

# Distributed Manufacturing for Electrified Chemical Processes in a Microgrid

Asha Ramanujam<sup>1</sup> | Gonzalo E. Constante-Flores<sup>1</sup> |  
Can Li<sup>1</sup>

<sup>1</sup>Davidson School of Chemical Engineering,  
Purdue University, 480 W. Stadium Ave,  
West Lafayette, IN, 47907 USA

## Correspondence

Can Li, Davidson School of Chemical  
Engineering, Purdue University, 480 W.  
Stadium Ave, West Lafayette, IN, 47907  
USA  
Email: canli@purdue.edu

## Funding information

The authors are grateful for the startup  
funding from the Davidson School of  
Chemical Engineering and the College of  
Engineering at Purdue University.

To alleviate the greenhouse gas emissions by the chemical industry, electrification has been proposed as a solution where electricity from renewable sources is used to power processes. The adoption of renewable energy is complicated by its spatial and temporal variations. To address this challenge, we investigate the potential of distributed manufacturing for electrified chemical processes installed in a microgrid. We propose a multi-scale mixed-integer linear programming model for locating modular electrified plants, renewable-based generating units, and power lines in a microgrid that includes monthly transportation and hourly scheduling decisions. We propose a K-means clustering-based aggregation disaggregation metaheuristic to solve the model efficiently. The model and algorithm are tested using a case study with 20 candidate locations in Western Texas. Additionally, we define a new concept, "the Value of the Multi-scale Model", to demonstrate the additional economic benefits of our model compared with a single-scale model.

## KEYWORDS

Electrification, Distributed Manufacturing, Multi-scale Integration, Mathematical Optimization

## 1 | INTRODUCTION

Chemical industries are a major source of greenhouse gas emissions and are responsible for 7% of the global greenhouse gas emissions [1] as they are mostly powered by the combustion of fossil fuels. To alleviate greenhouse gas emissions, electrification of the chemical industry is a promising solution that is currently being explored by chemical engineers [2, 3]. Electrification helps decarbonize the chemical industry and transition from fossil fuels to more renewable energy sources such as solar and wind. Electrification can be done in different parts of a process, including reactions, such as in electrochemical reactions [4], and heating, such as in resistive heating [3].

The major technical challenge to transitioning to renewable-based processes is that renewable resources, including solar and wind, have high spatial and temporal variations. A good way to take advantage of these variations is to adopt the concept of Distributed Chemical Manufacturing (DCheM). DCheM aims to improve chemical process industries by developing modular process plants, which take advantage of distributed resources and address distributed environmental problems. DCheM paves the way for introducing numerous new process technologies and simultaneously supports and enables energy and environmental sustainability while reducing chemical transportation costs. Modular processes increase the flexibility in dealing with the variability of conditions [5, 6, 7, 8].

While work has been done on the design and optimization of electrified plants and on the supply chain optimization of a particular set of chemicals produced by electricity, the combination of spatial and temporal variations (across different time scales) of renewable power-generating units has rarely been exploited with distributed manufacturing. Thus, there is a need to model both spatial and temporal features across time scales for electrified chemical processes and integrate the planning of the power resources with it. In this paper, we address this gap where we model a network of electrified chemical plants and power-generating units with three-time scales, i.e., single-time, monthly, and hourly. The problem is considered under the context of a microgrid, which typically consists of a network of low-voltage power-generating units, storage devices, and loads capable of supplying a local area such as a suburban area, an industrial park, or any commercial area with electric power and heat [9]. Microgrids support flexible operations, strengthen grid resilience and reduce energy losses making the electric power system efficient [10]. The location or point at which the microgrid is connected to the power utility is called the point of common coupling. Power can be purchased or sold to the power utility at the point of common coupling.

This research aims to design a network to facilitate DCheM for electrified chemical processes with the power demand satisfied by renewable sources and power from an external source coordinated by a microgrid operator by using a mixed-integer linear programming (MILP) model. The model incorporates both the microgrid (generating units and power lines) and chemical plant expansion in a single model. The transportation of chemicals and the transmission of power are two competing energy transportation approaches that are both captured in the lower-level decisions. Therefore, the tradeoffs between the transportation of chemicals and the transmission of power are studied. The model has more than millions of variables and constraints and is very difficult to solve directly using an MILP solver. To solve the model efficiently, we propose a tailored algorithm that combines mathematical programming and heuristics. Mathematical programming involves the study of techniques that can be used to generate provable optimal solutions to optimization problems. From a practical standpoint, the field of heuristics has a similar goal: to generate near-optimal and feasible solutions to optimization problems despite the lack of optimality guarantees. *Matheuristics* [11] are algorithms that combine mathematical programming and heuristics. Examples of matheuristics include local branching [12], and feasibility pump [13] that have been implemented in mixed-integer programming solvers such as CPLEX [14], and DICOPT [15]. Matheuristics are different from metaheuristic methods like the genetic algorithm and simulated annealing as matheuristic methods integrate mathematical programming techniques to construct solutions. The algorithm we propose in this paper is a decomposition-based matheuristic, where we break down the

problem into a sequence of subproblems where each subproblem is modeled as a mathematical program and solved to optimality.

The major contributions of this paper are listed below:

1. The proposed model encompasses three-time scales taking into account investment decisions, as well as monthly decisions and hourly decisions, and thus captures the temporal and spatial variations.
2. The size of the model can easily exceed millions of variables. We develop a tailored aggregation-disaggregation algorithm to solve the model efficiently, which is applicable to any multi-scale facility location problem.
3. We propose a new concept, "the Value of the multi-scale Model" (VMM), as a quantitative metric to characterize the additional economic savings of the multi-scale model compared with a single-scale model that ignores the temporal variations.
4. The proposed model and algorithms are made available as open-source software through our GitHub repository EDChem.jl with URL: <https://github.com/li-group/EDChem.jl>

The rest of this paper is organized as follows. In Section 2, a literature review of related works is provided. In Section 3, we describe the problem and the main assumptions. In Section 4, we present the problem formulation. In Section 5, we describe the proposed algorithm. In Section 6, the definition of the VMM is described. In Section 7, we illustrate the benefits of the proposed model and the effectiveness of the proposed solution techniques. The conclusions are drawn in Section 8.

## 2 | LITERATURE REVIEW

A number of works have been done in the optimization of the electrified production of various chemicals. Some work has been done with a focus on optimizing the design of processes, such as [16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. For example, Lazouski et al. [16] carried out a techno-economic analysis of the cost-optimal design and operation of a fully electrified ammonia generation process that utilizes intermittent renewable energy sources by modeling it as a MILP model. Sánchez and Martín [17] modeled the scale-up and scale-down of a renewable ammonia-based facility considering different technologies as a mixed-integer nonlinear programming (MINLP) model. Cooper et al. [18] developed a framework to optimize the design and operation of a large-scale electrolyzer hub under a variable power supply. Demirhan et al. [19] presented a process synthesis and global optimization framework for efficiently using renewable resources in ammonia production, where competing technologies were incorporated in a process superstructure.

Several works have been done with a focus on optimizing the scheduling of electrified processes, such as [26, 27, 28, 29, 30, 31, 32, 33, 34]. For example, Brée et al. [26] formulated and solved a model for the optimal scheduling of a chloralkali plant operating under different modes utilizing Demand Side Management. Allman and Daoutidis [27] developed a formulation for the optimal scheduling for a renewable ammonia plant. Bødal and M. Korpås [28] developed a two-stage stochastic model in a rolling horizon framework to consider uncertainties and used the model to study hydrogen production from electrolysis in a future scenario of a remote region in Norway with a focus on scheduling. Zheng et al. [29] modeled and optimized the operation of a grid-connected power-to-methanol system considering its flexibility.

