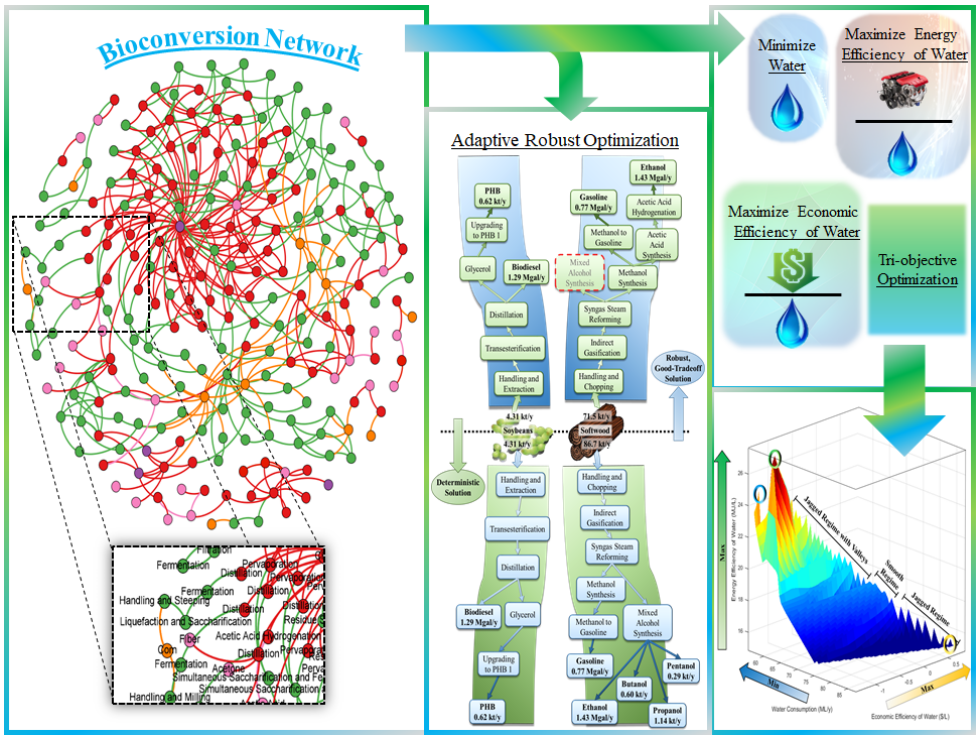
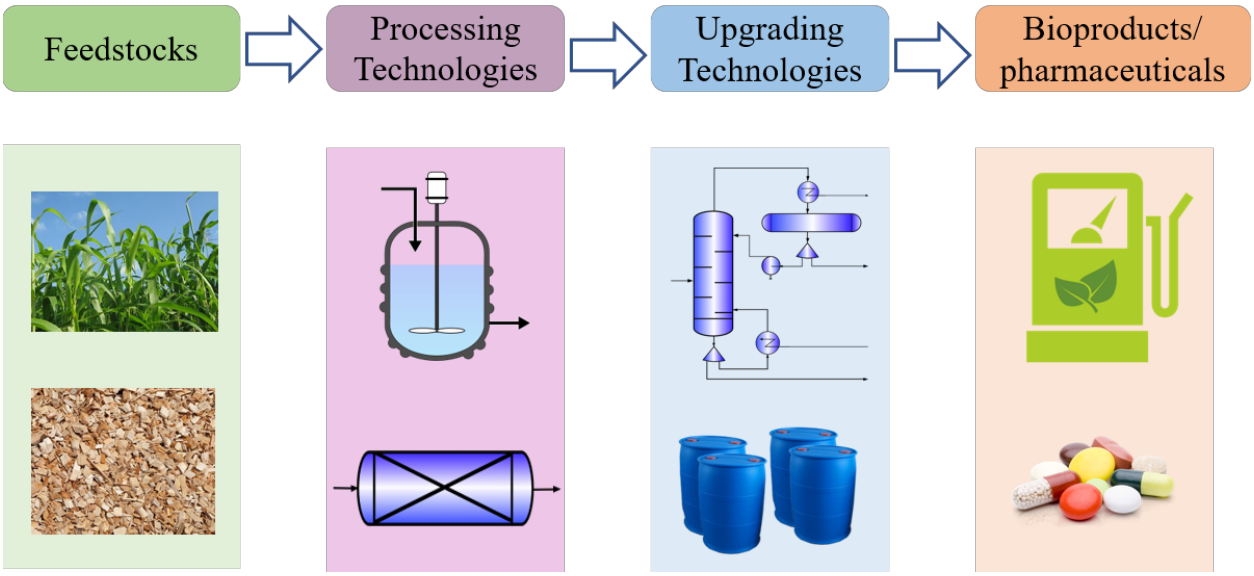
Data-Driven Distributionally Robust Optimization for Integrated Product-Process Design

Dr. Fengqi You, Roxanne E. and Michael J. Zak Professor
Process-Energy-Environmental Systems Engineering
fengq.you@cornell.edu | you.cbe.cornell.edu

With the development of machine learning and pattern recognition, intrinsic structures behind data can be efficiently extracted. However, in the past twenty years, such information has been primarily used for building regression and classification models, without speaking about what benefits can be brought to decision-making, which is of direct interest for industrial practitioners.



Product and process network design is a fundamental problem in process systems engineering. A complex product-process network usually consists of hundreds of feedstock, processing technologies, upgrading technologies and final products (e.g. see the figure above for an example of the bioconversion network). In a product-process network design problem, we need to determine technology pathways, sizing and operating levels of each process and the purchase and sale amounts. These decisions are heavily affected by various types of uncertainties, such as random prices and random demands. Big data themselves embody valuable information that could make a difference in decision-making, which is a promising topic worthy of our attentions.



This project aims to deal with decision-making problems from a data analytics point of view. Specifically, we adopt a new big data analytics strategy termed as distributionally robust optimization to address both the sequential nature of multi-stage uncertainties in the process-product network design problem, as well as the computational challenges induced by curse of dimensionality. Typically, a large number of data samples are used for extracting sufficient statistical information, which can be directly incorporated into optimization models. We will also investigate performances with different types of statistical information utilized. Case studies will be performed based on real-word demand data from an industrial-scale process-product network design problem.