCTCS + CCTLC + CCTUC + CCTOC$$

Objective Function 1:

$$Obj = Min\, CCTPC$$

A.2.7 Model parameters and inputs

Table A15

Density and Molecular Weight of Components

Component	**Value (**kg/m^3**)**	**Value (**kg/kmol**)**
Water	1000	18
Solid particles	1800	102
Metals	2500	98
Chemicals	1100	48
Active pharmaceutical ingredient	1400	748.996
Ozone	2.14	48
Sodium Hypochlorite	1110	74.44
Flocculant	2200	2200
Liquid chlorine	1470	70.91

Table A16

Purchase Cost of Added Components

Component	**Value ($/**kg**)**
Ozone	3.53
Sodium Hypochlorite	0.35
Flocculant	0.75
Liquid chlorine	2

Table A17

Standard Capacity, Cost, Laborers, and Power for Technologies

Unit Operation (costing capacity)	Standard Capacity (Units)	Standard Capacity Cost (million $)	Standard Laborers Required (#/h)	Standard Power Required (kW/h)
Flocculation (Volume)	2000 (m^3)	0.538	0.1	0.0002
Filtration (Area)	80 (m^2)	0.039	0.5	0.1
Sedimentation (Volume)	2500 (m^3)	1.128	0.1	0.0002
Adsorption (Volume)	100 (m^3)	0.12	0.2	0.3
Membrane Bioreactor (Area)	40 (m^2)	1.194	0.1	0.2
Rotational Biological Container (Area)	185 (m^2)	0.045	0.3	0.01
Activated Sludge (Volume)	250 (m^3)	0.569	0.4	0.2
Disinfection (Volume)	540 (m^3)	0.627	0.7	0.5
Membrane Processes (Area)	80 (m^2)	0.938	0.5	0.2
Advanced Oxidation Processes (Volume)	1000 (m^3)	0.735	0.1	0.5
Bleaching (Volume)	500 (m^3)	0.100	0.5	0.33

Table A18

Replacement Time for Technologies with Consumables

Technology	Value (h/yr)
Filtration	2000
Adsorption	720
Membrane Processes	1000
Membrane Bioreactor	1000

Table A19

Replacement Cost for Technologies with Consumables

Technology	Value ($/Unit)
Filtration	100
Adsorption	74.16
Membrane Processes	400
Membrane Bioreactor	400

A.2.7.1 Flocculation (Flc):

Flocculent added (kg/kg) – 0.005
Residence time (h) – 0.5

A.2.7.2 Sedimentation (Sdm):

Efficiency – 80%
Depth – 3m

A.2.7.3 Filtration (Ftt):

Retention factor: Water – 0.05, Solids – 0.80, Metals – 0.10, Chemicals – 0.05, API – 0.50
Flux ($m^3/m^2.h$): 0.2

A.2.7.4 Membrane Bioreactor (Mbrt):

Retention factor: Water – 0.005, Solids – 0.95, Metals – 0.85, Chemicals – 0.50, API – 0.90
Flux ($m^3/m^2.h$): 0.025

A.2.7.5 Adsorption (Ads):

Empty bed contact time (h): 0.25
Density of granular activated carbon (GAC) (kg/m^3): 1030
Efficiency: 90%
Void fraction of GAC: 50%

A.2.7.6 Activated Sludge (Asl):

Decay of biomass coefficient (h^{-1}): 0.0021
Biomass yield (kg/kg): 0.5
Hydraulic retention time (h): 2
Efficiency: 80%

A.2.7.7 Rotating Biological Contactors (Rbc):

Efficiency: 80%
Hydraulic loading (m^3/m^2h): 20

A.2.7.8 Disinfection (Dis):

Efficiency: 95%
Ratio of liquid chlorine to contaminant (kg/kg): 0.00173
Residence time (h): 2

A.2.7.9 Advanced Oxidation Processes (Aop):

Ratio of ozone to contaminant (kg/kg): 0.000173
Efficiency: 98%
Residence time (h): 0.21

A.2.7.10 Membrane Processes (Mbr):

Retention factor: Water – 0.05, Solids – 0.90, Metals – 0.90, Chemicals – 0.95, API – 0.95
Flux ($m^3/m^2.h$): 0.0856

A.2.7.11 Bleaching (blc):

Efficiency: 98%
Ratio of sodium hypochlorite to contaminant (kg/kg): 0.00173
Residence time (h): 2

A.2.8 Sustainable Process Index (SPI) Calculations

A.2.8.1 SPI for Pharmaceutical Wastewater Treatment Using Values from GAMS Code

Parameters

F_{RR} (kg/yr) - feed of a processed resource (752400000)
y_R ($kg/m^2.yr$) - specific yield (243.1542)
$y_{EI,RN}$ ($kWh/m^2.yr$) - mean industrial energy yield or mean industrial energy supply density (7)
F_{RN} (kg/yr) - feed of a processed resource (215463.6) [from GAMS Code]
C_N ($/kg) - price of the material (world market price, taxes excluded) (7)
C_E ($/kWh) - price of one kilowatt-hour of energy (industrial price, taxes excluded) (0.1)
F_E (kWh/yr) - energy used in the process (450584.64) [from GAMS code]
y_E ($kWh/m^2.yr$) - energy yield (43)
C_I ($) - total cost of energy for indirect installation (1329500) [from GAMS code]
LS (yr) - depreciated area over the life-span years (30)
$y_{EI,II}$ ($kWh/m^2.yr$) - industrial energy supply density or yield (43)
N_S (cap/yr) - total number of workers in the treatment plant (80.19)

y_S (cap/m^2.yr) - yield factor due to staff (3.59E-05)

$c_{c,m}$ (kg$_m$/kg) - allowable concentration of substance, *m* [Solids, Chemical, Metals, Water] in the compartment, *c* [air, water, soil]

R_c (kg/m^2.yr) - rate of renewal of the environmental compartments, c [air, water, soil]

Estimated Parameters

$E_{D,RN}$ - energy demand to supply one kilogram of the material in question for non-renewable energy (kWh/kg)

$$E_{D,RN} = \frac{0.95\ C_N}{C_E} = \frac{0.95\ \times 7}{0.1} = 66.5$$

$E_{D,II}$ - energy demand to supply one kilogram of the material in question for indirect land energy (kWh/yr)

$$E_{D,II} = \frac{0.54\ C_I}{C_E . LS} = \frac{0.54\ \times 1329500}{0.1 \times 30} = 239310$$

$S_{c,m}$ - dissipation to potential sink (kg$_m$/m^{2}yr)

$$S_{c,m} = R_c\ .\ c_{c,m}$$

Analysis for Rsoil (kg/m^2.yr)

Rate of soil renewal (RSN) in the US is 2.2E-04 m/yr

Assuming the soil is loamy with a 50% pore space, then the bulk density is 1330 kg/m3

$$R_{soil} = RSN \times Den_{bulk} = 0.00022 \times 1330 = 0.2926$$

Since we categorized contaminants into of solids, chemicals, and metals, we used contaminants that had the smallest allowable concentration for each category in the compartments to estimate S. For solids contaminant, we used lead (Pb), for chemical we used Chromium (Cr), and for metals we used lead (Pb).

Analysis for Ssoil,m (kgm/m^2.yr)

$S_{soil,m} = R_{soil} \cdot c_{soil,m}$

Table A20

Data on Soil Yield for Pharmaceutical Wastewater Contaminants

Component (m)	**$C_{(soil,m)}$ (kgm/kg)**	**$S_{(soil,m)}$ (kgm/m^2.y)**
Solids (Sslds)	1.00E+00	2.93E-01
Chemical (Chem)	1.00E-04	2.93E-05
Metals (Mtls)	2.00E-06	5.85E-07
Water (Wtr)	1.00E+00	2.93E-01
API	1.00E-04	2.93E-05

Analysis for Rwater (kg/m^2.yr)

Average rate of precipitation (RP) form Jan, 2009 to Dec, 2019 in the US is 31.91 in/yr (0.810514 m/yr)

Seeping ratio (SR) of water is 0.30

$R_{water} = RP \times SR \times Den_{water} = 0.810514 \times 0.30 \times 1000 = 243.1542$

Analysis for Swater,m (kgm/m^2.yr)

Table A21

Data on Water Yield for Pharmaceutical Wastewater Contaminants

Component (m)	**$C_{(water,m)}$ (kg_m/kg)**	**$S_{(water,m)}$ ($kg_m/m^2.y$)**
Solids (Sslds)	1.67E-06	4.05E-04
Chemical (Chem)	1.00E-07	2.43E-05
Metals (Mtls)	1.50E-08	3.65E-06
Water (Wtr)	1.00E+00	2.43E+02
API	1.00E-07	2.43E-05

$S_{air,chem}$ ($kgm/m^2.yr$) = 6.50E-03

Table A22

$F_{P,c,m}$ Product Flow Rate to Compartment c, by Substance, M (kg_m/yr) [From GAMS Code] for Pharmaceutical Wastewater Treatment

Component (m)	**Fraction of m into compartment c**			**$F_{P,c,m}$**		
	Air	Water	Soil	Air	Water	Soil
Solids (Sslds)	0	0.05	0.95	0.00E+00	7.92E+03	1.50E+05
Chemical (Chem)	0.03	0.9	0.07	9.93E+04	2.98E+06	2.32E+05
Metals (Mtls)	0	0.7	0.3	0.00E+00	4.99E+02	2.14E+02
Water (Wtr)	0	0.95	0.05	0.00E+00	7.15E+08	3.76E+07
API	0	0.5	0.5	0.00E+00	7.92E+03	7.92E+03

Variables

A_R - area for raw material production (m^2/yr)

A_{RR} - area for renewable raw material production (m^2/yr)

A_E - area for energy production (m^2/yr)

A_I - area for installation for equipment and other infrastructure (m^2/yr)

A_{ID} - area for direct installation (m^2/yr)

A_{II} - area for indirect installation (m^2/yr)

A_S - area for staff (m^2/yr)

$A_{P,c,m}$ - area for dissipating a single component of particular product flow to a given compartment (m^2/yr)

$A_{PS,c}$ - area assined to the dissipation of a particular product stream, S (m^2/yr)

A_P - area for product dissipation (m^2/yr)

A_{tot} - total area (m^2/m^3WW-yr)

Area for Raw Material Production (A_R)

Area for Renewable Raw Material Production (A_{RR})

$$A_{RR} = \frac{F_{RR}}{y_R} = \frac{7524000000}{243.1542} = 3.09E06$$

Area for Non-Renewable Raw Material Production (A_{RN})

$$A_{RN} = \frac{F_{RN}.E_{D,RN}}{y_{EI,RN}} = \frac{0 \times 66.5}{7} = 0$$

$$A_R = A_{RR} + A_{RN} = 3.09\text{E}06$$

Area for Energy Production (A_E)

$$A_E = \frac{F_E}{y_E} = \frac{450584.64}{43} = 1.05E4$$

Area for Installations (A_I)

$$A_I = A_{ID} + A_{II}$$

Area for Direct Installation (A_{ID})[from GAMS Code]

Flocculation technology = 17.43

Filtration technology = 494.63

Adsorption technology = 4.95

$A_{ID} = 517.00$

Area for Indirect Installation (A_{II})

$$A_{II} = \frac{E_{D,II}}{y_{EI,II}} = \frac{239310}{43} = 5.57E3$$

Area for Staff (A_S)

$$A_S = N_S . a_{in} = \frac{N_S}{y_S} = \frac{1037.19}{0.000035} = 2.89E7$$

Area for Product Dissipation into Various Environmental Compartment (A_P)

$$A_{P,c,m} = \frac{F_{P,c,m}}{S_{c,m}}$$

Table A23

Area Needed for Emission Dissipation for Pharmaceutical Wastewater Treatment

Component (m)	**$A_{P,c,m}$**		
	Air	Water	Soil
Solids (Sslds)	--	1.95E+07	5.14E+05
Chemical (Chem)	1.53E+07	1.23E+11	7.92E+09
Metals (Mtls)	--	1.37E+08	3.65E+08
Water (Wtr)	--	2.94E+06	1.29E+08
API	--	3.26E+08	2.71E+08

$$A_{PS,c} = max_m(A_{P,c,m})$$

Highlighted are the maximum values for each component.

Table A24

Highest Area Needed for Emission Dissipation for Pharmaceutical Wastewater Treatment

Component (m)	$A_{P,c,m}$		
	Air	Water	Soil
Solids (Sslds)	--	1.95E+07	5.14E+05
Chemical (Chem)	1.53E+07	1.23E+11	7.92E+09
Metals (Mtls)	--	1.37E+08	3.65E+08
Water (Wtr)	--	2.94E+06	1.29E+08
API	--	3.26E+08	2.71E+08

$$A_P = \sum_c A_{PS,c} = 1.30E11$$

$A_{tot} = A_R + A_E + A_I + A_S + A_P$ = 1.30E+11 m^2/yr =1.65E+05 (m^2/m^3WW-yr)

A.2.8.2 SPI for Direct Disposal of Pharmaceutical Wastewater

We considered on the area needed to sustainably embed the wastewater into the ecosystem, i.e. A_P

Area for Product Dissipation into Various Environmental Compartment (A_P)

Table A25

$F_{P,c,m}$ Product Flow Rate to Compartment c, By Substance, M (kg_m/yr) [From GAMS Code] for Direct Disposal of Pharmaceutical Wastewater

Component (m)	Fraction of		$F_{P,c,m}$	
	Water	Soil	Water	Soil
Solids (Sslds)	0.05	0.95	4.38E+05	8.32E+06
Chemical (Chem)	0.95	0.05	3.66E+07	1.93E+06
Metals (Mtls)	0.7	0.3	6.13E+03	2.63E+03
Water (Wtr)	0.95	0.05	8.32E+08	4.38E+07
API	0.5	0.5	1.75E+05	1.75E+05

$$A_{P,c,m} = \frac{F_{P,c,m}}{S_{c,m}}$$

Table A26

Area Needed for Emission Dissipation for Direct Disposal Of Pharmaceutical Wastewater

Component (m)	$A_{P,c,m}$	
	Water	Soil
Solids (Sslds)	1.08E+09	2.84E+07
Chemical (Chem)	1.51E+12	6.59E+10
Metals (Mtls)	1.68E+09	4.49E+09
Water (Wtr)	3.42E+06	1.50E+08
API	7.21E+09	5.99E+09

$A_{PS,c} = max_m(A_{P,c,m})$

Highlighted are the maximum values for each component.