For supply chain optimization of electrified plants, He et al. [35] developed a supply chain planning model that determines the least cost of hydrogen generation, storage, transmission, and compression facilities to meet hydro-

gen demand and is combined with power systems through electricity prices with the location of the renewable power-generating units fixed beforehand. Welder et al. [36] worked on potential future energy systems for power-to-hydrogen applications in Germany by developing a supply chain optimization model which takes the spatial and temporal resolution of the energy system into account with power through power lines not considered. These models consider the hourly demand as opposed to the monthly demand. Li et al. [37] proposed a co-planning approach for the regional wind resources-based ammonia industry and the electric network to optimize the wind-to-ammonia configuration and the expansion of the electrical network with ammonia demand being satisfied locally rather than transporting ammonia and for a one-time demand. Other papers such as [38, 39, 40, 41, 42] have also considered supply chain optimization of electrified plants. All these papers do not consider mode-based production of chemicals as well as local microgrid-based production, and most of them also do not consider multiple time scales and multiple products (chemicals) requirements in a network.

The co-expansion of the chemical and electricity network, like [37], is a less explored area. Another area that has not been studied in the electrification of the chemical industry is utilizing 3-time scales in the optimization model – single time, monthly and hourly. This is important because the output of renewable resources such as wind and solar varies hourly and by location, while the chemical delivery contract typically varies monthly. Multiple time scales have been explored in other areas, such as in the production of bioproducts [43] or in more general energy planning problems, such as [44], where the focus was more on central planning of the power utility as opposed to regional planning of microgrids.

As for algorithm development for solving integrated planning and scheduling models, Reinert et al. [45] proposed an algorithm to solve multi-time scale large-scale energy planning problems. Maravelias and Sung [46] reviewed techniques that can be used in integrating planning and scheduling. Li et al. [47] propose a tailored Benders decomposition algorithm for a multi-scale power systems planning problem. Allen et al. [48] propose improvements to decomposition algorithms by adding valid inequalities for multi-scale energy systems planning problems.

### 3 | PROBLEM STATEMENT

We first provide a general problem statement. An illustrative example shown in Figure 1 will be explained in detail after the generic description. We take the role of a company/conglomerate planning to invest in a regional microgrid or an industrial park. Our network has a set of consumers and suppliers of electrically produced chemical products, a point of common coupling (the location at which the microgrid is connected to the power utility), and a set of candidate locations. We can set up modular chemical plants or/and power-generating units at each candidate location. Modular chemical plants can be selected from several given technologies that involve electrochemical processes. For each plant, we are given the chemicals involved, the associated electrochemical reactions under different operating modes, and the equations to determine the power requirements. The modular power-generating units are all renewable, like solar panels and wind turbines. The capacity factors of the power-generating units for several historical years on an hourly basis are given. power lines can be installed between any two different locations or between any of the locations and the point of common coupling, whose location is predetermined. The time span of our problem is a given year.

Raw materials are obtained from suppliers with predetermined locations and transported to the installed chemical plants on a monthly basis. The required chemicals are produced in the plants and transported to consumers in certain given locations on a monthly basis. The electrochemical reactions in the plants consume the power obtained through the connected power lines. The consumed power can come from the installed generating units. In addition, when

the microgrid produces excess electricity, the excess electricity can be sold to the electric utility through the point of common coupling. On the other hand, the microgrid operator has the option of purchasing electricity from the electric utility as well. The monthly demand forecasts for each chemical for each of the consumers are given. In addition, we are given the resistance and inductance of the candidate power lines, the electricity price at each hour, the variable transportation costs, the capital cost of all the generating units, plants, and power lines, and the cost of different chemicals. We are also given the limits on production rates in different plants, as well as the capacity for power transmission in power lines, and the capacity for power generation of the generating units.

The proposed MILP model makes decisions across three-time scales: investment decisions at the beginning of the time horizon, monthly delivery of the chemical products to the customers as well as the monthly purchase of raw materials, and hourly operating decisions of chemical plants, power-generating units and power through power lines. To simplify the problem, we consider one year of operating decisions. The investment decisions include

- which chemical plants and power-generating units to install and the locations to install them,
- which power lines should be installed.

The monthly decisions include

- the amount of each chemical sold from each plant to each consumer,
- the amount of each chemical purchased from each supplier,
- the amount of consumer demand that is satisfied by other vendors,
- the amount of chemicals stored in the inventory of each plant at the end of each month.

The hourly operating decisions include

- the power flow of all the installed power lines,
- the amount of power required by each of the plants,
- the amount of each chemical produced in each plant,
- the mode at which each chemical plant operates,
- the amount of net electricity purchased/sold from/to the utility.

As an illustrative example, in the region considered, as shown in Figure 1, we are given three candidate locations denoted by rectangles. Three modular technologies, including chemical plants, wind turbines, and solar panels, are given, with the maximum number of each technology that can be installed in the network being one. Also, in the region shown, there are two consumers and two suppliers of raw materials and a point of common coupling whose locations are known. power lines can be installed between any two different locations and between any of the locations and the point of common coupling. A solution to the problem is shown in Figure 2. In the solution, one chemical plant is placed at R1, one solar panel at R3, and one wind turbine at R2 to meet the demand of the consumers most economically with appropriate power lines connected.

The assumptions we make are the following:

- The capacities of each type of modular plant and generating unit are assumed to be fixed to their standard values.
- At most, one reaction takes place in each electrochemical process at a particular time.
- The power lines in the network are of the same type i.e., have the same base voltage and base power.

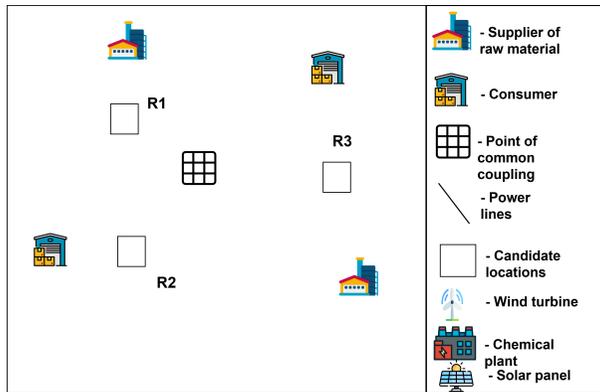


FIGURE 1 Region Representation

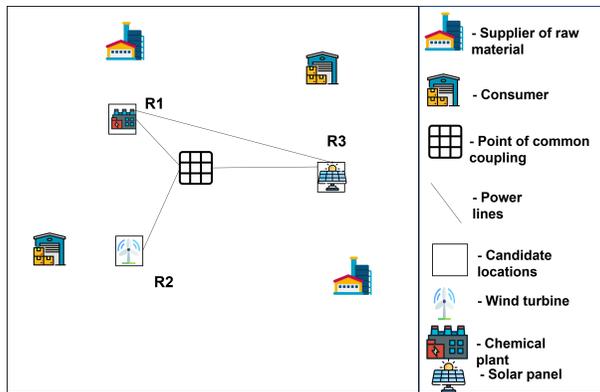


FIGURE 2 Solution Representation

- The chemical plant undergoes a mode-based operation.
- The time subperiods are hours.
- The transportation cost for transporting a specific amount of chemical is a linear function of the distance between the 2 locations and the amount of chemical transported.
- Chemicals are not transported between 2 plants
- Each consumer has a monthly demand. We do not have to satisfy the entire demand of the consumers with the chemicals produced by our installed chemical plants. However, a penalty cost higher than the chemical price is paid for the part of the demand satisfied by other vendors.
- The interface between the point of common coupling and the electric utility is set up, and the location of the point of common coupling is fixed.
- A chemical plant runs in different modes, and a chemical cannot be a reactant in one mode and a product in another mode for a chemical plant. In other words, a chemical is either a reactant for a plant or a product for the plant.
- Each plant has a shutdown mode where there is no production

- The differences between adjacent bus voltage angles are small.
- The chemical storage cost is assumed to be fixed and included in the capital cost of the plant.
- The ramping constraints, as well as constraints on the restrictions imposed on mode switching, which connect days, are relaxed.
- The transmission interconnection charges are neglected, i.e., the electricity from the microgrid can be sold at the point of common coupling at the market price without extra charges.
- The microgrid is assumed to be a price-taker in the electricity market, i.e., it does not impact the electricity price.

## 4 | MATHEMATICAL FORMULATION

### 4.1 | Temporal Simplification

Modeling each hour for each of the 365 days will make the problem extremely large. To reduce the number of variables while preserving the information in the data associated with the 365 days,  $k$  representative days can be selected in each month of the planning problem to represent the whole planning horizon where  $k \ll$  number of the days in the month, as shown in Figure 3. A clustering algorithm, such as  $k$ -means clustering or  $k$ -medoids clustering, can be used to divide the whole historical dataset into  $k$  clusters based on some of the characteristics of the historical days. The centroid or the medoid of each cluster is selected as the representative day. In this problem,  $k$  representative days are chosen for each month by applying a  $k$ -means clustering algorithm on the normalized power output of the renewable resources and electricity prices for the same month over several historical years [49].

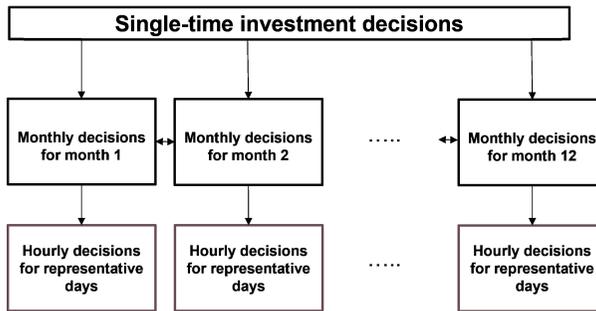


FIGURE 3 Decisions taken in the multi-scale model

### 4.2 | Notation

#### 4.2.1 | Indices and Sets

- $c \in C_i$  Set of chemicals, associated with plant  $i \in I_{plant}$ .
- $c \in C_j$  Set of chemicals supplied or demanded by  $j \in \mathcal{J}_c \cup \mathcal{J}_r$ .
- $i \in I_{plant}$  Set of plants.
- $i \in I_{power}$  Set of generating unit types.
- $j \in \mathcal{J}_c$  Set of consumers.
- $j \in \mathcal{J}_r$  Set of suppliers of raw materials.

$k \in \mathcal{K}_t$	Clusters for days of month $t \in \mathcal{T}$ (set of representative days in each month).
$(r, r') \in \mathcal{L}$	Set of candidate power lines between two different locations and between the locations and the point of common coupling (both ways included).
$m \in \mathcal{M}$	Set of modes for operating the plants.
$r \in \mathcal{R}$	Set of candidate locations to install the plants and generating units within the area considered.
$r_u$	A singleton that denotes the location of the point of common coupling.
$s \in \mathcal{S}$	Set of time periods (hours) within a day. We define the hours from 1 to 24 for a day, which represents 12 am to 11 pm.
$t \in \mathcal{T}$	Set of months of the year.

## 4.2.2 | Parameters

$C_c$	Price of chemical $c \in \cup_{i \in \mathcal{I}_{plant}} C_i$ . (\$/ton)
$C_{c,i,m}^{\min}$	Minimum production rate in mode $m$ of plant $i$ for $i \in \mathcal{I}_{plant}$ of chemical $c \in C_i$ . (ton/hr)
$C_{c,i,m}^{\max}$	Maximum allowable production rate increase/decrease in mode $m$ per time step in plant $i$ for chemical $c \in C_i$ . (ton/hr)
$C_{e,t,k,s}$	Electricity price during time $s$ on representative day $k$ of month $t$ . (\$/kWh)
$C_{f,i,r,t,k,s}$	Capacity factor of generator $i \in \mathcal{I}_{power}$ in location $r$ during time $s$ on representative day $k$ of month $t$ .
$C_{t,c}$	Cost of transportation of chemical $c \in \cup_{i \in \mathcal{I}_{plant}} C_i$ per unit distance per unit weight. (\$/(km · ton))
$D_{j,c,t}$	Demand of consumer $j \in \mathcal{J}_c$ for chemical $c \in C_j$ at the end of month $t$ . (ton)
$d_{r,r'}$	Distance between 2 locations $r, r'$ . (km)
$DIC_i$	Discounted investment cost of plant/power-generating unit $i \in \{\mathcal{I}_{plant} \cup \mathcal{I}_{power}\}$ . (\$)
$DIC_{(r,r')}$	Discounted investment cost of power line $(r, r') \in \mathcal{L}$ . (\$)
$F_{(r,r')}^{\max}$	Capacity of power line $(r, r') \in \mathcal{L}$ . (kW)
$FOC_i$	Fixed annual operating cost of plant/generating unit $i \in \{\mathcal{I}_{plant} \cup \mathcal{I}_{power}\}$ . (\$)
$FOC_{(r,r')}$	Fixed annual operating cost of line $(r, r') \in \mathcal{L}$ . (\$)
$G_{(r,r')}$	Conductance of power line $(r, r') \in \mathcal{L}$ . (pu)
$ir$	Annual interest rate. (%)
$mw_c$	Molecular weight of chemical $c \in \cup_{i \in \mathcal{I}_{plant}} C_i$ . (ton/mole)
$OCC_i$	Overnight capital cost of plant/power-generating unit $i \in \{\mathcal{I}_{plant} \cup \mathcal{I}_{power}\}$ . (\$)
$OCC_{(r,r')}$	Overnight capital cost of power line $(r, r') \in \mathcal{L}$ . (\$)
$P_{(r,r')}$	Active power in power line $(r, r') \in \mathcal{L}$ . (pu)
$pen$	Penalty given to consumers for not satisfying their demand.
$Q_{(r,r')}$	Reactive power in power line $(r, r') \in \mathcal{L}$ . (pu)
$Qg_i$	Nameplate (nominal) capacity of a generator $i \in \mathcal{I}_{power}$ . (kW)
$S_{i,c}$	Indicator parameter that denotes in plant $i \in \mathcal{I}_{plant}$ whether chemical $c \in C_i$ is a product or a reactant; equals to 1 if the chemical is a product, -1 if the chemical is a reactant.
$S_{base}$	Base power of power line $(r, r') \in \mathcal{L}$ . (kW)
$tl_i$	Lifespan of plant/power-generating unit $i \in \{\mathcal{I}_{plant} \cup \mathcal{I}_{power}\}$ . (years)
$tl_{(r,r')}$	Lifespan of of power line $(r, r') \in \mathcal{L}$ . (years)
$V_{max}$	Maximum voltage magnitude for any bus. (pu)
$V_{min}$	Minimum voltage magnitude for any bus. (pu)
$w_{k,t}$	Weight of representative day $k$ for month $t$ .

$x_i^{avail}$	Total number of resources $i \in \{I_{plant} \cup I_{power}\}$ available.
$x_i^{max}$	maximum number of resources $i \in \{I_{plant} \cup I_{power}\}$ which can be placed in any location.
$\alpha_{c,i,m}$	Stoichiometry of chemical $c \in C_i$ in plant $i \in I_{plant}$ in mode $m$ .
$\Delta C_{c,i,m}$	Maximum allowable production rate increase/decrease in mode $m$ per time step in plant $i$ for chemical $c \in C_i$ . (ton/hr)
$\theta_{i,m,m'}$	Minimum time required to stay in mode $m'$ after switching from $m$ to $m'$ for chemical plant $i \in I_{plant}$ . (hr)

### 4.2.3 | Discrete Variables

$nt_{(r,r')}$	Binary variable to indicate if power line $(r, r')$ is built.
$x_{i,r}$	Integer variable to indicate the number of component $i$ installed in location $r$ .
$Y_{i,r,m,t,k,s}$	Integer variable to indicate the number of plants $i \in I_{plant}$ installed in location $r$ in mode $m$ in time period $s$ on representative day $k$ in month $t$ .
$Z_{i,r,m,m',t,k,s}$	Integer variable to indicate the number of plants $i \in I_{plant}$ installed in location $r$ changing from mode $m$ to $m'$ in time period $s$ on representative day $k$ in month $t$ .