Table A27

Highest Area Needed for Emission Dissipation for Direct Disposal of Pharmaceutical Wastewater

Component (m)	**$A_{P,c,m}$**	
	Water	Soil
Solids (Sslds)	1.08E+09	2.84E+07
Chemical (Chem)	1.51E+12	6.59E+10
Metals (Mtls)	1.68E+09	4.49E+09
Water (Wtr)	3.42E+06	1.50E+08
API	7.21E+09	5.99E+09

$$A_P = \sum_c A_{PS,c} = 1.57E12$$

$A_{tot} = A_P$ = 1.57E+12 m^2/yr = 1.79E+06 (m^2/m^3WW-yr)

Appendix B

Supporting Information for Chapter 4

B.1 Model Equations (Applicable to all technologies)

B.1.1 Indices and Sets

i ∈ ***I*** – technologies (used as subscript to variables)

{UF-Ultrafiltration,
PVP-Pervaporation,
DST-Distillation,
SDM-Sedimentation,
DRY- Dryer,
ATPE- Aqueous Two-Phase Extraction,
CNF- Centrifugation,
FLT- Filter,
INCN- Incineration}

$j \in \boldsymbol{J}$ – stream (used as subscript to variables)

{1, 2, 3, 4 …}

$k \in \boldsymbol{K}$ – components (used as subscript to variables)

{IPA – isopropyl alcohol,
Wtr – water,
Salt1 – sodium chloride
Salt2 – sodium sulfate anhydrous
Hex--Hexane – hexane,
DME – dimethoxyethane,
EME – 1-ethoxy-1-methoxyethane,
Tol – toluene}

$c \in \boldsymbol{C}$ – the various compartment of the ecosystem

{Air_C,
Water_C,

Soil_C}

$e \in \boldsymbol{E}$ – emissions associated with the recovery process

{CO2 – carbon dioxide

CO – carbon monoxide

CH4 – methane

NOx – nitrogen oxides

NMVOC – non-methane volatile organic compounds

Others – other miscellaneous air emissions

WE – water emissions

SE – soil emissions}

B.1.2 Subsets

Subsets for Technologies

I^{CST} – technologies with costs

{UF, PVP, DST, SDM, DRY, ATPE, CNF, FLT, INCN}

I^{CF} – technologies with concentration factor

{PVP, UF, FLT, SDM, CNF}

I^{CONS} – technologies with consumables

{ATPE, PVP, FLT, UF}

Subsets for Components

Jbp_{ATPE} – bottom phase of ATPE

Jda_{DRY} – dry air inlet stream to DRY

$Jliq_{CNF}$ – stream containing no solids leaving CNF

Jin_i – inlet streams of technology i

$Jout_i$ – outlet streams of technology i

$Jpoly_{ATPE}$ – polymer feed stream to ATPE

$Jsalt_{ATPE}$ – salt feed stream to ATPE

$Jsld_{CNF}$ – stream containing solids leaving CNF

Jtp_{ATPE} – top phase of ATPE

K_i – components *k* in technology *i*

K_j – components *k* in stream *j*

K^{JP} – components in process streams

{IPA, Wtr, Salt1, Salt 2, Hex, DME, EME, Tol}

B.1.3 General Parameters

ρ (kg/m^3) = Density of component k

π_{feed} ($/kg biomass) = Entering feed cost in terms of per kg waste

π^{Rep}_i ($/unit) = Replacement cost of consumables per unit capacity in technology *i*

λ_{stm} (kJ/kg) = Latent heat of steam

$\lambda_{vap,k}$ (kJ/kg) = Heat of vaporization of component k

α_k = Relative volatility of component k for technology i

μ (N-s/m^2) = viscosity of fluid

η_{stage} = stage efficiency

θ_i^R (hr) – residence time in technology *i*

θ_i^{Rep} (h/year) = Replacement time for consumables in technology *i*

τ_{ann} (h/annum) = (330 days x 24 h/day = 7920 hours)

$C0_i$ ($/capacity) = Cost of a technology with standard capacity

Cp (KJ/kg-℃) = Specific heat of component k

$D_{p,SDM}$ = particle diameter in sedimentation unit

g (m/s^2) = gravitational constant

nc = cost scaling index (2/3 rule)

$Nlabr_i$ (#/h) = # of laborers required for technology *i* per hour

$Q0_i$ (m^3 or m^2 or m^3/h) = Standard capacity of a technology for costing, labor and power required

T_{amb} (℃) = ambient temperature

Tcw_i (℃) = Cooling water temperature in (25)

Tcw_o (℃) = Cooling water temperature out (30)

T_{sat} (℃) = saturation temp

B.1.4 Evaluated Parameters:

SOR_i (m/s) = surface overflow rate in sedimentation

U_i (m/s) = settling velocity of technology i

B.1.5 General Variables

B_i = volume ratio of equipment i

$Cc_{,i}$ ($) = Purchase cost of unit i

CF_i (m^3/m^3) = Concentration factor for technologies $i \in \boldsymbol{I^{CF}}$

$Cpur_k$ ($/h) = Purchase cost of added components ($k \in \mathbf{K^{ADD}}$)

D_i(m) = diameter of technology unit i

L_i(m) = length of technology unit i

Liq_{DST} = liquid molar flowrate in distillation column

$M_{j,k}$ (kg/h) = Mass flowrate of component k in stream j

N = actual number of stages

N_{min} = minimum number of stages

q = quality of mixture (for distillation, entering feed quality)

Qc, (m^3 or m^2 or m^3/h) = Costing variable for technologies $i \in \boldsymbol{I^{CST}}$

QC_{DST} = cooling requirement for distillation unit

QH_{DST} = heat duty for distillation unit

Qs_{DST} = heat required to bring the feed to saturation

PW_i (kW/h) = Power required for technologies $i \in \boldsymbol{I}^{CST}$

R_{min} = minimum reflux ratio

R = actual reflux ratio

U_v = Underwood variable

Vap_{DST} = vapor molar flowrate in distillation column

Wsp_i (kW/h) = Power required by technology i per hour

$X_{mj,k}$ = mole fraction of component k in stream j

B.1.6 General Equations

Component balances:

$$\sum_{j \in jin_i} M_{j,k} = \sum_{j \in jout_i} M_{j,k} \; ; \forall k \in K^{JP} \; and \; i \in I$$

Cost of technologies:

$$\left(\frac{C_{ci}}{C0_i}\right) = \left(\frac{Q_{ci}}{Q0_i}\right)^{nc} ; \forall i \in I^{CST}$$

Labor requirements of technologies:

$$Nlb_i Q0_i = Nlabr_i Qc_i; \forall i \in I^{CST}$$

Consumable costs:

$$Cons_i = \frac{\tau_{ann}}{\theta_i^{Rep}} \pi_i^{Rep} Qc_i; \forall i \in I^{CST}$$

Annualized Capital Cost:

$$CCAC = \frac{(1.66 * CRF * BMC * \sum_i Cc_i)}{10^3}$$

Labor Cost:

$$CCLB = \frac{(C_{lbr} * \tau_{ann} * \sum_i Nlbr_i)}{10^3}$$

Utility Cost:

$$CCUC = \frac{\left((\sum_i PW_i * C_{elec} + \sum_i Mstm_i * C_{stm}) * \tau_{ann}\right)}{10^3}$$

Membrane Cost:

$$CCMC = \frac{(\tau_{ann} * \sum_{i1} CPM_{i1})}{Rep_{time} * 10^3}$$

Other Cost:

$$CCTC = 2.78 * CCLB$$

Total Cost:

$$CCTC = CCAC + CCUC + CCMC + CCOC + CCLB$$

B.1.7 Sustainable Process Index (SPI) Equations

Parameters for SPI

γ_{RR} (kg/m^2.yr) = specific yield for renewable resource

γ_{RN} (kWh/m^2.yr) = specific yield for mean industrial energy supply to provide non-renewable resource

γ_E (kWh/m^2.yr) = specific energy yield for the process

γ_{EI} (m^2/#person) = specific energy yield for indirect installations

γ_S (labor/m^2.yr) = specific labor yield

C_{np} ($/kg) = price of material (world market price)

$\beta_{e,k}$ (kg/kg) = emission of component *e* per component *k*

$\delta_{e,C}$ (--) = fraction of emission component e, dissipated into compartment C

$\gamma_{P(e,C)}$ (kg/m^2.yr) = yearly allowable yield of component *e*, in compartment *C*

Variables for SPI

$E_{m(e,k)}$ (kg$_e$/yr) = emission due to components in process stream

$E_{m(e)}$ (kg$_e$/yr) = emissions due to steam or electricity usage

T_{Tot} (kg$_e$/yr) = total emission of component *e*

A_{RR} (km^2) = area for renewable raw material production

A_{RN} (km^2) = area for non-renewable raw material production

A_R (km^2)= total area for raw material production

A_E (km^2)= area for energy supply

$A_{C,\,i}$ (m^2)= area for direct installation of technology *i*

A_{TC} (m^2) = total area for direct installation of equipment

$E_{D,II}$ (kWh/yr) = energy demand to supply one kilogram of the material in question from indirect land use

A_{II} (m^2) = area for indirect installations

A_I (km^2) = total area for installations

A_S (km^2) = area needed to accommodate staff

$A_{P(e,C)}$ (km^2) = area needed for the dissipation of component *e*, into compartment *C*

A_{Ptot} (km^2) = total area needed for the dissipation of component *e*, into compartment *C*

A_{TOT} (km^2) = total area needed needed for the entire process

Equations:

Total Area for raw material usage:

A_R is an area that indicates the environmental pressures exerted by the provision of raw materials for the process in question. This raw material area comprises both non-renewable and renewable. Depending on the mass flow rate (kg/yr) and the annual specific yield per square meter (kg/m^2.yr) of the renewable raw material (e.g., corn, barley), the area required can be estimated. Furthermore, various raw material alternatives can be compared based on area requirements. The analysis for the area required for non-renewable consumption follows the same analogy by considering the consumption of minerals and fossils. Dividing the flow rate of fossil and mineral materials by their specific yields give the respective areas.

$$A_{RR} = \frac{\tau_{ann} * M_{1,RR}}{\gamma_{RR} * 10^6} = \frac{\left(\frac{hr}{yr}\right)\left(\frac{kg}{hr}\right)}{\left(\frac{kg}{m^2 yr}\right)\left(\frac{m^2}{km^2}\right)} \; [=] \; km^2$$

$$A_{RN} = \frac{\tau_{ann} * M_{1,RN}}{\gamma_{RN} * 10^6} = \frac{\left(\frac{hr}{yr}\right)\left(\frac{kg}{hr}\right)}{\left(\frac{kg}{m^2 yr}\right)\left(\frac{m^2}{km^2}\right)} \; [=] \; km^2$$

$$A_R = A_{RR} + A_{RN}$$

Total area for energy use:

A_E, the energy area, is estimated based on the electricity requirements of the recovery process. It accounts for the environmental pressure caused by energy provision to the process. We estimated this area by dividing the annual electricity demand (kWh/yr) of the process by the average specific yield of electricity (kWh/m^2.yr).

$$A_E = \frac{\tau_{ann} \sum_i PW_i}{\gamma_E * 10^6} = \frac{\left(\frac{kWh}{hr}\right)\left(\frac{hr}{yr}\right)}{\left(\frac{kWh}{m^2 yr}\right)\left(\frac{m^2}{km^2}\right)} [=] km^2$$

Total area for installation:

A_I, the installation area, comprises the areas needed for the direct and indirect installation of equipment. The direct installation area comes directly from the capacity of the technologies. The indirect area is associated with the environmental pressures of the piping and other instrumentation installations. We calculated the indirect area by dividing the indirect installation cost ($/yr) by the unit cost of electricity ($/kWh), which estimated the energy demand (kWh/yr). This energy demand was further divided by the specific yield of electricity to obtain the area. A factor of 0.5 – 0.6 of the annualized cost is sufficiently precise in estimating the indirect installation cost for most industrial processes.

$$A_{TC} = \sum_i A_{C_i} [=] m^2$$

$$E_{D_{II}} = \frac{0.54 * CCAC}{\mathrm{C}_{elec}} = \frac{\left(\frac{\$}{yr}\right)}{\left(\frac{\$}{kWh}\right)} [=] \left(\frac{kWh}{yr}\right)$$

$$A_{II} = \frac{E_{D_{II}}}{\gamma_E} = \frac{\left(\frac{kWh}{yr}\right)}{\left(\frac{kWh}{m^2 yr}\right)} [=] m^2$$

$$A_I = \frac{A_{TC} + A_{II}}{10^6} = \frac{m^2 + m^2}{\left(\frac{m^2}{km^2}\right)} [=] km^2$$

Area for staff:

A_S, the area needed to accommodate the working staff, is a function of the total number of employees. The working area was obtained by multiplying the number of employees (person) by the specific arable land per person within the United States (m^2/person).

$$A_S = \frac{\left(\frac{24}{8}\right) * \sum_i Nlb_i * \gamma_S}{10^6} = \frac{\left(\frac{hr}{hr}\right)\left(\frac{\#person - hr}{hr}\right)\left(\frac{m^2}{\#person}\right)}{\left(\frac{m^2}{km^2}\right)} [=] km^2$$

Total Area Needed to Embed Emissions:

A_P constitutes the area needed to embed emissions into the biosphere's air, water, and soil compartments. We estimated the annual emissions (kg_e/yr) due to the process by multiplying the emission factors (kg_e/kg_c) and the component flow rates of the waste streams (kg_c/yr). Other emissions that were considered were from steam and electricity usage. We used the rate of renewal of the various compartments to estimate the specific yield ($kg_e/m^2.yr$) of emission into the biosphere. This was achieved by multiplying the rate of renewal ($kg/m^2.yr$) and the allowable concentration of emissions (kg_e/kg) into the respective compartments. For the water compartment, we used the seeping rate to groundwater bodies. This value is usually a fraction (0.3 – 0.5) of the precipitation rate per square meter ($kg/m^2.yr$). For the soil compartment, we used the annual rate of topsoil renewal through composting of 1 m^2 of biomass ($kg/m^2.yr$) while the natural emissions of relevant gases by forests per square meter ($kg/m^2.yr$) was used for the air compartment. Finally, the area was estimated by dividing the annual emissions by their respective specific yields. SimaPro® software was used to estimate the emission factors (LCIs).