### 4.2.4 | Continuous Variables

$C_{r,t,k,s}^p$	Square of the voltage of bus $r \in \{R \cup r^u\}$ during time $t$ on representative day $k$ of month $t$ . (pu)
$C_{(r,r'),t,k,s}^v$	Product of the voltage of bus of the sending and receiving locations of the line $(r, r')$ during time $t$ on the representative day $k$ of month $t$ . (pu)
$F_{i,c,r,t,k,s}$	Sum of rate of production by all plants $i \in I_{plant}$ of chemical $c \in C_i$ in location $r$ in time period $s$ on representative day $k$ for month $t$ . (ton/hr)
$F_{i,c,m,r,t,k,s}^{mod}$	Sum of rate of production by all plants $i \in I_{plant}$ of chemical $c \in C_i$ in location $r$ in time period $s$ on representative day $k$ for month $t$ in mode $m$ . (ton/hr)
$P_{i,r,t,k,s}$	Power output of generator $i \in I_{power}$ in location $r$ in time period $s$ on representative day $k$ for month $t$ . (kW)
$P_{r,t,k,s}^{cu}$	Power curtailed in location $r$ in time period $s$ on representative day $k$ for month $t$ . (pu)
$P_{(r,r'),t,k,s}^{flow}$	Power transfer through line $(r, r')$ during time $s$ on representative day $k$ of month $t$ . (pu)
$P_{t,k,s}^{flow,ext}$	Power transfer from external source to microgrid during time $s$ on representative day $k$ of month $t$ . (pu)
$P_{i,r,t,k,s}^{plant}$	Sum of power required by all plants $i \in I_{plant}$ in location $r$ in time period $s$ on representative day $k$ for month $t$ for $i \in I_{plant}$ . (kW)
$Q_{i,c,r,t}$	Sum of inventory of all plants $i \in I_{plant}$ for chemical $c \in C_i$ in location $r$ at the end of month $t$ . (ton)
$S_{j,c,t}^I$	Amount of unserved demand for consumer $j$ of chemical $c \in C_i \cap C_j$ at the end of month $t$ . (ton)
$T_{i,c,r,j,t}^f$	Net sum of amount of chemical $c \in C_i \cap C_j$ transported from all $i \in I_{plant}$ in location $r$ to consumer $j$ at the end of month $t$ and negative of amount of chemical $c \in C_i \cap C_j$ transported to $i$ in location $r$ from supplier $j$ at the beginning of month $t$ for $i \in I_{plant}$ . (ton)
$V_{r,t,k,s}$	Magnitude of voltage of bus $r \in \{R \cup r^u\}$ during time $t$ on representative day $k$ of month $t$ . (pu)
$\phi^{material}$	Net revenue obtained from materials.
$\phi_{tr}$	Net transportation cost.
$\phi_{elec}$	Net electricity cost obtained from selling and buying electricity.
$\phi_{FIXOP}$	Net fixed operating cost.
$\phi_{CAPEX}$	Net amortized capital cost/

### 4.3 | Constraints

#### 4.3.1 | Constraints on chemical plants

The following constraints are applicable to chemical plants.

We assume that the chemical plants operate in different modes  $m \in \mathcal{M}$ . The following constraint (1) assures that each plant runs only in one operating mode at each time. This is ensured by equating the number of plants  $x_{i,r}$  to the sum of the variables indicating the number of plants in a particular mode,  $y_{i,r,m,t,k,s}$  over all the modes.

$$\sum_m y_{i,r,m,t,k,s} = x_{i,r}, \quad \forall i \in \mathcal{I}_{plant}, r, t, k \in \mathcal{K}_t, s. \quad (1)$$

The transition between different modes is modeled using the integer variable  $z_{i,r,m,m',t,k,s}$ , which represents the number of plants  $i \in \mathcal{I}_{plant}$  installed in location  $r$  changing from mode  $m$  to  $m'$  in time period  $s$  on the representative day  $k$  in month  $t$ . The constraint (2) is a way to determine the transition variables.

$$\sum_{m' \in \mathcal{M}} z_{i,r,m',m,t,k,s-1} - \sum_{m' \in \mathcal{M}} z_{i,r,m,m',t,k,s-1} = y_{i,r,m,t,k,s} - y_{i,r,m,t,k,s-1}, \quad \forall i \in \mathcal{I}_{plant}, r, s \geq 2, t, k \in \mathcal{K}_t, m. \quad (2)$$

A restriction is imposed that plants  $i \in \mathcal{I}_{plant}$  have to remain in a certain mode for a minimum amount of time after there is a transition to that mode using the following constraint (3)

$$y_{i,r,m,t,k,s} \geq \sum_{s'=1}^{\theta_{i,m',m}} z_{i,r,m',m,t,k,s-s'}, \quad \forall i \in \mathcal{I}_{plant}, r, t, k \in \mathcal{K}_t, m, m'. \quad (3)$$

with  $\theta_{i,m,m'}$  being the minimum time required to stay in mode  $m'$  after switching from  $m$  to  $m'$  for  $i \in \mathcal{I}_{plant}$ .

For a chemical plant,  $F_{i,c,r,t,k,s}$  represents the total rate of production in all plants of chemical  $c \in C_i$  in location  $r$ .  $F_{i,c,m,r,t,k,s}^{\text{mod}}$  represents the sum of rate of production of chemical  $c \in C_i$  in mode  $m$ .  $F_{i,c,r,t,k,s} > 0$  for  $c$  being a product and  $F_{i,c,r,t,k,s} < 0$  for  $c$  being a reactant. Similarly,  $F_{i,c,m,r,t,k,s}^{\text{mod}} > 0$  for  $c$  being a product and  $F_{i,c,m,r,t,k,s}^{\text{mod}} < 0$  for  $c$  being a reactant.

Let the reaction be  $\sum_c \alpha_{c,i,m} c = 0$ , where  $\alpha_{c,i,m}$  represents the stoichiometry of chemical  $c \in C_i$  in plant  $i$  in mode  $m$ . The following constraint (4) describes the stoichiometry for all  $c, c' \in C$  such that  $c, c'$  are in the reaction for plant  $i$  operating in mode  $m$ .

$$\frac{F_{i,c,m,r,t,k,s}^{\text{mod}}}{mw_c} \alpha_{c',i,m} = \frac{F_{i,c',m,r,t,k,s}^{\text{mod}}}{mw_{c'}} \alpha_{c,i,m}, \quad \forall c, c' \in C_i, m, t, k \in \mathcal{K}_t, s. \quad (4)$$

The production rate of chemical  $c$  in plant  $i$  is the summation of production over all modes and is described by equation (5).

$$\sum_{m \in \mathcal{M}} F_{i,c,m,r,t,k,s}^{\text{mod}} = F_{i,c,r,t,k,s}, \quad \forall i \in \mathcal{I}_{plant}, c \in C_i, r, t, k \in \mathcal{K}_t, s. \quad (5)$$

For each plant, there is a base chemical denoted by  $c_{base,i}$ . By default a product is taken as the base chemical. The constraints (6),(7), and (8) are formulated on the base chemical of each plant. This is because the production rates of all the chemicals are stoichiometrically related by (4).

The following constraint (6) imposes a restriction on the maximum and minimum rate of production or consump-

tion of the base chemical in all the plants in each mode.  $C_{c,i,m}^{\max}$  and  $C_{c,i,m}^{\min}$  represent the maximum and minimum production in mode  $m$  of chemical  $c \in C_i$  respectively.

$$C_{c_{base},i,m}^{\min} Y_{i,r,m,t,k,s} \leq F_{i,c_{base},m,r,t,k,s}^{\text{mod}} \leq C_{c_{base},i,m}^{\max} Y_{i,r,m,t,k,s}, \quad \forall i \in I_{plant}, r, t, k \in \mathcal{K}_t, s, m. \quad (6)$$

Constraints (7) and (8) bound the change in production level when there is no transition in modes in the plants of a particular location and all the plants are in the same mode. The maximum allowable production increase/decrease in mode  $m$  per time step in plant  $i$  for chemical  $c \in C_i$  is given by  $\Delta C_{c,i,m}$  and  $M$  is a sufficiently large number.

$$F_{i,c_{base},m,r,t,k,s+1}^{\text{mod}} \leq F_{i,c_{base},m,r,t,k,s}^{\text{mod}} + \Delta C_{c_{base},i,m} Y_{i,r,m,t,k,s} + M(2x_{i,r} - Y_{i,r,m,t,k,s+1} - Y_{i,r,m,t,k,s}), \quad \forall i \in I_{plant}, r, t, k \in \mathcal{K}_t, m, s <= 23, \quad (7)$$

$$F_{i,c_{base},m,r,t,k,s+1}^{\text{mod}} \geq F_{i,c_{base},m,r,t,k,s}^{\text{mod}} - \Delta C_{c_{base},i,m} Y_{i,r,m,t,k,s} - M(2x_{i,r} - Y_{i,r,m,t,k,s+1} - Y_{i,r,m,t,k,s}), \quad \forall i \in I_{plant}, r, t, k \in \mathcal{K}_t, m, s <= 23. \quad (8)$$

Additionally, the constraint that the production rates of all chemicals when the plant is in shut down mode is zero is added, i.e., when  $m = \text{Shut down}$ ,  $F_{i,c,m,r,t,k,s}^{\text{mod}} = 0$  for all the chemicals in the plant.

The following constraints (9) and (10) describe the mass balance of chemicals in each location. The inventory level of a chemical at a given month  $t$ , denoted by  $Q_{i,c,r,t}$  which is non-negative, equals the inventory at the previous month plus the summation of chemicals produced during this month minus the summation of transportation of the chemical out of this location (denoted by  $T_{i,c,r,j,t}^r$ ).

$$Q_{i,c,r,1} = \sum_{k \in \mathcal{K}_1} \sum_{s=1}^{24} w_{k,1} F_{i,c,r,t,k,s} - \sum_{j|c \in C_j \cap C_i} T_{i,c,r,j,1}^r, \quad \forall i \in I_{plant}, c \in C_i, r, \quad (9)$$

$$Q_{i,c,r,t} = \sum_{k \in \mathcal{K}_t} \sum_{s=1}^{24} w_{k,t} F_{i,c,r,t,k,s} + Q_{i,c,r,t-1} - \sum_{j|c \in C_j \cap C_i} T_{i,c,r,j,t}^r, \quad \forall i \in I_{plant}, c \in C_i, r, t \geq 2. \quad (10)$$

Constraints (11) ensure that we try to satisfy the demand of each consumer  $j$ ,  $D_{j,c,t}$ . The part of the demand not met for the consumer,  $S_{j,c,t}^l$  is penalized in the objective function.

$$\sum_{i \in I_{plant}} \sum_{c \in C_i} \sum_r T_{i,c,r,j,t}^r + S_{j,c,t}^l = D_{j,c,t}, \quad \forall j \in \mathcal{J}_c, t, c \in C_j. \quad (11)$$

Constraint (12) represent the total power  $P_{i,r,t,k,s}^{\text{plant}}$  required by plants where  $f_{c,m,i}(F)$  is the linear function connecting the power requirement to the production rate for each chemical in a particular mode

$$P_{i,r,t,k,s}^{\text{plant}} = \sum_{c \in C_i} \sum_m f_{c,m,i}(F_{i,c,r,t,k,s}^{\text{mod}}), \quad \forall t, k \in \mathcal{K}_t, s, r, i \in I_{plant}. \quad (12)$$

Strictly speaking, the power requirement is a nonlinear function of the production rate due to overpotential. Here, we use a linear approximation [50] that is accurate for our chlorine cast study [26]. When the linear approximation is inaccurate due to overpotential, we can change (12) to a piecewise linear approximation shown in [16].

### 4.3.2 | Modelling of Generating Units

The power output of a generator in a particular location,  $p_{i,r,t,k,s}$  is the product of the nameplate (nominal) capacity of the generator  $Qg_{i,r}$  and the capacity factor of generator  $Cf_{i,r,t,k,s}$ . The power output of generating unit  $i$  is modeled as follows:

$$p_{i,r,t,k,s} = Qg_{i,r} Cf_{i,r,t,k,s} X_{i,r}, \quad \forall i \in I_{power}, r, t, k \in \mathcal{K}_t, s. \quad (13)$$

### 4.3.3 | Constraints on Power Balance in the microgrid

The power transferred through line  $l$  is denoted by  $p_{l,t,k,s}^{flow}$  and is expressed in per unit. The per-unit system is widely used in the field of electrical engineering to express values of currents, voltages, powers, and impedances of various power equipment [51]. The quantities are divided by a corresponding base quantity to get a value in per unit (pu). The power balance in each candidate location is enforced by the following constraints:

$$\sum_{i \in I_{power}} p_{i,r,t,k,s} X_{i,r} - \sum_{i \in I_{plant}} P_{i,r,t,k,s}^{plant} - S_{base} p_{r,t,k,s}^{cu} = S_{base} \sum_{(r,r') \in \mathcal{L}} p_{(r,r'),t,k,s}^{flow}, \quad \forall t, k \in \mathcal{K}_t, s, r \in \mathcal{R}. \quad (14)$$

Constraints (15) represent the power balance in the point of common coupling

$$p_{t,k,s}^{flow,ext} - \sum_{(r_u,r) \in \mathcal{L}} p_{(r_u,r),t,k,s}^{flow} = 0, \quad \forall t, k \in \mathcal{K}_t, s, \quad (15)$$

where  $p_{t,k,s}^{flow,ext}$  is the power transferred from the electric utility to the microgrid or the negative of the power transferred from the microgrid to the electric utility.

### 4.3.4 | Constraints on power lines

The governing physical laws describing the flow of electricity, known as AC power flow equations, in an electric grid are quadratic and non-convex. Generally, such equations are linearly approximated based on certain valid operational assumptions, i.e., the reactance to resistance (X/R) ratio of transmission lines is generally between 2 and 10. The most common linear approximation is known as the DC power flow, which has been used for decades to plan and operate high-voltage transmission systems [47]. However, for electric utilities, such an approximation is not applicable since the X/R ratio of electric utility lines is significantly smaller than its high-voltage transmission counterpart.

We derive an approximation of the power flows in a microgrid based on the following considerations: (i) the reactive power is neglected, (ii) the angle difference between two connected nodes is small, and (iii) the resistance of power lines is greater than their reactance. The proposed approximation of the power flow equations is presented as follows:

$$p_{(r,r'),t,k,s}^{flow} = nt_{(r,r')} G_{(r,r')} \left( C_{r,t,k,s}^p - C_{(r,r'),t,k,s}^v \right), \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (16)$$

$$C_{r,t,k,s}^p = V_{r,t,k,s}^2, \quad \forall r, t, k \in \mathcal{K}_t, s, \quad (17)$$

$$C_{(r,r'),t,k,s}^v = V_{r,t,k,s} V_{r',t,k,s}, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (18)$$

where  $C_{r,t,k,s}^p$  and  $C_{(r,r'),t,k,s}^v$  are auxiliary variables encoding the quadratic terms in the power flow constraints (16),

and  $nt_{(r,r')}$  is a binary variable that indicates whether power line  $(r, r')$  is installed. The derivation of the proposed approximation is described in Appendix 1.

Constraints (16)-(18) are nonconvex due to (i) the products of binary and continuous variables in constraints (16) and (ii) the quadratic terms in constraints (17) and (18).

The nonconvexities in (16), the product of binary and continuous variables, are linearized using a Big-M formulation as follows:

$$-M(1 - nt_{(r,r')}) \leq \rho_{(r,r'),t,k,s}^{\text{flow}} - G_{(r,r')} (C_{r,t,k,s}^{\text{P}} - C_{(r,r'),t,k,s}^{\text{V}}) \leq M(1 - nt_{(r,r')}), \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s. \quad (19)$$

The tightness of the LP relaxation of the proposed MILP model depends on the value of  $M$ . Here, we choose  $M$  to be a number greater than the maximum absolute possible value of  $G_{(r,r')} (C_{r,t,k,s}^{\text{P}} - C_{(r,r'),t,k,s}^{\text{V}})$ .