Annual emissions from process stream:

$$E_{m\_PS_{(e,k)}} = M_{18,k} * \beta_{e,k} * \tau_{ann} = \left(\frac{kg}{hr}\right)\left(\frac{kg_e}{kg}\right)\left(\frac{hr}{yr}\right)[=]\left(\frac{kg_e}{yr}\right)$$

$$E_{m\_PS_{(e,k)}} = M_{29,k} * \beta_{e,k} * \tau_{ann} = \left(\frac{kg}{hr}\right)\left(\frac{kg_e}{kg}\right)\left(\frac{hr}{yr}\right)[=]\left(\frac{kg_e}{yr}\right)$$

Annual emissions from steam usage:

$$E_{m\_St_{(e)}} = M_{stm} * \beta_{e,stm} * \tau_{ann} = \left(\frac{kg}{hr}\right)\left(\frac{kg_e}{kg}\right)\left(\frac{hr}{yr}\right)[=]\left(\frac{kg_e}{yr}\right)$$

Annual emissions from electricity usage:

$$E_{m\_El_{(e)}} = \sum_i PW_i * \beta_{e,El} * \tau_{ann} = \left(\frac{kWh}{hr}\right)\left(\frac{kg_e}{kWh}\right)\left(\frac{hr}{yr}\right)[=]\left(\frac{kg_e}{yr}\right)$$

Total Emissions for each component:

$$T_{Tot_{(e)}} = \sum_{k} E_{m\_PS_{(e,k)}} + E_{m\_St_{(e)}} + E_{m\_El_{(e)}} [=] \left(\frac{kg_e}{yr}\right)$$

Area needed for each component emitted:

$$A_{P_{e,C}} = \frac{T_{Tot_{(e)}} * \delta_{e,C}}{\gamma_{P_{e,C}} * 10^6} = \frac{\left(\frac{kg_e}{yr}\right)\left(\frac{kg_e}{kg_e}\right)}{\left(\frac{kg_e}{m^2 yr}\right)\left(\frac{m^2}{km^2}\right)} [=] km^2$$

Total Area needed for each component emitted:

$$A_P = \sum_{e,C} A_{P_{e,C}} [=] km^2$$

Total Area (SPI):

$$A_{Tot} = A_E + A_I + A_S + A_{P_{Tot}} [=] km^2$$

B.1.9 Emergy Equations

Parameters for Emergy

τr_{ste} (sej/kg) = transformity of steel

τr_{sol} (sej/kg) = average transformity of industrial chemicals

τr_{con} (sej/kg) = transformity of concrete

τr_{apc} (sej/$) = transformity of annualized purchase cost

τr_{mo} (sej/$) = transformity of maintenance and overhead cost

τr_{lbr} (sej/$) = transformity of labor

τr_{ele} (sej/kWh) = transformity of electricity

τr_{wtr} (sej/kg) = transformity of water

τr_{dsl} (sej/kg) = transformity of diesel

τr_{air} (sej/kg) = transformity of air

αg (sej/km^2.yr) = specific global emergy per annum

Variables for Emergy

SolEm (sej/yr) = emergy due to solvents

QEm (sej/yr) = emergy due to material of construction of technology

PEm (sej/yr) = emergy associated with annualized purchase cost of technology

LEm (sej/yr) = emergy associated with labor

MOEm (sej/yr) = emergy associated with maintenance and overhead cost

UEm (sej/yr) = emergy due to utilities

TotEm (sej/yr) = total emergy

SEY (m^2) = specific emergy yield (emergy carrying capacity)

SEI (--) = specific emergy intensity

Ren (%) = percentage renewability

EYR (--) = emergy yield ratio

ELR (--) = emergy loading ratio

ESI (--) = emergy sustainability index

Equations:

Emergy due to solvents:

$$SolEm = \tau_{ann}\tau r_{sol}\sum_{k} M_{1,k} = \frac{hr}{yr}\frac{sej}{kg}\frac{kg}{hr}\ [=]\ \frac{sej}{yr}$$

Emergy associated with material of construction of technology:

$$QEm = \left(\rho_{ste}\tau r_{ste}Qc_{i=dst,pvp,uf,dry,cnt}\right) + \left(\rho_{con}\tau r_{con}Qc_{i=sdm,flt}\right) = \frac{kg}{m^3}\frac{sej}{kg}\frac{m^3}{yr}\ [=]\frac{sej}{yr}$$

NB: if Qc is in m^2, it was multiplied by the height of the technology in Table B.1.1 to get m^3.

Emergy associated with annualized purchase cost of technology:

$$PEm = \tau r_{apc}\, CCAC = \frac{sej}{\$}\frac{\$}{yr}[=]\frac{sej}{yr}$$

Emergy associated with labor

$$LEm = \tau r_{lbr} CCLB = \frac{sej}{\$}\frac{\$}{yr}\ [=]\frac{sej}{yr}$$

Emergy associated with maintenance and overhead cost:

$$MOEm = \tau r_{mo} CCOC = \frac{sej}{\$}\frac{\$}{yr}\ [=]\frac{sej}{yr}$$

Emergy due to utilities:

$$UEm = \left(\tau_{ann}\tau r_{ele}\sum_i PW_i\right) + \left(\tau_{ann}\tau r_{wtr}\sum_i (M_{stm_i} + M_{cw_i})\right)$$
$$= \frac{hr}{yr}\frac{sej}{kWh}\frac{kWh}{hr} + \frac{hr}{yr}\frac{sej}{kg}\frac{kg}{hr}\ [=]\frac{sej}{yr}$$

Total emergy:

$$TotEm = SolEm + QEm + PEm + LEm + MOEm + UEm\ [=]\frac{sej}{yr}$$

Emergy carrying capacity:

$$SEY = \frac{TotEm}{\alpha g} = \frac{(\frac{sej}{yr})}{(\frac{sej}{km^2 yr})}\ [=]\ km^2$$

Specific emergy intensity:

$$SEI = \frac{A_{Tot}}{SEY} = \frac{km^2}{km^2}\ [=]\ (--)$$

Renewability percentage:

$$Ren = \frac{\tau_{ann}\tau r_{wtr}\sum_i (M_{stm_i} + M_{cw_i}) * 100}{TotEm} = \frac{(\frac{sej}{yr})(\%)}{(\frac{sej}{yr})}\ [=]\ \%$$

Emergy yield ratio:

$$EYR = \frac{TotEm}{QEm + PEm + LEm + MOEm + UEm} = \frac{(\frac{sej}{yr})}{(\frac{sej}{yr})} = (--)$$

Emergy loading ratio:

$$ELR = \frac{TotEm - (\tau_{ann}\tau r_{wtr}\sum_i (M_{stm_i} + M_{cw_i}))}{(\tau_{ann}\tau r_{wtr}\sum_i (M_{stm_i} + M_{cw_i}))} = \frac{(\frac{sej}{yr})}{(\frac{sej}{yr})} = (--)$$

Emergy sustainability index:

$$ESI = \frac{EYR}{ELR} = (--)$$

B.1.10 Equations for Individual Technologies

Ultrafiltration (UF)

Unit Specific Parameters

ζ_i (m^3/m^2h) = Flux of technology *i*

ξ_k, (—) = Retention factor of component k for technology *i*

Retention factor equations:

$$\xi_{k,UF} = \frac{M_{Jr_{uf},k}}{M_{Jin_{uf},k}}; \forall k \in K_j$$

Concentration factor:

$$CF_{UF} = \frac{\sum_{k \in K_j, j \in Jin_{UF}} \left(\frac{M_{j,k}}{\rho_k}\right)}{\sum_{k \in K_j, j \in Jretentate_{UF}} \left(\frac{M_{j,k}}{\rho_k}\right)}$$

$$1.01 \leq CF_{UF} \leq 35$$

Flux balance:

$$\zeta_{UF} Qc_{UF} = [\sum_{k \in K_j, j \in Jin_{UF}} \left(\frac{M_{j,k}}{\rho_k}\right)] \left(1 - \frac{1}{CF_{UF}}\right)$$

Power required:

$$PW_{UF} = Wsp_{UF} Qc_{UF}$$

Direct installation area:

$$A_{ID_{UF}} = Q_{UF}$$

Pervaporation (PVP)

Unit specific parameters

λ_{stm} (KJ/kg) = Latent heat of steam

$\lambda_{vap,k}$ (KJ/kg) = Heat of vaporization of component k

Retention factor:

$$\xi_{k,PVP} = \frac{M_{Jr_{PVP},k}}{M_{Jin_{PVP},k}}; \forall k \in K_j$$

Concentration factor:

$$CF_{PVP} = \frac{\sum_{k \in K_j, j \in Jin_{PVP}} \left(\frac{M_{j,k}}{\rho_k}\right)}{\sum_{k \in K_j, j \in Jr_{PVP}} \left(\frac{M_{j,k}}{\rho_k}\right)}$$

Flux balance:

$$\zeta_{PVP} Q_{C_{PVP}} = \sum_{k \in K_j, j \in Jin_{PVP}} \left(\frac{M_{j,k}}{\rho_k}\right)] \left(1 - \frac{1}{CF_{PVP}}\right)$$

Power required:

$$PW_{PVP} = Wsp_{PVP} Qc_{PVP}$$

Heat required for vaporization:

$$Mstm_{PVP} \lambda_{stm} = \sum_{k \in K_j, j \in Jprm_{PVP}} M_{j,k} \lambda_k^{vap}$$

Direct installation area:

$$A_{ID_{PVP}} = Q_{PVP}$$

Distillation (DST)

Terms

LK = Light Key (Top Product)

HK = Heavy Key (Bottom Product)

λ_{stm} (kJ/kg) = Latent heat of steam

$\lambda_{vap,k}$ (kJ/kg) = Heat of vaporization of component k

T_{amb} (°C) = ambient temp

T_{sat} (°C) = saturation temp

α_k = Relative volatility of component k for technology i

η_{stage} = stage efficiency

Cp (KJ/kg-°C) = Specific heat of component k

Tcw_i (°C) = Cooling water temperature in (25)

Tcw_o (°C) = Cooling water temperature out (30)

$X_{mj,k}$ – mole fraction of component *k* in stream *j*

R_{min} – minimum reflux ratio

R – actual reflux ratio

N – the actual number of stages

Qs_{DST} – heat required to bring the feed to saturation

QH_{DST} - heat duty for the distillation unit

QC_{DST} – the cooling requirement for the distillation unit

Liq_{DST}- liquid molar flowrate in the distillation column

Vap_{DST} – vapor molar flowrate in the distillation column

U_v – Underwood variable

q – the quality of the mixture

Unit Specific Model Equations:

Molar flow rates in DST:

$$F_{j,k} = \frac{M_{j,k}}{MW_k}; \forall\, j \in J^{DST}, k \in K^{DST}$$

Component balance in DST:

$$\sum_{j \in Jin_i} F_{j,k} = \sum_{j \in Jout_i} F_{j,k}\,; \forall\, j \in J^{DST}, k \in K^{DST}$$

Mole fractions in DST:

$$X_{mj,k} = \frac{F_{j,k}}{\sum_{k \in K^{DST}} F_{j,k}};\ \forall\, j \in J^{DST}, k \in K^{DST}$$

Constraints on recovery:

$$X_{m_{Jtop_{DST},k}} when(\alpha_k < \alpha_{HK}) = 0;\ \forall\, k \in K^{DST}$$

$$X_{m_{Jtop_{DST},k}} when(\alpha_k > \alpha_{LK}) = 0; \forall\, k \in K^{DST}$$

Distillate recovery constraints:

$$X_{m_{Jtop_{DST},HK}} = 0.08$$

$$X_{m_{Jtop_{DST},LK}} = 0.92$$

Minimum number of stages with Fenske's equation:

$$N_{min} \log(\alpha_i) = log\,[\frac{X_{m_{Jtop_{DST},LK}}}{X_{m_{Jtop_{DST},HK}}} \frac{X_{m_{Jbot_{DST},HK}}}{X_{m_{Jbot_{DST},LK}}}]$$

Underwood's variable:

$$(1-q) = \sum_{k \in K^{DST}, j \in Jin_{DST}} \frac{\alpha_k X_{m_{j,k}}}{\alpha_k - U_v}$$

Assume feed is a saturated liquid (q=1):

$$0 = \sum_{k \in K^{DST}, j \in Jin_{DST}} \frac{\alpha_k X_{m_{j,k}}}{\alpha_k - U_v}$$

Minimum reflux ratio:

$$R_{min} = \sum_{k \in K^{DST}, j \in Jtop_{DST}} \frac{\alpha_k X_{m_{j,k}}}{\alpha_k - U_v} - 1$$

Reflux ratio:

$$R = 1.3 R_{min}\ (assumption)$$

Number of stages:

$$0.6N = N_{min}$$

Number of actual stages:

$$N_{act} = \frac{N}{\eta_{stage}}$$

Height of column:

$$H_{DST} = H_{stage} N_{act}$$

Liquid and vapor flowrates:

$$Liq_{DST} = R \sum_{k \in K^{DST}, j \in Jtop_{DST}} M_{j,k}$$

$$Vap_{DST} = Liq_{DST} + R \sum_{k \in K^{DST}, j \in Jtop_{DST}} M_{j,k}$$

Column diameter:

$$D_{DST} = \sqrt{\frac{4 Vap_{DST}}{\pi u_{vap}}}$$

$$u_{vap} = vapor\ linear\ velocity$$

Costing variable of column:

$$Qc_{DST} = \frac{\pi}{4} D_{DST}^2 H$$

Initial heating of feed to reach saturation:

$$QS_{DST} = \sum_{k \in K^{DST}, j \in Jin_{DST}} M_{j,k} Cp_k (T_{sat} - T_{amb})$$

Heat duty:

$$QH_{DST} = (1 + R) \sum_{k \in K^{DST}, j \in Jtop_{DST}} F_{j,k} MW_k \lambda_k^{vap}$$

Cooling:

$$QC_{DST} = R \sum_{k \in K^{DST}, j \in Jtop_{DST}} F_{j,k} MW_k \lambda_k^{vap}$$

Steam required:

$$Mstm_{DST} \lambda_{stm} = QS_{DST} + QH_{DST}$$

Cooling water required:

$$Mcw_{DST} Cp_w (Tcw_{out} - Tcw_{in}) = QC_{DST}$$

Variable bounds:

$$N_{min} \geq y_{DST}$$

$$R_{min} \geq 1.01 y_{DST}$$

Sedimentation (SDM)

Unit Specific Parameters

Ce_i^0 ($) – equipment cost for technology *i* of known capacity

n_i – cost exponent for technology *i*

V_i^0 (m^3) – vessel volume for technology *i* of known capacity A_i^0 (m^2) – area for technology i of known capacity

W_i^0 (kW) – power consumption for technology *i* of known capacity

θ_i^R (hr) – residence time in technology *i*

ρ_k (kg/m3) – density of component k

π_k ($/kg) –market price of k^{th} component

η_{SDM} (−) – efficiency of removal in typical sedimentation unit (75%)

Variables

V_i (m^3) – vessel volume for technology *i*

A_i (m^2) – area for technology *i*

CP_k ($/hr) – cost price for component k consumed per hour

Ce_i ($) – equipment cost for technology *i*

Equations:

Settling velocity (evaluated parameter**)**:

$$U_{S,SDM} = \frac{g D_p^2 (\rho_s - \rho_L)}{18\mu}$$

$U_{S,SDM}$ – settling velocity (m/s)

D_p – particle diameter (m) (5E-3 m or 5 mm)

g – acceleration due to gravity (m/s^2)

ρ_s – density of solid (kg/m^3)

ρ_l – density of liquid (kg/m^3)

μ – viscosity of fluid ($N\text{-}s/m^2$)

Efficiency:

$$\eta_{SDM} = \frac{M_{Jtp_{SDM},Sol}}{M_{Jin_{SDM},Ssol}}$$

Concentration factor: (volume concentration factor)

$$CF_{SDM} = \frac{\sum_{k \in K_j, j \in Jin_i} \left(\frac{M_{j,k}}{\rho_k}\right)}{\sum_{k \in K_j, j \in Jtp_{SDM}} \left(\frac{M_{j,k}}{\rho_k}\right)}$$

$$1.01 \leq CF_{SDM} \leq 15$$

Surface overflow rate:

$$SOR_{SDM} = \frac{U_{S,SDM}}{\eta_{SDM}}$$

Area of sedimentation tank:

$$A_{SDM} = \frac{\sum_{k \in K_j j \in Jin_{SDM}} \left(\frac{M_{j,k}}{\rho_k}\right)}{SOR_{SDM}}$$

Direct installation area:

$$A_{ID_{SDM}} = Q_{SDM}$$

Dryer (DRY)

Unit specific parameters

v_{air}(m/s) = velocity of air flow in the dryer

Moisture Content in stream j:

$$X_j = \frac{M_{j,Wtr}}{M_{j,k}}; \forall\, k \in K_j$$

Diameter of the Drum:

$$D_{DRY} = \frac{M_{Jda_{DRY},k}}{\sqrt{v_{air}\pi * 900\rho_{air}}}; \forall k \in K_j$$

Length of Drum:

$$L_{DRY} = B_{DRY} * D_{DRY}$$

$$4 \leq B_{DRY} \leq 15$$

Heat required for vaporization:

$$Q_{c,DRY} = \frac{\pi}{4} D_{DRY}^2 L_{DRY}$$

Power required:

$$PW_{DRY} = \frac{3.19995 M_{Jda_{DRY},k}}{MW_{air}};\ k \in K_j$$

Aqueous Two-Phase Extraction (ATPE)

Unit Specific Parameters

$\Psi_{k\text{-}k'}$- solubility of component k in component k'

κP_k- partition coefficient of component k

Solubility Equations:

$$M_{Jbp_{ATPE},poly} = \psi_{poly-bp} M_{Jbp_{ATPE},salt}$$

$$M_{Jtp_{ATPE},salt} = \psi_{salt-tp} M_{Jtp_{ATPE},poly}$$

Extraction Factor:

$$EF_{ATPE} = \frac{\kappa P_k M_{Jpoly_{ATPE},k}}{M_{Jsalt_{ATPE},k}}$$

Number of Stages:

$$\left(\frac{EF-1}{EF^{NAE+1}-1}\right) = \frac{M_{Jfeed_{ATPE},k} - M_{Jtop_{ATPE},k}}{M_{Jfeed_{ATPE},k}}$$

Size of unit:

$$Q_{c,ATPE} = \sum_{k\in K_j, j\in Jfeed_{ATPE}} \frac{M_{j,k}}{\rho_k} + \sum_{k\in K_j j\in Jpoly_{ATPE}} \frac{M_{j,k}}{\rho_k} + \sum_{k\in K_j, j\in Jsalt_{ATPE}} \frac{M_{j,k}}{\rho_k}$$

Power Required:

$$PW_{ATPE} = Wsp_{ATPE}\ Qc_{,ATPE}$$

Cooling Duty:

$$M_{cw,ATPE} = \frac{3600\ PW_{ATPE}}{c_p\left(T_{cw,out} - T_{cw,in}\right)}$$

Centrifugation (CNF)

Unit Specific Parameters

Efficiency Equation:

$$\eta_{water} = \frac{M_{Jsld_{CNF},WTR}}{M_{Jsld_{CNF},WTR}}$$

$$\eta_{solvent} = \frac{M_{Jliq_{CNF},solvent}}{M_{Jfeed_{CNF},solvent}}$$

Concentration Factor:

$$CF_{CNF} = \frac{\left[\sum_{k\in K_j, j\in Jfeed_{CNF}} \left(\frac{M_{j,k}}{\rho_k}\right)\right]}{\left[\sum_{k\in K_j j\in Jliq_{CNF}} \left(\frac{M_{j,k}}{\rho_k}\right)\right]}$$

$$2 \le CF_{CNF} \le 20$$

Sigma Factor Equation:

$$Qc_{CNF}U_{CNF} = \left[\sum_{k\in K_j j\in Jfeed_{CNF}} \left(\frac{M_{j,k}}{\rho_k}\right)\right]$$

Power Required:

$$PW_{CNF} = Wsp_{CNF}\left[\sum_{k\in K_j, j\in Jfeed_{CNF}} \left(\frac{M_{j,k}}{\rho_k}\right)\right]$$

Power dissipation to heat it about 40%, therefore cooling duty is required:

$$Mcw_{CNF}c_{pw}\left(T_{cw,out} - T_{cw,in}\right) = 0.4PW$$

Filtration (FLT)

Retention factor:

$$\xi_{k,FLT} = \frac{M_{Jflt_{FLT},k}}{M_{Jfeed_{FLT},k}}; \forall k \in K_j$$

Concentration factor:

$$CF_{FLT} = \frac{\sum_{k\in K_j, Jfeed_{FLT}} \left(\frac{M_{j,k}}{\rho_k}\right)}{\sum_{k\in K_j, Jflt_{FLT}} \left(\frac{M_{j,k}}{\rho_k}\right)}$$

$$2 \leq CF_{FLT} \leq 30$$

Flux balance:

$$\zeta_{FLT}Q_{C_{FLT}} = \sum_{k\in K_j, j\in Jfeed_{FLT}} \left(\frac{M_{j,k}}{\rho_k}\right)\left(1 - \frac{1}{CF_{FLT}}\right)$$

Power required:

$$PW_{FLT} = Wsp_{FLT}Qc_{FLT}$$

Direct installation area:

$$A_{ID_{FLT}} = Q_{FLT}$$

Incineration (INCN)

Process Equations

Heating value of waste stream:

$$Q_{solv,INCN} * 1000 = 14544 * C + 62208\left(H - \frac{O}{8}\right) + 4050 * S\ [=]\ MJ/kg$$

Mass of fuel needed for heating:

$$m_{fuel,INCN} * NE_{fuel} = Q_{solv} * m_{solv}\ [=]\ MJ/s$$

$$NE_{fuel}\ = 38.9\ [=]\ MJ/kg$$

*Fuel oils are products of petroleum distillation, consists of hydrocarbons

Mass of air fed (textbook):

$$m_{air,INCN} = air_{rate} * O * m_{solv}\ [=]\ kg/s$$

$$air_{rate}\ =\ 4.35$$

Energy consumed during process:

$$E_{con,INCN} = Q_{solv,INCN} * m_{solv,INCN}\ [=]\ MJ/s$$

Energy produced during process:

$$E_{prod,INCN} = eff_{INCN} * E_{con,INCN}\ [=]\ MJ/s$$

Efficiency of energy production ranges from 30-45%

Net energy:

$$E_{net,INCN} = E_{prod,INCN} - E_{con,INCN}\ [=]\ MJ/s$$

Costing Equations

Annual fuel cost:

$$Mat_{INCN} = m_{fuel} * C_{fuel} * 3600 * 24 * 340\ [=]\ \$/yr$$

$$Cost\ of\ fuel = C_{fuel} = 0.81\ [=]\ \$/kg$$

Annual energy cost:

$$E_{cost,INCN} * 3.6 = E_{net} * Ce * 3600 * 24 * 340\ [=]\ \$/yr$$

$$Ce = 0.10\ [=]\ \$/kWh$$

Conversion factor: 3.6 MJ/kWh

Annual air cost (hydraulics and pneumatics):

$$OC_{INCN} * \rho_a = C_{air,INCN} * Mass_{air} * 3600 * 24 * 340 \text{ [=]} \$ \frac{kg}{m^3\ yr}$$

$$C_{air,INCN} = 0.0004 \text{ [=] } \$/m^3$$

Capital cost:

$$Cap_{cost,INCN} = Cap_{old,INCN} * \left(\frac{m_{solv,incn} * 3600}{m_{std,incn}}\right)^{nc} \text{ [=] } \$$$

$$Cap_{old,INCN} = 967000 \text{ [=] } \$$$

$$M_{std,INCN} = 100000 \text{ [=] } kg/hr$$

Number of laborers (SuperPro):

$$Nlb_{INCN} * m_{std,INCN} = Nlabr_{INCN} * m_{solv,INCN} * 3600$$

$$Nlab_{INCN} = 0.1$$

Annual cost of labor:

$$Ncost_{INCN} = Nlb_{INCN} * Pay * 24 * 340 \text{ [=] } \$/yr$$

$$Pay = 30 \text{ [=] } \$/hr$$

Total annual cost (objective to be minimized):

$$CCT_{INCN} = N_{cost,INCN} + Cap_{cost,INCN} + OC_{INCN} + Mat_{INCN} + E_{cost,INCN}$$

B.1.11 Model specifications and input data (standard capacities and costs, parameters, feed compositions)

Table B1

Standard Capacity, Costs, Scaling Factors, Labor Requirements for Technologies

Unit operation (costing capacity)	Standard capacity (units)	Base costs (million $)	Scaling exponent (n)	Laborers required (#/hr)	Power required (kWh)	Consumable Costs ($/unit)	Height of Unit (m)
Sedimentation (Area)	2500 m^2	1.128	0.67	0.1	0	0	2
Filtration (Area)	80 m^2	0.039	0.67	0.5	0.1	400 ($/$m^2$)[c]	1.5
Microfiltration (Area)	80 m^2	0.75	0.67	1	0.1	400 ($/$m^2$)[c]	1
Centrifuge (Sigma factor)	60000 m^2	0.66	0.67	1	19.2	0	1
Distillation (Volume)	22.58 m^3	0.082	0.67	1	0	0	-
Pervaporation (Area)	80 m^2	0.0261	0.67	1	0.33	400 ($/$m^2$)[c]	1
Aqueous Two-Phase Extraction (volumetric flowrate)	185 m^3/hr	0.362	0.67	1	0.5	2.6 ($/kg)[a]	-
Ultrafiltration (Area)	80 m^2	0.938	0.67	1	0.2	400 ($/$m^2$)[c]	1
Dryer (Volume)	106 m^3	0.024	0.67	0.5	0	0	-
Incineration (Mass flowrate, kg/hr)	100000	0.967	0.67	0.1	~ [b]	0	-

a. This cost is the consumable cost associated with adding in the hexane and salt into the aqueous two-phase extraction unit. The unit cost of hexane is $2/kg and the unit cost of sodium chloride salt is $0.6/kg

b. This value is dependent on the composition of the incoming stream. Different compounds have different heat of combustions, which will cause variation in the power required.

c. The replacement time for all filter consumables in assumed to be 2000 hours.

Table B2

Utility and Labor Costs (SuperPro Designer v8.5)

Utility	**Cost per unit ($/unit)**
Electricity	$0.1/kWh
Cooling Water	$5E-5/kg
Steam	$0.012/kg
Labor	$30/laborer*hr

Table B3

Input Component Parameters for Case Study 1 – Pharmaceutical Waste Stream

Component	**Molecular weight of component (kg kmol^{-1})**	**Density (kg m^{-3})**	**Heat of vaporization of component c(kJ kg^{-1})**	**Heat capacity of component c (kJ kg^{-1} C^{-1})**	**Feed mass fraction (kg component c kg feed^{-1})**
Isopropanol	60	786	664	2.32	0.51
Water	18	1000	a.	4.2	0.49
Salt	138	2430	a.	a.	a.
Hexane	86	655	a.	1.58	a.

a. This value was not a required input for the model

Table B4

Input Component Parameters for Case Study 2 – Specialty Chemical Waste Stream

Component	**Molecular weight of component (g mol^{-1})**	**Density (Kg cm^{-3})**	**Heat of vaporization of component c (kJ kg^{-1})**	**Heat capacity of component c (kJ kg^{-1} C^{-1})**	**Feed mass fraction (kg component c kg feed^{-1})**
Dimethoxyethane	90	867	418.6	1.42	0.167
Water	18	1000	a.	4.2	0.276
Toluene	92	876	401.6	1.71	0.323
Ethoxy methoxy ethane	104	800	400	1.5	0.01
Salt	142	2671	a.	0.9	0.218
Air	29	0.864	a.	a.	a.

a. This value was not a required input for the model

B.1.11.1 Aqueous two phase extraction (ATPE):

Residence time: 2 h

Partition coefficient: Isopropanol – 8, Water – 0.05

Solubility Parameter: Hexane in bottom phase – 0.005, Salt in top phase – 0.005

B.1.11.2 Sedimentation (SDM):

Residence time: 6 h

Efficiency: 70%

B.1.11.3 Ultrafiltration (UF):

Flux: 0.0856 m^3 m^{-2} h^{-1}

Retention Factor (UF1): Isopropanol – 0%, Water – 100%, Salt -100%, Hexane – 100%

Retention Factor (UF2): Isopropanol – 0%, Water – 100%

B.1.11.4 Pervaporation (PVP):

Flux: 0.55 m^3 m^{-2} h^{-1}

Retention Factor (PVP1): Isopropanol – 5%, Water – 90%

Retention Factor (PVP2): Isopropanol – 1%, Water – 90%

B.1.11.5 Filtration (FLT):

Flux: 0.2 m^3 m^{-2} h^{-1}

Retention Factor: Dimethoxy ethane - 10%, Toluene – 10%, Ethoxy methoxy ethane – 10%, Salt – 100%, Water – 100%

B.1.11.6 Pervaporation (PVP):

Flux: 0.55 m^3 m^{-2} h^{-1}

Retention Factor: Dimethoxy ethane - 5%, Toluene – 97%, Ethoxy methoxy ethane – 5%

B.1.11.7 Ultrafiltration (UF):

Flux: 0.0856 m^3 m^{-2} h^{-1}

Retention Factor: Dimethoxy ethane - 5%, Toluene – 5%, Ethoxy methoxy ethane – 97%

B.1.12 Logic Equations for case study 1:

$$Y_{ATPE} \qquad \neg Y_{ATPE}$$
$$\left[M_{1,k} = M_{2,k}\right] \qquad \left[M_{2,k} = 0\right]$$

$$Y_{PVP1} \qquad \neg Y_{PVP1}$$
$$\left[M_{1,k} = M_{3,k}\right] \qquad \left[M_{3,k} = 0\right]$$

$$Y_{DST} \qquad \neg Y_{DST}$$
$$\begin{bmatrix} M_{1,k} = M_{4,k} \\ M_{20,IPA} \geq 0.87 * M_{4,IPA} \end{bmatrix} \qquad \left[M_{24,kC} = M_{19,kC} * Y_{DST}\right]$$

$$Y_{ATPE} \vee Y_{PVP1} \vee Y_{DST}$$

B.1.13 Logic Equations for case study 2:

$$Y_{FLT} \qquad \neg Y_{FLT}$$
$$\left[M_{1,k} = M_{3,k}\right] \qquad \left[M_{3,k} = 0\right]$$

$$Y_{SDM} \qquad \neg Y_{SDM}$$
$$\left[M_{1,k} = M_{4,k}\right] \qquad \left[M_{4,k} = 0\right]$$

$$Y_{CNF} \qquad \neg Y_{CNF}$$
$$\left[M_{1,k} = M_{5,k}\right] \qquad \left[M_{5,k} = 0\right]$$

$$Y_{PVP} \qquad \neg Y_{PVP}$$
$$\left[M_{17,kC} = M_{15,kC}\right] \qquad \left[M_{17,kC} = 0\right]$$

$$Y_{DST1} \qquad \neg Y_{DST1}$$
$$\left[M_{16,kC} = M_{15,kC}\right] \qquad \left[M_{19,kC} = M_{32,kC} * Y_{DST1}\right]$$

$$Y_{UF1} \qquad \neg Y_{UF1}$$
$$[M_{26,kC} = M_{22,kC}] \qquad [M_{26,kC} = 0]$$

$$Y_{DST2} \qquad \neg Y_{DST2}$$
$$[M_{23,kC} = M_{22,kC}] \qquad [M_{25,kf} = M_{33,kf} * Y_{DST2}]$$

$$Y_{FLT} \vee Y_{SDM} \vee Y_{CNF}$$
$$Y_{PVP} \vee Y_{DST1}$$
$$Y_{UF1} \vee Y_{DST2}$$

Table B5

Emission Factors for Case Study 2

Component	DME (kg emission/kg DME)	EME (kg emission/kg EME)	TOL (kg emission/kg TOL)	Water (kg emission/kg Water)	Salt (kg emission/kg Salt)
CO_2	2.02E+00	2.02E+00	1.24E+00	5.13E-04	6.92E-01
CO	2.45E-03	2.45E-03	2.28E-03	1.86E-06	1.20E-03
CH_4	1.18E-02	1.18E-02	1.24E-02	1.48E-06	1.77E-03
NO_X	4.35E-03	4.35E-03	2.44E-03	1.34E-06	2.22E-03
NMVOCs	1.77E-03	1.77E-03	1.92E-03	2.32E-07	2.77E-04
Others	1.11E-02	1.11E-02	3.04E-03	3.86E-06	9.27E-03
WE	2.50E-01	2.50E-01	8.16E-03	3.56E-04	1.41E-01
SE	1.73E-03	1.73E-03	1.83E-03	2.22E-07	4.86E-04

Table B6

Emission Factors for Case Study 2 (Cont.)