The nonconvexities in (17)-(18), products of continuous variables, are relaxed using McCormick envelopes [52] as follows:

$$C_{r,t,k,s}^{\text{P}} \geq 2V_{r,t,k,s}V_{\min} - V_{\min}^2 \quad \forall r, t, k \in \mathcal{K}_t, s, \quad (20a)$$

$$C_{r,t,k,s}^{\text{P}} \geq 2V_{r,t,k,s}V_{\max} - V_{\max}^2, \quad \forall r, t, k \in \mathcal{K}_t, s, \quad (20b)$$

$$C_{r,t,k,s}^{\text{P}} \leq V_{r,t,k,s}(V_{\min} + V_{\max}) - V_{\min}V_{\max}, \quad \forall r, t, k \in \mathcal{K}_t, s, \quad (20c)$$

$$C_{(r,r'),t,k,s}^{\text{V}} \geq V_{r,t,k,s}V_{\min} + V_{r',t,k,s}V_{\min} - V_{\min}^2, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (20d)$$

$$C_{(r,r'),t,k,s}^{\text{V}} \geq V_{r,t,k,s}V_{\max} + V_{r',t,k,s}V_{\max} - V_{\max}^2, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (20e)$$

$$C_{(r,r'),t,k,s}^{\text{V}} \leq V_{r,t,k,s}V_{\min} + V_{r',t,k,s}V_{\max} - V_{\min}V_{\max}, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (20f)$$

$$C_{(r,r'),t,k,s}^{\text{V}} \leq V_{r,t,k,s}V_{\max} + V_{r',t,k,s}V_{\min} - V_{\min}V_{\max}, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (20g)$$

where the bounds on the voltage magnitudes  $V_{\min}$  and  $V_{\max}$  are known parameters. In particular, we consider that the voltage magnitudes vary  $\pm 5\%$  of the nominal voltage. It is noteworthy that even though the McCormick relaxation (20) is often not tight, it provides a proxy of the power flow equations, which balances computational tractability and model accuracy. Additionally, it is important to note that such a proxy is only used as a simplified model for the hourly operation of the representative days to determine long-term investment decisions.

We ensure that the power flows of installed power lines are bounded by their capacity. If a power line is not installed, its corresponding flow is enforced to be zero. The power flow bounds are enforced as follows:

$$-F_{(r,r')}^{\text{max}} nt_{(r,r')} \leq \rho_{(r,r'),t,k,s}^{\text{flow}} \leq F_{(r,r')}^{\text{max}} nt_{(r,r')}, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s. \quad (21)$$

Additionally, constraint (22) is also added to ensure that the sum of power loss contributions for a line is positive and constraint (23) to ensure that there is the same number of lines both ways.

$$\rho_{(r,r'),t,k,s}^{\text{flow}} + \rho_{(r',r),t,k,s}^{\text{flow}} \geq 0, \quad \forall (r, r') \in \mathcal{L}, t, k \in \mathcal{K}_t, s, \quad (22)$$

$$nt_{(r,r')} = nt_{(r',r)}, \quad \forall (r, r') \in \mathcal{L}. \quad (23)$$

### 4.3.5 | Additional constraints

Several constraints corresponding to nontrivial variable bounds are added. The number of plants and generating units that can be installed in each location is bounded. The inventory level  $Q_{i,c,r,t}$  is bounded by its maximum capacity. Other constraints corresponding to the non-negativeness of the variables such as the slack variables  $Sl_{j,c,t}$  and the variables corresponding to the power curtailed  $p_{r,t,k,s}^{cu}$  as well as the sign of the transportation variables  $Tr_{i,c,r,j,t}$  are also added. The equations of these constraints are skipped due to their simplicity. Additionally, there could be constraints on the investment decisions  $x_i$  based on the availability of resources  $x_i^{avail}$ . We can write these additional constraints using (24).

$$\sum_r x_{i,r} \leq x_i^{avail} \quad \forall i \in \{I_{plant} \cup I_{power}\}. \quad (24)$$

In certain instances, constraints may be required on the maximum number of plants or power-generating units allowable in a particular location, represented by equation (25).

$$x_{i,r} \leq x_i^{max} \quad \forall i \in \{I_{plant} \cup I_{power}\}. \quad (25)$$

Lastly, to ensure no plant or power-generating unit exists without a connected power line, we introduce constraint(26):

$$\sum_{i \in \{I_{plant} \cup I_{power}\}} x_{i,r} \leq \left( \sum_{i \in \{I_{plant} \cup I_{power}\}} x_i^{avail} \right) \left( \sum_{(r,r') \in \mathcal{L}} (nt_{(r,r')} + nt_{(r',r)}) \right) \quad \forall r. \quad (26)$$

## 4.4 | Objective Function

The objective is to maximize the total annual profit, which is the difference between revenue and cost. We try to meet the demand of each consumer. For the part of the demand unmet  $Sl_{j,c,t}$ , we pay a penalty of  $pen.C_c Sl_{j,c,t}$ . The net revenue from materials in the microgrid is given by equation (27).

$$\phi_{material} = \sum_t \left( \sum_{i \in I_{plant}} \sum_r \sum_j \sum_{c \in C_i \cap C_j} C_c Tr_{i,c,r,j,t}^r \right) - \sum_t \left( \sum_j \sum_{c \in C_j \cap C_j} pen.C_c Sl_{j,c,t} \right). \quad (27)$$

The transportation cost is calculated as a linear function of the distance between the 2 locations, and the amount of chemicals transported between the 2 locations. The total transportation cost is given by equation (28).

$$\phi_{tr} = \sum_t \sum_j \sum_r \sum_{i \in I_{plant}} \sum_{c \in C_i \cap C_j} C_{tc} Tr_{i,c,r,j,t}^r d_{r,j} S_{i,c}. \quad (28)$$

The total electricity cost (or negative of total revenue obtained from selling electricity) of the microgrid is given by equation (29).

$$\phi_{elec} = \sum_t \sum_{k \in \mathcal{K}_t} \sum_{s=1}^{24} w_{k,t} C_{e,t,k,s} S_{base} p_{t,k,s}^{flow,ext}. \quad (29)$$

The storage cost of chemicals is included in the capital cost by assumption.

The fixed operating costs for plants, generators, and power lines are given by equation (30). The costs of power lines are divided by 2 to prevent double counting of power lines as the same number of power lines go both ways.

$$\phi_{FIXOP} = \sum_i \sum_r FOC_i x_{i,r} + \frac{\sum_{(r,r') \in \mathcal{L}} FOC_{(r,r')} nt_{(r,r')}}{2}. \quad (30)$$

The discounted (amortized) investment cost of  $i \in \{I_{plant} \cup I_{power}\}$ ,  $DIC_i$  is given by equation (31).

$$DIC_i = \left( \frac{OCC_i ir}{\left(1 - \frac{1}{(1+ir)^{t_i}}\right)} \right). \quad (31)$$

The discounted (amortized) investment cost of power line  $(r, r') \in \mathcal{L}$ ,  $DIC_{(r,r')}$  is given by equation (32).

$$DIC_{(r,r')} = \left( \frac{OCC_{(r,r')} ir}{\left(1 - \frac{1}{(1+ir)^{t_{(r,r')}}}\right)} \right). \quad (32)$$

The total amortized capital cost is given by equation (33).

$$\phi_{CAPEX} = \sum_i \sum_r DIC_i x_{i,r} + \frac{\sum_{(r,r') \in \mathcal{L}} DIC_{(r,r')} nt_{(r,r')}}{2}. \quad (33)$$

The entire optimization model is summarized in (34)

$$\max \quad \phi_{material} - \phi_{tr} - \phi_{elec} - \phi_{FIXOP} - \phi_{CAPEX}, \quad (34a)$$

$$\text{s.t.} \quad (1) - (15), (19) - (33), \quad (34b)$$

$$\text{Constraints on variable bounds.} \quad (34c)$$

## 5 | ALGORITHM

### 5.1 | Overview

The model presented in Section 4 can easily have millions of variables. Even for a small number of locations (5~10), it is very difficult to solve directly using an MILP solver. In addition, decomposition algorithms such as Benders decomposition and Lagrangian decomposition cannot be applied to solve this problem efficiently. We provide the details of the applicability of existing decomposition methods in Supplementary B. To solve it efficiently, we propose a matheuristic to obtain a good feasible solution in a reasonable amount of time. Before presenting the details of our algorithm, we provide an overview of the algorithm to help the reader build some intuition. The algorithm has two major steps. The first step is to aggregate variables such that a smaller MILP is solved. The second step is a disaggregation heuristic to find a feasible solution to the original problem.