Component	Electricity (kg emission/kWh electricity)	Steam (kg emission/kg Steam)
CO_2	6.38E-01	2.96E-01
CO	2.50E-04	1.58E-04
CH_4	1.05E-03	6.13E-04
NO_X	7.04E-04	3.85E-04
NMVOCs	7.25E-05	6.54E-05
Others	4.26E-03	1.17E-02
WE	1.22E-01	1.26E-02
SE	5.04E-05	1.20E-04

Table B7

Emission Factors for Case Study 1

Component	IPA (kg emission/ kg IPA)	Water (kg emission/kg Water)	Steam (kg emission/kg Steam)	Electricity (kg emission/kWh electricity)
CO_2	1.82E+00	5.13E-04	2.96E-01	6.38E-01
CO	2.32E-03	1.86E-06	1.58E-04	2.50E-04
CH_4	1.05E-02	1.48E-06	6.13E-04	1.05E-03
NO_X	3.49E-03	1.34E-06	3.85E-04	7.04E-04
NMVOCs	1.81E-03	2.32E-07	6.54E-05	7.25E-05
Others	2.93E-02	3.86E-06	1.17E-02	4.26E-03
WE	2.85E-01	3.56E-04	1.26E-02	1.22E-01
SE	4.56E-04	2.22E-07	1.20E-04	5.04E-05

Table B8

Fraction of Emitted Component that is Dissipated into Compartment c

Component	Air Compartment	Water Compartment	Soil Compartment
CO_2	1	0	0
CO	1	0	0
CH_4	1	0	0
NO_X	1	0	0
NMVOCs	1	0	0
Others	1	0	0
WE	0	1	0
SE	0	0	1

Table B9

Yearly Allowable Yield of Component Per Square Meter

Component	Air Compartment	Water Compartment	Soil Compartment
CO_2 (kg $CO_2/m^2.yr$)	6.53E-01	-	-
CO (kg $CO/m^2.yr$)	9.80E-03	-	-
CH_4 (kg $CH_4/m^2.yr$)	4.50E-03	-	-
NO_X (kg $NO_X/m^2.yr$)	1.31E-04	-	-
NMVOCs (kg NMVOCs /$m^2.yr$)	6.50E-03	-	-
Others (kg Others /$m^2.yr$)	3.90E-04	-	-
WE (kg WE /$m^2.yr$)	-	1.22E-04	-
SE (kg SE /$m^2.yr$)	-	-	1.32E-06

Table B10

Yearly Specific Yield

Yield factor	Value
Non-renewable resource ($kg/m^2.yr$)	2
Staff (m^2/person)	4.74E+3
Indirect installation ($kWh/m^2.yr$)	6
Energy ($kWh/m^2.yr$)	43

Table B11

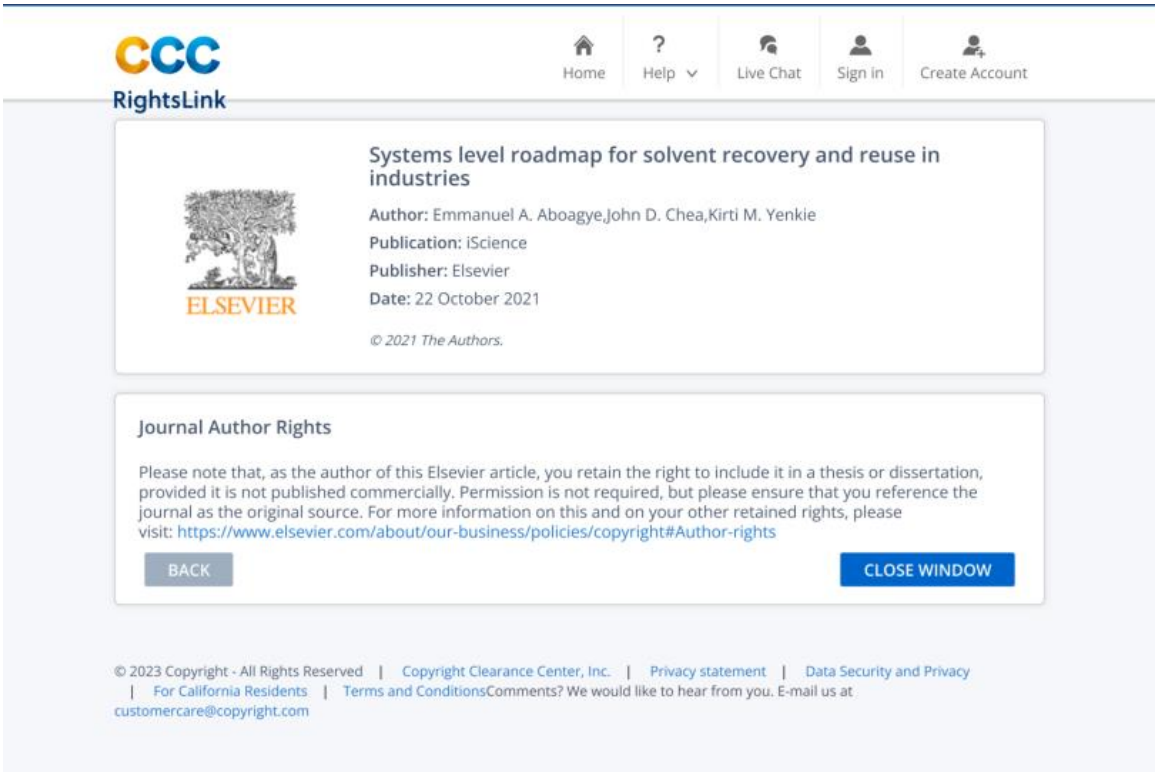
Transformity and other Parameters for Emergy Analysis

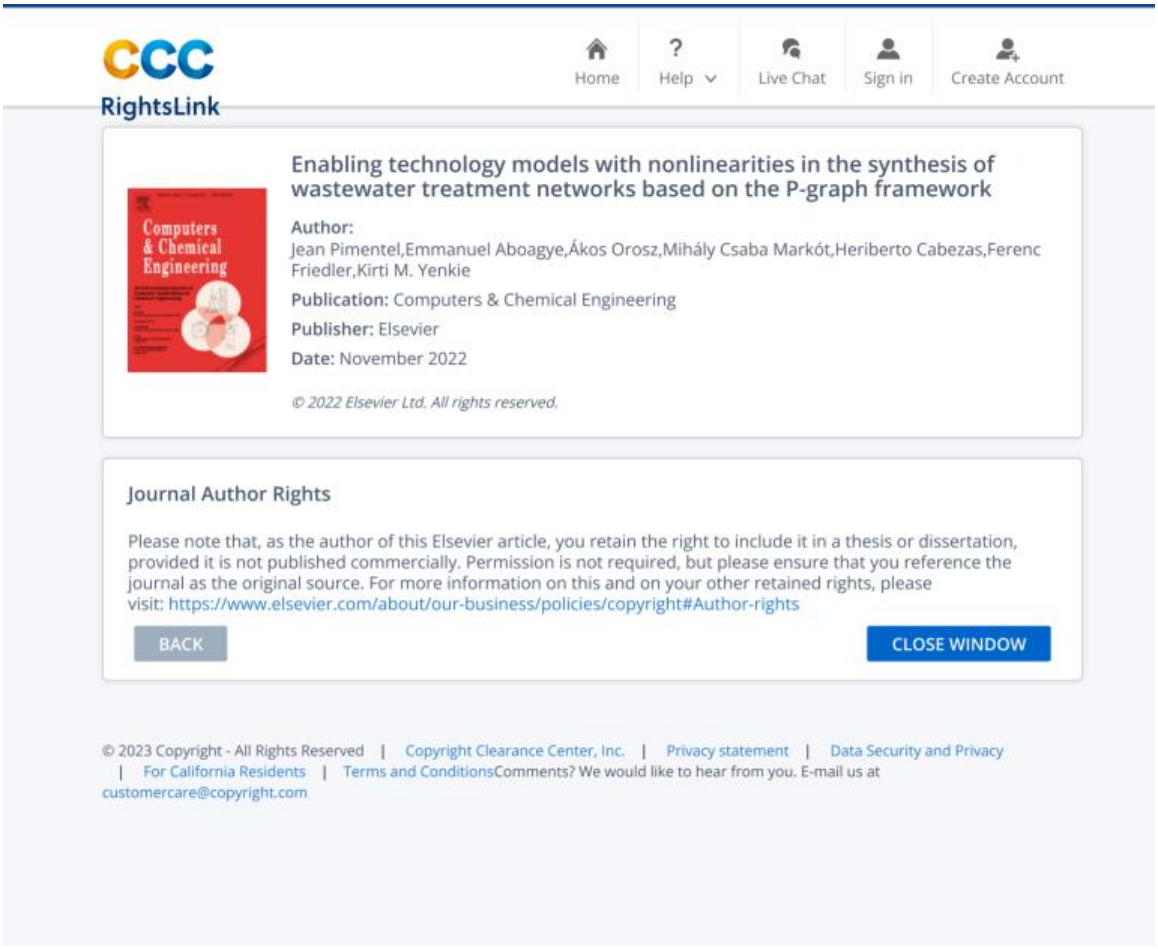
Component	Unit	Value ($x10^{10}$)
Steel	sej/kg	8.70
Concrete	sej/kg	227
Purchase Cost	sej/$	494
Maintenance Cost	sej/$	494
Labor	sej/$	280
Water	sej/kg	0.159
Electricity	sej/kWh	122
Diesel	sej/kg	40
Air (Wind)	sej/kg	1.03
Annual global emergy density	$sej/m^2.yr$	3100000
Density of Steel	kg/m^3	8050
Density of Concrete	kg/m^3	2400

Appendix C

Copyright Permissions and Educational Contributions from this Work

Chapter 1 Texts, Figures, & Tables




CCC
RightsLink
Home
Help
Live Chat
Sign in
Create Account
Enabling technology models with nonlinearities in the synthesis of wastewater treatment networks based on the P-graph framework
Author:
Jean Pimentel,Emmanuel Aboagye,Ákos Orosz,Mihály Csaba Markót,Heriberto Cabezas,Ferenc Friedler,Kirti M. Yenkie
Publication: Computers & Chemical Engineering
Publisher: Elsevier
Date: November 2022
© 2022 Elsevier Ltd. All rights reserved.
Journal Author Rights
Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights
BACK
CLOSE WINDOW
© 2023 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy | For California Residents | Terms and ConditionsComments? We would like to hear from you. E-mail us at customercare@copyright.com

Open Access Feature Paper Article

Systematic Design, Optimization, and Sustainability Assessment for Generation of Efficient Wastewater Treatment Networks

by Emmanuel A. Aboagye, Sean M. Burnham, James Dailey, Rohan Zia, Carley Tran, Maya Desai and Kirti M. Yenkie *

Department of Chemical Engineering, Henry M. Rowan College of Engineering, Rowan University, Glassboro, NJ 08028, USA

* Author to whom correspondence should be addressed.

Water **2021**, *13*(9), 1326; https://doi.org/10.3390/w13091326

Received: 6 March 2021 / Revised: 1 May 2021 / Accepted: 6 May 2021 / Published: 10 May 2021

(This article belongs to the Section Wastewater Treatment and Reuse)

CHEMICAL ENGINEERING TRANSACTIONS
VOL. 88, 2021
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Petar S. Varbanov, Yee Van Fan, Jiří J. Klemeš

ISBN 978-88-95608-86-0; **ISSN** 2283-9216
DOI: 10.3303/CET2188082

Efficient Design and Sustainability Assessment of Wastewater Treatment Networks using the P-graph Approach: A Tannery Waste Case Study

Emmanuel A. Aboagye[a], Jean Pimentel[b], Ákos Orosz[c], Heriberto Cabezas[d], Ferenc Friedler[e], Kirti M. Yenkie[a,*]

Publication Policy

CHEMICAL ENGINEERING TRANSACTIONS is indexed by:

SCOPUS since Vol. 17, 2009

Google Scholar

Each Volume is identified by proper **ISBN, ISSN** and since Volume 17 by a **DOI** number for each paper.

Each article is evaluated for acceptance by at least **two Peer Reviewers** (single blind peer review)

The list of the Peer Reviewers is available inside the Issue.

Each article published in the Volume is covered by a **copyright** assigned to **AIDIC Servizi S.r.l.**

REQUIREMENTS FOR PUBLICATION

- original article (never published before)
- reviewers acceptance
- full compliance with the Chemical Engineering Transactions formatting instructions
- publication fee paid
- transfer of copyright agreement duly signed

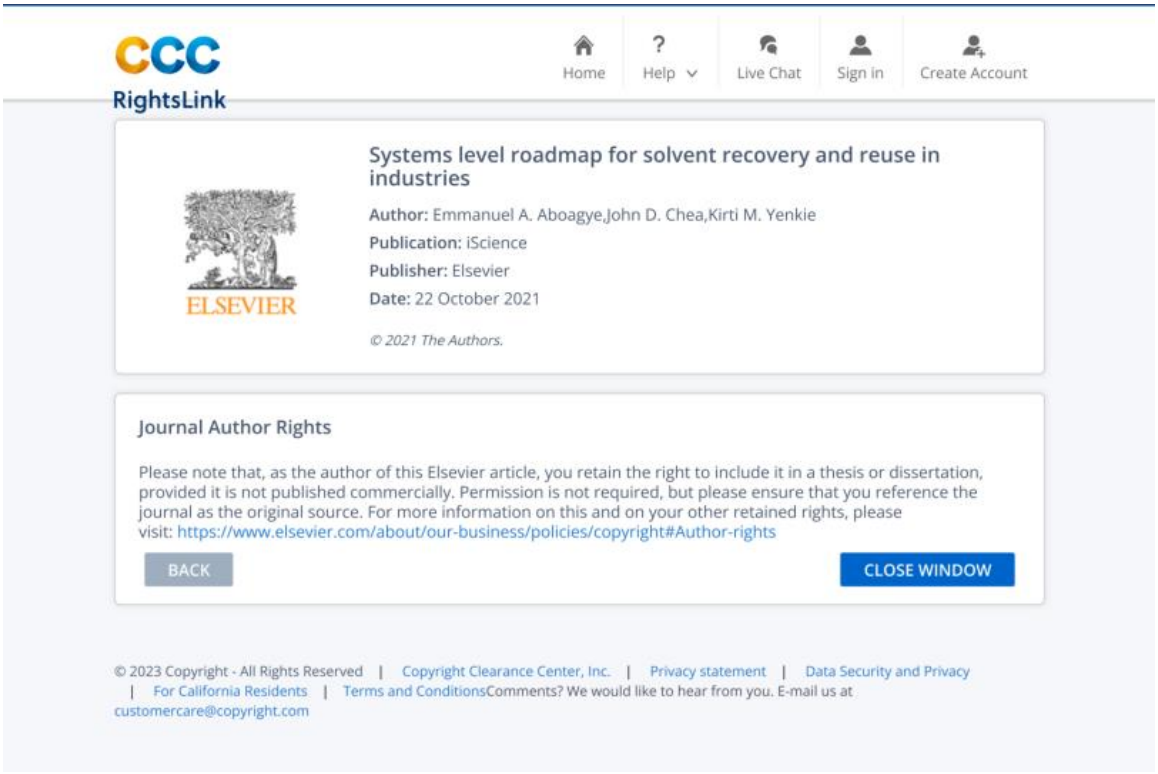
Open Access Policy

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Our policy is to permit Authors to reuse part of their CET articles or to self-archive the published version of their work in Institutional Repository, provided

---

that AIDIC/CET is acknowledged as the source. The version to be used is the Publisher's PDF.