For the aggregation step, the integral operating variables are initially relaxed and the locations are grouped into a smaller number of clusters by applying the k-means algorithm on their coordinates. The variables of the locations belonging to the same cluster are then aggregated and the cluster centers are used as candidate locations. Thus the

problem has been reduced from a large number of candidate locations to a problem with a reasonable number of candidate locations that can be solved using an MILP solver in a reasonable amount of time. From this, the number of plants and power-generating units in each cluster, as well as a fair idea of the number of power lines between any two cluster centers, can be obtained. Figure 4c represents an example of the investment decisions obtained from the aggregated problem.

In the disaggregation step, an MILP to disaggregate each individual cluster is solved while fixing most of the variables of the rest of the clusters including the number of plants, power-generating units, and the power lines among them. Note that the clusters are disaggregated independently and not sequentially here. Therefore, these disaggregating MILPs can be parallelized.

After completely disaggregating each cluster, the number of plants and the number of power-generating units in each location can be fixed. However, the number of power lines between any two locations is not fully determined since we are disaggregating each cluster independently. Essentially, the solution to the disaggregation MILP has two types of power lines: a) power lines between a candidate location and the point of common coupling or between two candidate locations that are within the same cluster in the disaggregating MILPs and b) power lines between a candidate location and a cluster center. Note that the second type of power line does not correspond to a line in the original problem. We propose a heuristic to obtain the number of power lines between any two locations based on the number of the second type of lines from the disaggregating MILP. To this end, an integer program is solved to match the power lines to the solution of the disaggregating MILPs. While more detailed steps are given in the next subsection, detailed mathematical formulations of the first few steps are given in Appendix 2 of the supplementary text.

## 5.2 | Details

In this subsection, more details are given on the implementation of the algorithm.

---

### Algorithm 1: Aggregate Disaggregate Algorithm

---

```

1 Function Aggregate_Disaggregate Algorithm:
2   Step 1: Cluster the locations based on coordinates
3   Step 2: Solve the aggregated problem (A.4) after relaxing the integral operational variables
4   Step 3:
5   foreach cluster do
6     Solve Problem (A.5) to disaggregate the cluster keeping all the other clusters and their investment
       decisions fixed to the optimal solution of the aggregated problem.
7   end
8   Step 4: Solve Problem (35) to match power lines between clusters obtained from the previous step.
9   Step 5: Fix all the investment decisions and solve for the lower-level decisions to obtain a feasible
       solution to the original full-scale MILP (34) based on all the original candidate locations.
10 return

```

---

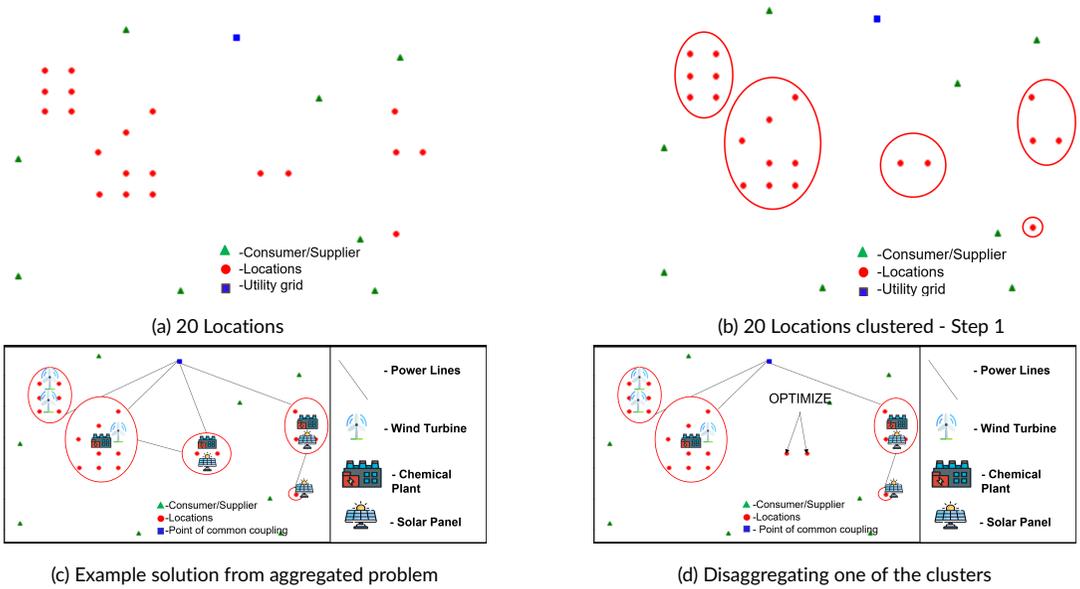
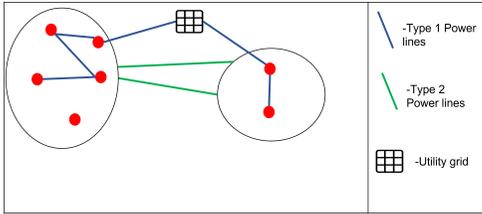


FIGURE 4 Steps of the Algorithm

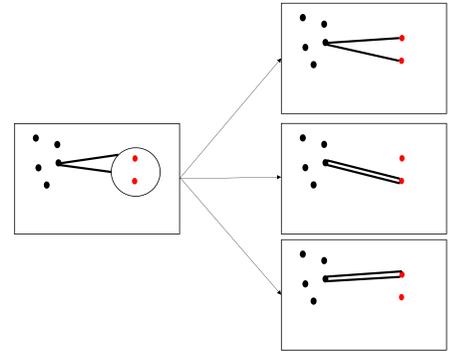
In Step 1 of the algorithm, the k-means clustering algorithm is run on the coordinates of the locations to cluster the locations. The number of clusters is chosen such that the aggregated MILP in Step 2 of the algorithm, can be solved in a reasonable amount of time. For example, for the locations shown in Figure 4a, we cluster the locations into 5 clusters and Figure 4b shows an image of the clusters derived from this step.

Using the centers of the clusters obtained from Step 1 of the algorithm as candidate locations, an aggregated problem is solved in Step 2. Here, the constraints on the investment decisions are aggregated based on the number of locations in each cluster. The operating variables are also relaxed.

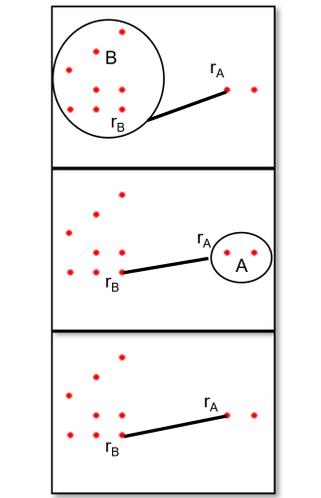
After solving the aggregated problem in Step 2 of the algorithm, the number of plants and power-generating units in each cluster, as well as a rough estimate of power lines between locations in different clusters, are obtained. Each cluster is then disaggregated in Step 3 of the algorithm, keeping the other clusters' investment decisions fixed. For example, Figure 4c represents the investment decisions obtained from the aggregated problem, Figure 4d represents the process of disaggregating one of the clusters, where the investment decisions concerning the two locations marked are solved, keeping the investment decisions of all the other clusters fixed.