No embargo period is required.


CCC
RightsLink
Home
Help
Live Chat
Sign in
Create Account
Systems level roadmap for solvent recovery and reuse in industries
Author: Emmanuel A. Aboagye,John D. Chea,Kirti M. Yenkie
Publication: iScience
Publisher: Elsevier
Date: 22 October 2021
© 2021 The Authors.
ELSEVIER
Journal Author Rights
Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights
BACK
CLOSE WINDOW
© 2023 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy | For California Residents | Terms and ConditionsComments? We would like to hear from you. E-mail us at customercare@copyright.com

Systematic Design of Solvent Recovery Pathways: Integrating Economics and Environmental Metrics

Author: Emmanuel A. Aboagye, John D. Chea, Austin L. Lehr, et al
Publication: ACS Sustainable Chemistry & Engineering
Publisher: American Chemical Society
Date: Aug 1, 2022

Teaching Sustainable Design through Simultaneous Evaluation of Economics and Environmental Impacts

Kirti M. Yenkie[1*], John D. Chea[1], Emmanuel A. Aboagye[1], Mariano J. Savelski[1], and C. Stewart Slater[1]

[1]Department of Chemical Engineering, Henry M. Rowan College of Engineering, Rowan University, Glassboro, New Jersey, USA.

yenkie@rowan.edu

Abstract

The ever-increasing human population and industrial growth have posed a considerable burden on existing resources and have led to an increase in environmental pollution and climate change. The Engineering Clinics offered at the Henry M. Rowan College of Engineering at Rowan University is the hallmark of our program that enables our undergraduate students to actively participate in solving real-world problems through collaborative activities. Our graduate students get an opportunity to engage in stakeholder (*i.e.,* industries, federal and regional funding agencies) interactions and student mentoring in conjunction with developing their research ability. Thus, through these synergistic undergraduate-graduate-faculty-stakeholder collaborations this work envisions to develop awareness about sustainable design and environmental impact in the community. The clinic problems include; (i) solvent recovery in process industries, and (ii) systematic synthesis of wastewater treatment (WWT) networks. These problems are important because imprudent use of industrial solvents and water resources have exacerbated the challenges relating to availability, quality as well as safe disposal of harmful solvents and wastewater. Through these challenging and relevant problems, we can teach our students multiple skills such as information collection, selective extraction of valuable content, economic and sustainability evaluation of multiple pathways through mathematical modeling, computer programming, technical writing, and presentation. The overall impact of these efforts is evident in the peer-reviewed conference and journal publications, oral and poster presentations at regional and national conferences, as well as our students choosing careers which value sustainability.

1 Introduction

The unique feature of the undergraduate curriculum at the Henry M. Rowan College of Engineering (HMRCOE) at Rowan University (RU) is the Engineering Clinics, which are offered in conjunction with all the required courses every semester. The undergraduate students from all the engineering disciplines are part of a common clinic activity in their first two years, which are aimed at enhancing the basic engineering skills and to increase an aptitude for reason-based learning. They also learn basic technical writing and presentation skills in these two years. In their junior (3^{rd}) and senior (4^{th}) years, these students get an opportunity to participate in discipline-specific research-based clinics where they have an opportunity to engage with stakeholders from industries and federal agencies and work on real-world problems. In this paper, we have placed an emphasis on one specific clinic project: solvent recovery in process industries. This project is offered in the Chemical Engineering department at RU to teach our students the importance of sustainable design and the impacts of chemical processes and their effluents on the environment. In the following sub-sections, the background and motivation in choosing this clinic project are emphasized.

1.1 Solvent Recovery and Reuse

The demand for solvents has expanded across many industries such as the pharmaceutical, adhesives, food, cosmetics, cleaning, and personal care industries. Solvents are typically used as dissolution medium, materials to aid in reaction, mass separation, and cleaning operations (Slater et al., 2010; Wypych, 2014). However, there are inefficiencies in the existing industrial manufacturing processes, which can be caused by large-scale production challenges such as inefficient mixing, insufficient reaction time, inappropriate technologies, quality of raw materials, measurement control anomalies, etc. (Cavanagh et al., 2014a; Raymond et al., 2010). The global chemical market is projected to double between 2017 and 2030. However, waste generation due to poor solvent selection and processing inefficiencies in the chemical industry have led to a growing concern for chemical releases, exposures, environmental impacts, and health safety (United Nations Environment Programme, 2019). The US EPA has estimated that solvent emissions resulting from the chemical market growth can reach up to 10 million metric tons of carbon dioxide equivalent (US EPA, 2016, p. 2).

1.2 Role of Process Systems Engineering (PSE)

The selection of appropriate solvent recovery technologies is a function of the physicochemical properties of solvents, other components present in the waste stream, and the desired final purity levels to be achieved after separation. These separation technology options may include sedimentation, filtration, precipitation, distillation, liquid-liquid extraction, and pervaporation (Chea et al., 2019). Hence, this problem belongs to the process systems engineering (PSE) area, which comprises multiple methods and their associated computational tools to systematically solve the problem of generation of solvent recovery framework.

Furthermore, the availability of multiple recovery technologies, such as distillation, pervaporation, and aqueous two-phase extraction, adds complexity to the selection process A comparative assessment of the solvent recovery methods to the existing waste handling methods such as incineration is crucial to change the mindset of the people working in process industries as well as our undergraduates, who are the future workforce of the nation. Through PSE tools, we can selectively choose appropriate materials/methods for the efficient design of treatment systems and their sustainability over the desired period. Through planned projects, educational activities, and result dissemination, we aim to create an appreciation for ‘Sustainable Design in Engineering’ and motivate students to pursue it as their career path.

2 Methodology

2.1 Project Teams & Management

The clinic project team is composed of 2-3 undergraduate students, a graduate student mentor, and faculty advisors. The faculty develops contact with industries and other universities, applies for research and educational grants to federal and regional agencies, and private funding organizations. The faculty is responsible for developing the project goals and learning objectives for the students. The graduate student mentor is responsible for ensuring the project continuity, documentation, and partial supervision of undergraduate students. The engineering clinic is a 2-credit course every semester with biweekly meeting slots of 3 hours each. This course provides ample time for required student training, progress assessment as well as consultations with industrial liaisons and collaborators.

2.2 Tutorials for Basic Research Skills

As faculties, we provide students initial training on the necessary research tools and resources. The most crucial aspect for both these clinic projects was a literature review to collect relevant information about existing industrial processes and their waste streams, characterization metrics, existing case studies, technology information, and modelling. To this end, the students were trained to use literature review resources such as Google Scholar, and SciFinder Scholar. Instructions were provided on reading research papers effectively as well as categorizing them into reviews, model information, case studies, optimization, and simulations. Furthermore, they were trained to use citation managers such as Zotero and Mendeley to create a systematic database of references and cite them in research reports and manuscripts.

The next set of tutorials included training in PSE tools for mathematical modelling and optimization. Since both, the clinic projects involve a selection decision between multiple waste treatment and resource recovery technologies to meet the cost criteria and minimization of overall environmental impacts, the optimization tools needed were non-linear programming as well as discrete programming (Biegler et al., 1997; Diwekar, 2013). The theory, as well as software training in Matlab, GAMS, and P-graph (Heckl, Friedler, and Fan 2010), were provided to the students. Training for the environmental impact assessment tools such as SimaPro (Cavanagh et al., 2014b) and Sustainable Process Index (Narodoslawsky and Krotscheck, 1995) were also provided. Additionally, resources for enhanced technical writing and presentation skills were taught. These tutorials were scheduled appropriately as per the project's progress and requirements. Figure 1 highlights the resources and tools from PSE that our research lab (the Sustainable Design & Systems Medicine Lab) has access to at Rowan University.

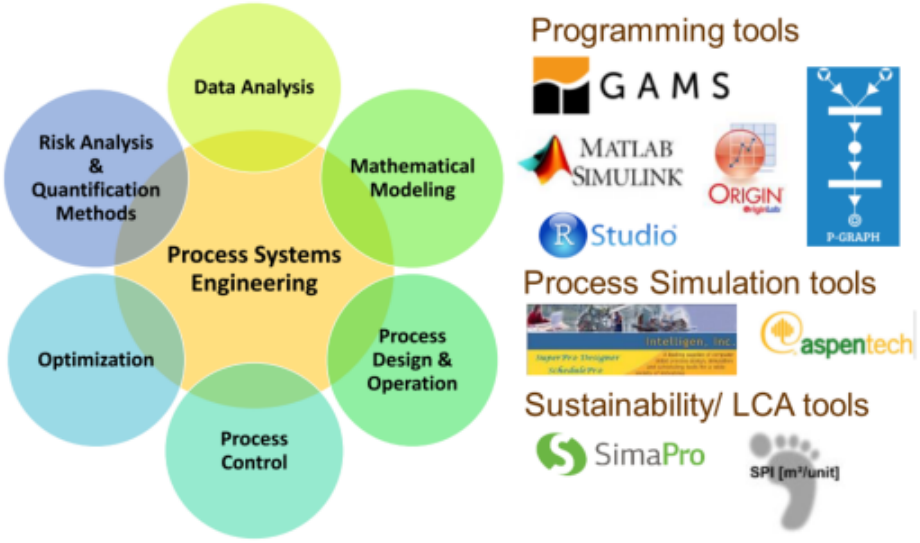


Figure 1: Process Systems Engineering (PSE) tools and computational resources at the Sustainable Design & Systems Medicine Laboratory at Rowan University.

2.3 Clinic Project; Solvent Recovery in Process Industries

This clinic project is funded by the US Environmental Protection Agency's Pollution Prevention Program. It addresses the two important national emphasis areas of (1) Business-based pollution prevention solutions

supporting the Toxic Substances Control Act (TSCA) Priorities and (2) Hazardous materials source reduction approaches in States or Communities. The overall goal of this project is to develop a computational software tool that can help the chemical industry minimize solvent waste from chemical processes. The research strategy for the proposed project has been divided into the following specific aims:

Aim#1: Collect information and consult industries about solvent recovery issues in current practices
Aim#2: Create a list of potential solvent recovery technologies based on information collected about solvent applicability, toxicity, and physicochemical properties.
Aim#3: Develop technology models comprising of mathematical equations involving material and energy balances, utility (electricity, cooling water) requirements, equipment design, and costing
Aim#4: Based on properties of the solvent rich stream, devise a ranked list of the best recovery pathway which minimizes cost, reduces environmental impact, and limits the waste discharge
Aim#5: Develop a user-friendly computer-aided software program for the solvent recovery roadmap

An example case study from the pharmaceutical industry is analysed, and the results are explained in section#3.

3 Results

3.1 Economic Evaluation of IPA Recovery from Pharmaceutical Waste Stream

Pfizer and Rowan University had carried out an investigation with aims to recover and purify isopropanol (IPA) and minimize waste from the celecoxib process, which produces the API for an arthritis pain medicine known as Celebrex® (Slater et al., 2012). The waste stream following the final purification stage contains a significant amount of IPA. However, the results of laboratory-scale distillation and extraction conducted at the plant site failed to reach the purity requirement (Slater et al., 2012). The case study is a classic representation of an API purification process. In a batch process, the celecoxib process required 4,205 kg of IPA/batch. If incineration is selected as the waste solvent disposal method, then approximately 14.51 kg of steam and 0.83 kWh of electricity/kg IPA is required. Life cycle analysis (LCA) has determined that there is 2.19 kg of total emission/kg of IPA used within the process (Slater et al., 2012).

Azeotropic points are anticipated at 87.7 wt.% and 80.37°C, which means that separation solely through distillation will not be able to achieve the desired purity. A summary of IPA recovery model specifications is provided in Table 1, where we assumed a waste stream feed basis of 1000 kg/hr.

Table 1: Isopropyl Alcohol (IPA) recovery case study model specification for optimization

Feed Conditions	Feed Rates (kg/hr)	Outlet Requirements (%)
IPA 51%	510	Recovery: 99.5% IPA
Water 49%	490	Purity 99%

The general equations for process streams, costs, energy requirements, and theories concerning technologies are composed of linear and non-linear equations. The selection or non-selection is represented via binary variables in the superstructure. This example is formulated as a mixed-integer non-linear programming (MINLP) problem and solved in the GAMS programming language through Branch-and-

Reduce Optimization Navigator (BARON) algorithm. Although solvent recovery is inherently a multi-stakeholder problem, we concentrated our objective toward only cost minimization. The optimized path is presented in Figure 2, with an annualized cost of $524,000 (i.e., 14 cents/kg solvent recovered) over 25 years (Chea et al., 2020). This pathway was able to reach the desired output specification from Table 1 and presents a solution with the lowest potential cost in comparison to other alternative pathways. Figure 2B presents the cost distribution of the optimal pathway. The annualized capital cost accounts for up to 47% of the total costs of the optimal pathway, followed by other costs (overhead), membrane replacement, labor, and utility. The price of selecting this pathway may be reduced further if the pervaporation and ultrafiltration units are available onsite for retrofit.

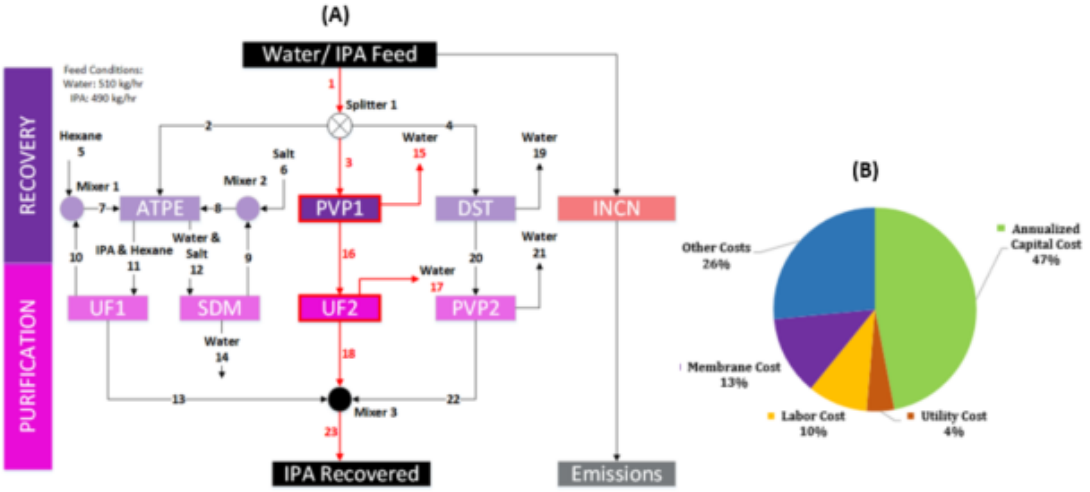


Figure 2: (A) A superstructure of the possible solvent recovery methods to separate IPA from the water. ATPE, UF, SDM, PVP, DST, and INCN represent aqueous two-phase extraction, ultrafiltration, sedimentation, pervaporation, distillation, and incineration, respectively. The most economically viable pathway for IPA recovery is highlighted in red. (B) The cost distribution of the optimal pathway (PVP1—UF2).

In comparison to incineration, solvent recovery is more economically viable. The cost required to incinerate the hypothetical waste flow rate of 1000 kg/hr requires $8.1 million/yr., which equates to $2.01/kg incinerated. The considerable increase in cost is attributed to the requirement for the heat of combustion. The organic solvent's chemical identity is irreversibly altered and thus cannot be reused within the process.

3.2 Environmental Impacts Assessment of IPA Recovery from Pharmaceutical Waste Stream

The environmental impacts of the optimized solvent recovery pathway were compared against conventional waste disposal methods. Sustainable process index (SPI) is an ecological footprint that measures the total arable area needed to embed a process into the ecosystem. SPI quantifies the environmental impacts of goods and services using material and energy flows. The primary assumption on which SPI is built on is that the natural source of environmental income to a sustainable economy is solar energy or radiation. Since the planet is finite, the area available to convert this income (solar radiation) into products and services is also finite. Therefore, the arable area needed to provide a service or goods is a convenient measure for the

SPI from an ecological sustainability point of view. Higher arable area needed to provide service goods corresponds to the increased impact on the ecosystem (Krotscheck and Narodoslawsky, 1996; Narodoslawsky, 2015; Narodoslawsky and Krotscheck, 1995, 2004). Human activities exert pressure on the ecosystem. To build up a process, humans depend on the ecosystem for resources such as both renewable and non-renewable energy, installation of equipment, and extraction of raw materials. Emissions are generated after the production of a product from a process. Therefore, an area in the ecosystem is needed to embed these air, water, and soil emissions aside from the areas needed for resource generation. The summation of these individual areas gives the total arable area needed to provide one unit of a product. Figure 3 shows the schematics for SPI.

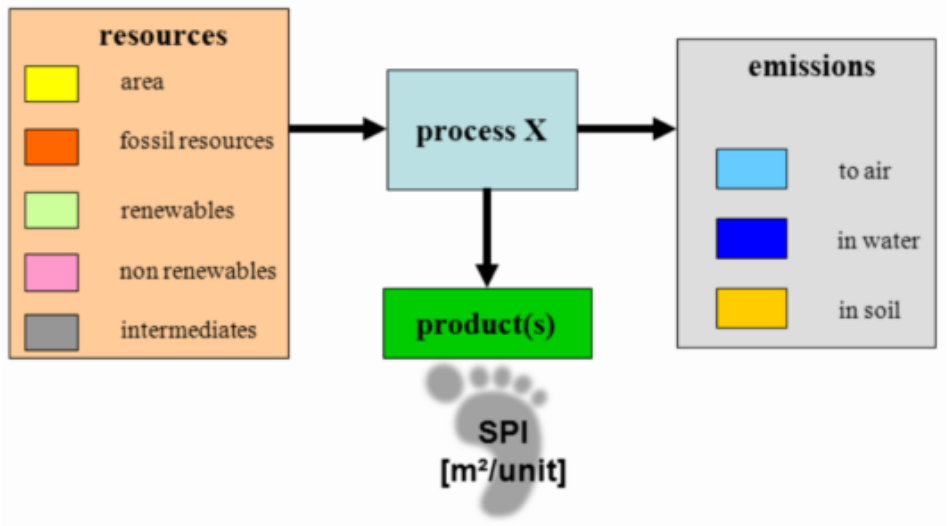


Figure 3. Schematics for sustainable process index. The resources are the inputs (quantified as arable area) to the process. Every process produces some emissions in the form of air, water, and soil. These emissions need to be embedded within an area in the ecosystem. The summation of these individual areas per unit of product(s) produced gives the SPI value.

The sustainability analysis for this case study was modelled using the sustainable process index footprint in SPIonWeb – an open-source software. For environmental impacts comparison, we considered three case scenarios, which include solvent recovery, direct disposal of the solvent waste into the environments, and incineration of the solvent waste. Table 2 shows the results for the case study from SPI analysis.

Table 2: Annual arable area (from SPI) needed to provide the services of direct disposal, solvent recovery, and incineration and the co_2 emissions and global warming potential associated with these services.

	SPI (m^2.a/unit)	**SPI (m^2.a/yr.)**	**CO_2 (kg/yr.)**	**Global Warming Potential (kg CO_2-eq/yr.)**
Direct Disposal	1988	8.03E+09	2.03E+07	2.17E+07
Solvent Recovery	128	4.93E+08	1.60E+06	1.69E+06
Incineration	405	3.21E+09	1.57E+07	1.71E+07

The total arable area needed for direct disposal and incineration supersedes that of solvent recovery by 93.9% and 84.6%, respectively. Thus, it will cost the ecosystem, an extra 93.9%, and 84.6% of natural income (arable area) if direct disposal and incineration were selected as the method of waste disposal. The annual CO_2 emission and global warming potential for both direct disposal and incineration supersede solvent recovery by 92.2% and 89.2%. Therefore, in all three scenarios, solvent recovery provides the best option for the treatment of hazardous waste.

Currently, we have completed the assessment of the economic and environmental impact separately, with greater emphasis on economics. If the cost of solvent recovery processes exceeds the price of common waste disposal methods significantly, then there is little incentive to choose recovery. Depending on the values of the company, more expensive recovery options may be chosen to minimize the overall environmental impacts. The next step in this work involves integrating this multi-objective complexity through the simultaneous modelling of both objectives using GAMS.

4 Summary

Through our unique engineering clinic program as well as synergistic efforts of the students, faculty, and staff at Rowan University, we were able to teach our students the importance of Sustainable Design in Chemical Engineering. In addition to project-based technical skills, the students also learned the importance of teamwork, technical writing, and presentation. Our students have presented this work at the AIChE (American Institute of Chemical Engineers) regional and national meetings, and in this process, they gained networking and communication skills. We believe that as engineering educators, it is our responsibility to teach the students the impact of systems-inspired design. Through all these activities, we were able to achieve our goals.

Acknowledgments

This research was supported by the United States Environmental Protection Agency's Pollution Prevention (P2) Program (NP96259218-0). The Design Engineering approach to education has been supported by an internal KEEN Curricular Reimagination Grant.

References

Biegler, L.T., Grossmann, I.E., Westerberg, A.W., 1997. Systematic methods of chemical process design. Prentice Hall PTR.

Cavanagh, E.J., Savelski, M.J., Slater, C.S., 2014a. Optimization of environmental impact reduction and economic feasibility of solvent waste recovery using a new software tool. Chemical Engineering Research and Design 92, 1942–1954. https://doi.org/10.1016/j.cherd.2014.02.022

Cavanagh, E.J., Savelski, M.J., Slater, C.S., 2014b. Optimization of environmental impact reduction and economic feasibility of solvent waste recovery using a new software tool. Chemical Engineering Research and Design, Green Processes and Eco-technologies 92, 1942–1954. https://doi.org/10.1016/j.cherd.2014.02.022

Chea, J., Lehr, A., Stengel, J., Savelski, M.J., Slater, C.S., Yenkie, K., 2020. Evaluation of Solvent Recovery Options for Economic Feasibility through a Superstructure-Based Optimization

Framework. Industrial & Engineering Chemistry Research. https://doi.org/10.1021/acs.iecr.9b06725

Chea, J.D., Christon, A., Pierce, V., Reilly, J.H., Russ, M., Savelski, M., Slater, C.S., Yenkie, K.M., 2019. Framework for Solvent Recovery, Reuse, and Recycling In Industries, in: Muñoz, S.G., Laird, C.D., Realff, M.J. (Eds.), Computer Aided Chemical Engineering, Proceedings of the 9 International Conference on Foundations of Computer-Aided Process Design. Elsevier, pp. 199–204. https://doi.org/10.1016/B978-0-12-818597-1.50032-1

Diwekar, U., 2013. Introduction to Applied Optimization. Springer Science & Business Media.

Krotscheck, C., Narodoslawsky, M., 1996. The Sustainable Process Index a new dimension in ecological evaluation. Ecological Engineering 6, 241–258. https://doi.org/10.1016/0925-8574(95)00060-7

Narodoslawsky, M., 2015. Chapter 3 - Sustainable process index, in: Klemeš, J.J. (Ed.), Assessing and Measuring Environmental Impact and Sustainability. Butterworth-Heinemann, Oxford, pp. 73–86. https://doi.org/10.1016/B978-0-12-799968-5.00003-8

Narodoslawsky, M., Krotscheck, C., 1995. The sustainable process index (SPI): evaluating processes according to environmental compatibility. Journal of Hazardous Materials, Selected papers presented at the Conference on Hazardous Waste Remediation 41, 383–397. https://doi.org/10.1016/0304-3894(94)00114-V

Narodoslawsky, M., Krotscheck, Ch., 2004. What can we learn from ecological valuation of processes with the sustainable process index (SPI) — the case study of energy production systems. Journal of Cleaner Production, Advances in cleaner production technologies 12, 111–115. https://doi.org/10.1016/S0959-6526(02)00184-1

Raymond, M.J., Slater, C.S., Savelski, M.J., 2010. LCA approach to the analysis of solvent waste issues in the pharmaceutical industry. Green Chemistry 12, 1826. https://doi.org/10.1039/c003666h

Slater, C.S., Savelski, M., Hounsell, G., Pilipauskas, D., Urbanski, F., 2012. Green design alternatives for isopropanol recovery in the celecoxib process. Clean Technologies and Environmental Policy 14, 687–698. https://doi.org/10.1007/s10098-011-0433-6

Slater, C.S., Savelski, M.J., Carole, W.A., Constable, D.J.C., 2010. Solvent Use and Waste Issues, in: Dunn, P.J., Wells, A.S., Williams, M.T. (Eds.), Green Chemistry in the Pharmaceutical Industry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 49–82. https://doi.org/10.1002/9783527629688.ch3

United Nations Environment Programme, 2019. Global Chemicals Outlook II - From Legacies to Innovative Solutions - Synthesis Report (2019), Global Chemicals Outlook II. United Nations Environment Programme.

US EPA, O., 2016. Global Mitigation of Non-CO2 Greenhouse Gases: Solvents [WWW Document]. US EPA. URL https://www.epa.gov/global-mitigation-non-co2-greenhouse-gases/global-mitigation-non-co2-greenhouse-gases-solvents (accessed 3.3.19).

Wypych, G., 2014. 13 - Solvent Use In Various Industries. Oxford: ChemTech Publishing, Handbook of Solvents 1–261.

Teaching Sustainability: Using Machine Learning to Predict Life Cycle Inventory data for Environmental Impact Assessment

Emmanuel A. Aboagye[1], Austin L. Lehr[1], C. Stewart Slater[1], Mariano J. Savelski[1], Robert P. Hesketh[1], Kirti M. Yenkie[1]*

[1]Department of Chemical Engineering, Rowan University, NJ, 08028
*yenkie@rowan.edu

Abstract

This paper explores the teaching of sustainability assessment and designs through the Engineering Clinics course offered by the College of Engineering at Rowan University. The study outlines the systematic methods involved, including how the project group is managed, student evaluation and assessment, and how to predict environmental impact metrics of chemicals using machine learning. The study further highlights the potential of machine learning techniques in promoting sustainability and developing sustainable chemicals and processes. Furthermore, it underscores the importance of sustainability education in equipping students with the necessary skills and knowledge for assessing the environmental impact of chemicals. Additionally, graduate students get an opportunity to engage in stakeholder interactions and mentoring while developing their research ability. Through the course, students work on real-world problems, such as the one described in this paper, using machine learning to predict the environmental impact metrics of chemicals. This project is essential because evaluating the environmental impacts of chemicals at early-stage process synthesis can be time-consuming and resource-intensive, making it challenging. Thus, through machine learning, impact assessment of chemicals can be performed at early-stage process synthesis to help choose chemicals with a less environmental burden. This challenging and relevant project teaches students skills such as information collection, computational programming, evaluation and prediction of sustainability metrics, technical writing, and oral presentation. Furthermore, students get the opportunity to present their research at regional and national conferences as well as students opting for a career in sustainability.

Keyword Sustainability, machine learning, global warming potential, environmental impact assessment, sustainability-teaching

1. Introduction

The Engineering Clinics (EC) offered by Henry M. Rowan College of Engineering at Rowan University is the hallmark of our undergraduate program, where students acquire hands-on experiences by solving real-world problems (Slater et al. 2005; Slater and Savelski, 2011). During the junior (third) and senior (fourth) years of undergraduate studies, the students have the opportunity to participate with stakeholders from industries and various federal agencies to work on challenging projects. Through the EC course, students can solve complex engineering problems by applying fundamental and advanced engineering concepts to produce solutions that meet specific needs with special consideration for public, health, environmental, and economic factors. Furthermore, the students are able to acquire hands-on experience regarding the systematic design of processes, interpretation, and analysis of data from experiments, and the ability to draw inferences based on sound engineering principles. Additionally, the students get the opportunity to present their research findings at national and regional conferences, thus improving their ability to write and communicate effectively to various audiences. In this paper, we have highlighted a specific project within the Chemical Engineering department at Rowan University where students implement Machine Learning (ML) approaches to predict the environmental impacts of chemicals from cradle-to-gate. Thus, students are introduced to the fundamental concepts of sustainability and the need to incorporate sustainable design principles at early-stage process design. In the subsequent sections, we discuss the background and motivation for choosing this project, the methodology implemented, some results from the project, and some concluding thoughts on the influence of this project on the environmental and social context for students.

1.1 Machine Learning for Environmental Impact Assessment

The advancement of technology has facilitated the exploration and identification of novel chemicals by the chemical industry (Johnson et al. 2020). However, the recent climate crisis has indicated the need to incorporate sustainable design at early-stage process synthesis to help improve the greenness of these processes. One of the main challenges associated with the sustainable design of these processes is the lack of life cycle inventories of novel and extant chemicals and their corresponding environmental impacts during the early-stage design synthesis (Argoti, Orjuela, and Narváez 2019; Karka, Papadokonstantakis, and Kokossis 2019; Papadokonstantakis et al. 2016). The absence of such metrics hinders the accurate evaluation of the environmental friendliness of these novel and extant chemicals and their potential as safer alternatives to commonly utilized chemicals during early-stage process synthesis. In recent decades, Machine Learning (ML) has garnered significant attention primarily owing to the predictive capabilities of its algorithms (Zhong et al. 2021).

As a result, the chemical industry has been presented with a fresh wave of possibilities in various domains of ML, such as supervised, unsupervised, and reinforcement learning (Yan, Borhani, and Clough 2020, 14). Hence, leveraging the ML approaches to our advantage, we can predict Life Cycle Impact Assessment (LCIA) metrics of novel and extant chemicals using their molecular descriptors and thermodynamic properties at the onset of process synthesis. In this work, we describe how students used machine learning for a project to predict the environmental impacts of chemicals. We begin by outlining the methods used to collect and preprocess data, including gathering thermodynamic and molecular properties data of chemicals and their corresponding environmental impact assessment data. We then describe how the students developed a supervised machine learning model using Python, which was trained using a dataset of the known chemicals and their corresponding environmental impact metrics. Finally, we present the study results, including the performance evaluation of the machine learning model and its potential implications for chemical development and sustainability. Through this project, we aim to create an appreciation for sustainable engineering by introducing students to the new trends in sustainable chemical process design using ML. Thus, we motivate students to pursue this career path.

2. Methodology

2.1 Project Team and Management

The clinic project team is comprised of 3-4 ambitious undergraduate students, a graduate student mentor, and experienced faculty advisors. The faculty advisors are accountable for establishing connections and networks with industries and other academic institutions and submitting research and educational grants to federal and regional organizations and industrial agencies. The faculty members further define the objectives and tasks associated with each objective and the overall learning outcomes for the students. The graduate student is responsible for mentoring and partially supervising undergraduate students during the project life cycle. Additionally, the student is responsible for documenting project progress and ensuring the overall continuity of the project till completion. The EC is a 2-credit course offered every semester with meeting times scheduled twice a week. The graduate student partially assigns weekly tasks to the undergraduate students. Furthermore, the students attend biweekly meetings with faculty advisors to present their findings based on the assigned tasks. During these biweekly meetings, the faculty advisors can provide sufficient directions for the research, highlight new tasks, and discuss challenges the students face. Thus, this undergraduate-graduate-faculty synergy provides adequate technical input for successful project completion.

2.2 Materials for Research and Student Assessment

The faculty advisors provide initial training regarding research tools and resources. At the beginning of the EC project, the advisors organize a one-day workshop where students, through the assistance of the graduate student, undergraduate students are taught how to perform a literature review to know the state-of-the-art trends and also how to collect relevant information about industrial processes, characterization metrics, chemical properties, and any other important modeling information relevant to the project. The students are introduced to platforms such as Google Scholar and ACS SciFinder. Furthermore, students are exposed to other web-based literature search platforms, such as Knovel, Web of Science, and Scopus, available through the Campbell Library of Rowan University. Additionally, students are trained to use citation managers such as Zotero and Mendeley to cite manuscripts. Furthermore, students are taught Python programming language, the standard platform for training the machine learning model. Additionally, students are exposed to using SimaPro for acquiring life cycle emissions and environmental impacts of processes and chemicals. At the end of the semester, students submit a comprehensive final report and also participate in a college-wide final presentation. Using a rubric grading system, the students are assessed based on their presentations and final report. During the presentation, students are evaluated on various categories such as project introduction effectiveness, organization and clarity of presentation slides, visual aids for effective communication, and overall handling of discussions.

2.3 Clinic Project: Sustainable Design of Chemical Processes

The funding source for the clinic project is the United States Environmental Protection Agency (USEPA) Bipartisan Infrastructure Law grant. The USEPA Pollution Prevention Program was established to help promote industries to implement more sustainable practices. In this project, students are trained to use computational modeling and machine learning approaches for sustainable process synthesis, safer alternative chemicals, and the environmental impact assessments of the processes and chemicals. The project is divided into four specific aims as follows:

Aim#1: Develop ML algorithms to estimate the life cycle inventories for new chemicals

Aim#2: Find environmental emissions scale-up factors for chemicals and technologies

Aim#3: Calculate operating parameters for greener technology options

Aim#4: Integrate aims#1-3 to design a case study, test and validate it via industrial consultations

In this paper, we focus on the first specific aim where students are trained on how to use Supervised Machine Learning (SML) to predict key environmental impact metrics such as Global Warming Potential (GWP), Human Health Impacts (HHI),

Ecosystem Quality Impacts (EQI), and Resource Utilization Impacts (RUI) as shown in Fig. 1.

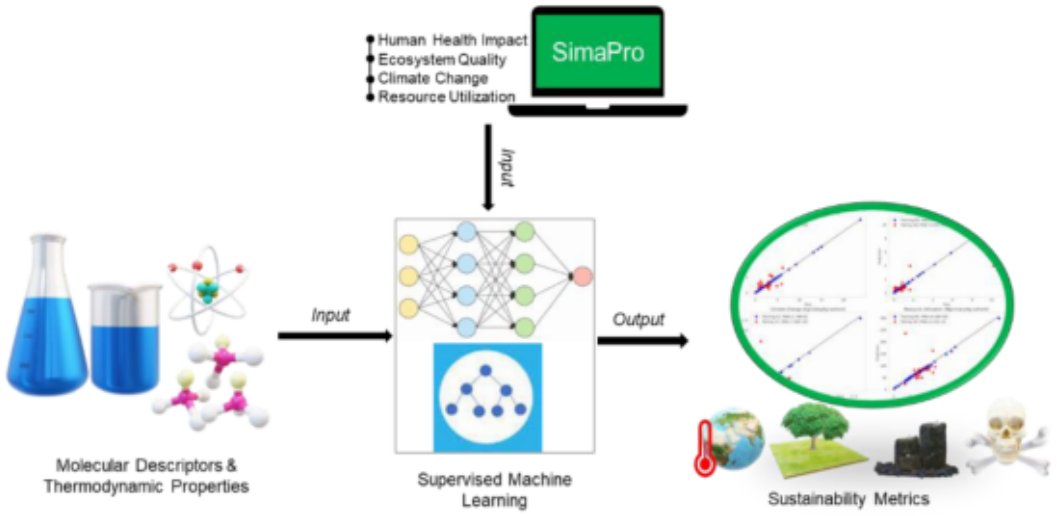


Fig 1. Framework for predicting environmental impacts of chemicals using ML.

Firstly, the students used available chemical databases such as PubChem, National Institute of Standards and Technology (NIST), and ChemSpider to acquire molecular descriptors and thermodynamic properties of existing chemicals. Overall, the students gathered 23 and 200 thermodynamic and molecular properties, respectively. Examples of the thermodynamic properties gathered are critical temperature, critical pressure, critical volume, acentric factor, saturated molar volume, and standard Gibbs free energy. Molecular weight, functional group, heavy atom count, number of aromatic rings, number of carbon, hydrogen, oxygen, nitrogen, sulfur atoms, radical electrons, and rotational bonds, among others, were some of the molecular properties. In the next step, the students use SimaPro, a well-recognized sustainability assessment software tool, to acquire data on human health, ecosystem quality, climate change, and resource utilization for each chemical. Using SimaPro, the students acquire tremendous experience in performing chemical life cycle analysis. Data was collected for over 350 chemicals. After the data collection process is completed, an analysis is performed to preprocess the data. During the preprocessing stage, data analysis is conducted for outlier detection, duplicated data detection, and missing data handling. All this analysis was done in Python using packages such as NumPy and Pandas. Upon completion of the data preprocessing and analysis, a supervised machine learning model is developed using the dataset of known chemical properties and their corresponding environmental impact metrics in Python. The algorithm used in this work is XGBoost (Extreme Gradient Boosting), a type of ensemble learning combining multiple decision trees to achieve high-accuracy predictions (Chen and Guestrin 2016; Ibrahem Ahmed Osman et al. 2021). It works by iteratively adding decision trees to a model. Each newly added tree focuses on reducing the residual between the predicted and actual values of the previous tree. One of the key advantages of using the is algorithm is its ability to help prevent overfitting via regularization and cross-validation, thus, improving the accuracy of

the model. The hyperparameters of the model are tuned using the hyperopt library (Bergstra, Yamins, and Cox 2013) to find the optimal parameter values, such as the number of trees and learning rate that maximizes the prediction accuracy.

3. Results and Discussion

Fig. 2 shows the predicted results for the four environmental impact metrics from the machine learning model. The model performance evaluation shows that both GWP (measured in kg-CO_2-eq/kg-chemical) and HHI (measured in Disability Adjusted Life-Years (DALY)/kg-chemical) for the testing set have a correlation coefficient greater than 0.7, with a root-mean-squared-error of 1.15 and 2.63 respectively. This shows that the model performs well at predicting these two metrics and can be used to predict these metrics for novel chemicals during early-stage process synthesis with an accuracy greater than 70%, thus, reducing the environmental footprint of chemical products and promoting the development of more sustainable and environmentally friendly alternatives. However, the ecosystem quality and resource utilization metrics need further hyperparameter tuning to improve their predictions.

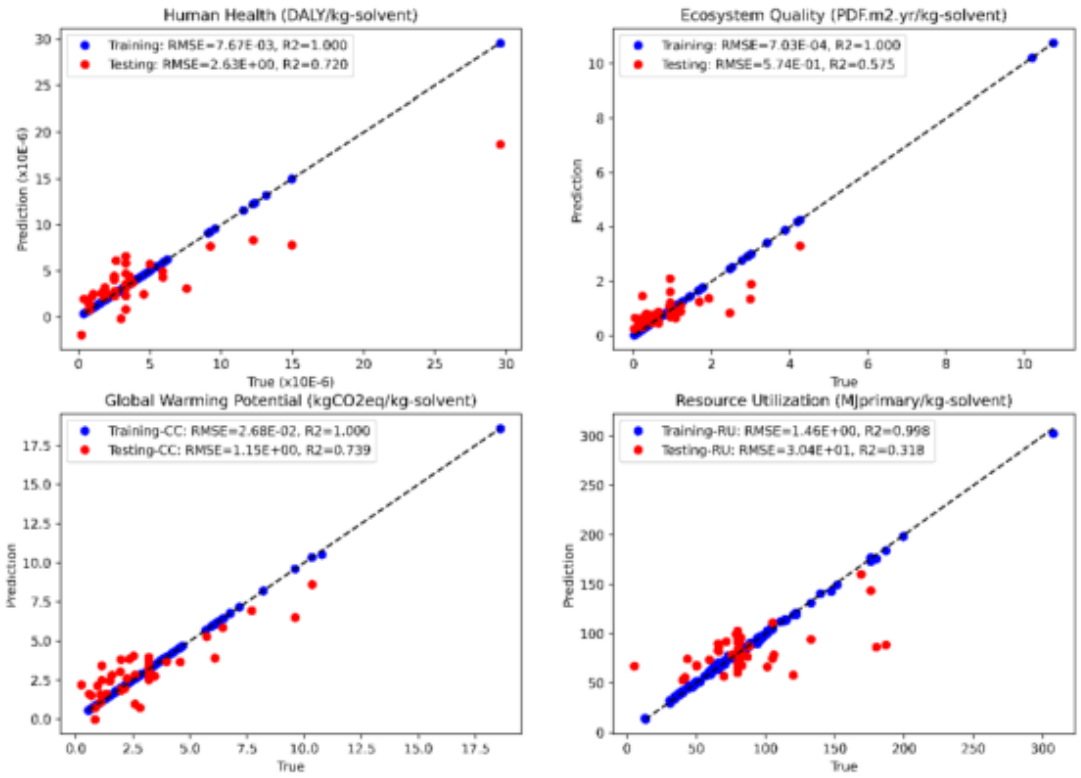


Fig 2. Parity plot for each metric showing the root-mean-squared-error (RMSE) and regression coefficient (R^2) for both training and testing set

4. Conclusion

In this work, we have demonstrated how through a step-by-step description approach, students gained hands-on experience in data collection, preprocessing, and machine learning to predict the environmental impacts of chemicals. This project further allowed students to develop critical skill sets that are essential in the field of environmental sustainability and impact assessment. The results of this study have important implications for the future of chemical development and sustainability. The experiences gained by the students in this study can prepare them for careers in environmental impact assessment and sustainability, where data analysis and machine learning techniques are becoming increasingly critical for informed decision-making. As the need for sustainable practices grows, future generations of chemical engineers must have the necessary skills to identify potential impacts and develop better solutions. Machine learning is just one example of the many techniques that can be used to promote sustainability, and it is essential that these techniques are integrated into the chemical engineering education curricula.

References

Argoti, Andres, Alvaro Orjuela, and Paulo C Narváez. 2019. "Challenges and Opportunities in Assessing Sustainability during Chemical Process Design." *Current Opinion in Chemical Engineering*, Energy, Environment & Sustainability: Sustainability Modeling Reaction engineering and catalysis: Green Reaction Engineering, 26 (December): 96–103. https://doi.org/10.1016/j.coche.2019.09.003.

Bergstra, James, Dan Yamins, and David Cox. 2013. "Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms." In , 13–19. Austin, Texas. https://doi.org/10.25080/Majora-8b375195-003.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. San Francisco California USA: ACM. https://doi.org/10.1145/2939672.2939785.

Ibrahem Ahmed Osman, Ahmedbahaaaldin, Ali Najah Ahmed, Ming Fai Chow, Yuk Feng Huang, and Ahmed El-Shafie. 2021. "Extreme Gradient Boosting (Xgboost) Model to Predict the Groundwater Levels in Selangor Malaysia." *Ain Shams Engineering Journal* 12 (2): 1545–56. https://doi.org/10.1016/j.asej.2020.11.011.

Johnson, Andrew C., Xiaowei Jin, Norihide Nakada, and John P. Sumpter. 2020. "Learning from the Past and Considering the Future of Chemicals in the Environment." *Science* 367 (6476): 384–87. https://doi.org/10.1126/science.aay6637.

Karka, Paraskevi, Stavros Papadokonstantakis, and Antonis Kokossis. 2019. "Predictive LCA - a Systems Approach to Integrate LCA Decisions Ahead of Design." In *Computer*

Aided Chemical Engineering, 46:97–102. Elsevier. https://doi.org/10.1016/B978-0-12-818634-3.50017-5.

Papadokonstantakis, S., P. Karka, Y. Kikuchi, and A. Kokossis. 2016. "Challenges for Model-Based Life Cycle Inventories and Impact Assessment in Early to Basic Process Design Stages." In *Sustainability in the Design, Synthesis and Analysis of Chemical Engineering Processes*, 295–326. Elsevier. https://doi.org/10.1016/B978-0-12-802032-6.00013-X.

Slater, C. S., Savelski, M. J. 2011. "Partnerships with the Pharmaceutical Industry to Promote Sustainability," *World Transactions on Engineering and Technology Education*, 9, 6-11.

Slater, C. S., Hesketh, R. P., Newell, J. A., Farrell, S., Gephardt, Z. O., Savelski, M. J., Dahm, K. D., Lefebvre, B.G. 2005. "ChE at Rowan University," *Chemical Engineering Education*, 39, 82-87.

Yan, Y., T. N. Borhani, and P. T. Clough. 2020. "Chapter 14:Machine Learning Applications in Chemical Engineering." In *Machine Learning in Chemistry*, 340–71. https://doi.org/10.1039/9781839160233-00340.

Zhong, Shifa, Kai Zhang, Majid Bagheri, Joel G. Burken, April Gu, Baikun Li, Xingmao Ma, et al. 2021. "Machine Learning: New Ideas and Tools in Environmental Science and Engineering." *Environmental Science & Technology*, August, acs.est.1c01339. https://doi.org/10.1021/acs.est.1c01339.