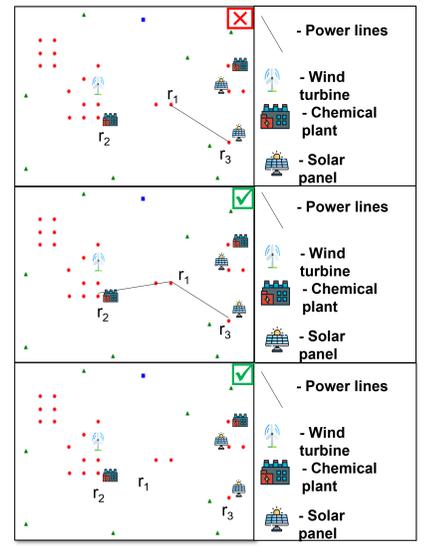
(a) Two types of power lines



(b) Constraining the sum of power lines in Step 4 (35b)



(c) Symmetry heuristic in Step 4 (35c)



(d) Configurations for power lines with no plant or generator in Step 4 (35f)-(35g)

FIGURE 5 Step 4 of Algorithm

From Step 3, the number of plants and power-generating units in each location are obtained as well as the power lines. As shown in Figure 5a, there are two types of power lines:

- Power lines between 2 candidate locations, which happens between 2 candidate locations in the same cluster on which we solve the disaggregation MILP, or a candidate location and the point of common coupling (there is a candidate power line from each location in the cluster to the point of common coupling when we solve the disaggregation MILP) shown by the blue lines in Figure 5a.

- Power lines between a candidate location and a cluster center, which happens as there are candidate lines from the locations of a cluster to the other cluster centers when we solve the disaggregation MILP shown by the green lines in Figure 5a.

Note that cluster centers are not real candidate locations and the investment decisions must involve lines between two real candidate locations. In order to convert the second type of power lines into the form of the first type, we solve the integer programming problem (35).

With a slight abuse of notation, we use the notations in Table 1 in this section.

TABLE 1 Notations in Section 5

Notation	Definition
$R_c$	Set of locations in the cluster with center $c$
$R^{empty}$	Set of locations with no plant or power-generating unit i.e., $r \in R^{empty}$ if $r \in R, \sum_{i \in \{I_{plant} \cup I_{power}\}} x_{i,r} = 0$ .
$nt_{r,c}^{disagg}$	Number of power lines which are of type two between $r$ and $c$ obtained from the disaggregation step
$\mathcal{L}^{td}$	Set of potential candidates for all power lines of type two i.e., $\forall nt_{r,c}^{disagg} \geq 1, (r, r') \in \mathcal{L}^{td}, r' \in R_c$
$c$	Cluster center
$r$	Candidate location
$nt_{r,r'}$	Integer variable denoting the number of power lines between locations $r$ and $r'$
$p_{r,r'}$	Binary variable indicating if there is at least 1 power line connecting locations $r$ and $r'$
$q_r$	Binary variable indicating if there are more than 2 different sets of power lines connected to $r$

Problem (35) aims to minimize the sum of the product of the number of power lines between 2 locations and the distance between the 2 corresponding locations, over all the potential power lines (defined in  $\mathcal{L}^{td}$ ), as the distance between 2 locations reflects the price of a power line between the 2 locations. Note that when we define the set  $\mathcal{L}^{td}$ , we only define one set of power lines and not the lines going both ways, i.e if we have both (p,v) and (v,p) as potential power line, we include only one of them in set  $\mathcal{L}^{td}$ .

The sum of the power lines connected to locations in a cluster is constrained based on the number of power lines connected to these locations which is obtained from the previous step. This is shown in Figure 5b and is described by equation (35b). To make the problem more accurate, the power lines of type 1 which are not connected to the point of common coupling can be included as well in the set  $\mathcal{L}^{td}$ , and the clusters can be defined as individual locations.

The symmetry of power lines is used as part of the heuristic, which is depicted in Figure 5c. If there is at least one power line between one location  $r_A$  in cluster A and the center of cluster B, and at least one line between a location  $r_B$  in cluster B and the center of cluster A obtained from the solution of the disaggregation step, then at least one power line can be constructed between the locations  $r_A$  and  $r_B$ . This heuristic rule is described by equation (35c).

If there is a location with no plant or power-generating unit, there cannot be exactly 1 set of power lines connected to it, as shown in Figure 5d. For example, if there is only one power line connected to location  $r_1$  (from  $r_3$ ) in Figure 5d, power transmitted from  $r_3$  is neither consumed nor transmitted from  $r_1$ . To ensure the power balance, there would be no power flowing through it, making it redundant. Therefore, there should be either no power line or at least two distinct sets of power lines connected to a location with no plant or power-generating unit (where power can flow in and out of the node). To ensure this, constraints (35f)-(35g) are added.

$$\min \sum_{r,r' \in \mathcal{L}^{td}} d_{r,r'} nt_{r,r'}, \quad (35a)$$

$$\text{s.t.} \quad \sum_{r' \in \mathcal{R}_{clus,c}} nt_{r,r'} \geq nt_{r,c}^{disagg}, \quad \forall nt_{r,c}^{disagg} \geq 1, \quad (35b)$$

$$nt_{r,r'} \geq 1, \quad \forall nt_{r,c'}^{disagg} \geq 1 \text{ and } nt_{c,r'}^{disagg} \geq 1, r \in \mathcal{R}_c, r' \in \mathcal{R}_{c'}, \quad (35c)$$

$$nt_{r,r'} \leq Mp_{r,r'}, \quad \forall (r, r') \in \mathcal{L}^{td}, \quad (35d)$$

$$nt_{r,r'} \geq p_{r,r'}, \quad \forall (r, r') \in \mathcal{L}^{td}, \quad (35e)$$

$$\sum_{(r,r') \in \mathcal{L}^{td}} p_{r,r'} + \sum_{(r',r) \in \mathcal{L}^{td}} p_{r',r} \leq Mq_r, \quad \forall r \in \mathcal{R}^{empty}, \quad (35f)$$

$$\sum_{(r,r') \in \mathcal{L}^{td}} p_{r,r'} + \sum_{(r',r) \in \mathcal{L}^{td}} p_{r',r} \geq 2q_r, \quad \forall r_1 \in \mathcal{R}^{empty}. \quad (35g)$$

After Step 4 of the algorithm, we have all the investment decisions for the model. In Step 5, we fix the investment decisions in the original model to the values obtained in the previous steps and solve the restricted problem (34) to find feasible operating decisions.

### 5.3 | Strategies to further speed up the algorithm

There is no control over the number of locations in each cluster, meaning that even the disaggregation MILP can be intractable for clusters containing a large number of locations. Therefore, the same aggregation-disaggregation strategy can be applied recursively within a cluster. A cluster can be directly disaggregated only if it has few locations. If the number of locations in the cluster exceeds a threshold, a  $k$ -means clustering can be done in that cluster to get new subclusters and the aggregation and disaggregation step can be done on the new subclusters. This process can be done till there are only a few locations in a subcluster for which the required MILP can be solved. If there are no plants or power-generating units in a particular cluster, it is not necessary to disaggregate it.

While disaggregating the clusters, the mode decisions for the other clusters can be fixed to further speed up the implementation of the algorithm. Further, the power lines can be bundled in such a way as to keep at least one variable to signify a power line connecting two real locations to reduce the number of binary variables in the model. The number of bundles can be chosen approximately based on the approximate number of plants (determined by the demand data), the number of solar panels, and the number of wind turbines as well as an average of the maximum possible power outputs through all the power lines.

## 6 | THE VALUE OF THE MULTI-SCALE MODEL (VMM)

Although the value of solving multi-scale models to consider lower-level operational decisions is well-recognized by the PSE community, there has not been a metric to quantify the additional value a multi-scale model can create compared with a single time-scale model. In this section, we introduce such a metric that we call *the Value of the Multi-scale Model* (VMM). The metric is inspired by the Value of the Stochastic Solution (VSS) introduced by the stochastic programming community to quantify the additional value stochastic solutions can bring compared with the optimal solutions to the deterministic model [53].

To define VMM, we introduce several new notations. Note that these new notations are only valid for this section and should be distinguished from the notations already introduced in previous sections. Here, we use  $x$  to represent

all the upper-level decisions, e.g., the investment decisions and the monthly decisions in our problem. Lower-level decisions are represented as  $\mathbf{y}_s$  defined for each subperiod  $s$ . In our model,  $\mathbf{y}_s$  corresponds to the hourly operating decisions on each representative day where  $s$  is the index for each day. The multi-scale model (MM) can be defined as,

$$\text{MM} = \min_{\mathbf{x}, \mathbf{y}_s} f(\mathbf{x}) + \sum_{s \in \mathcal{S}} w_s q(\mathbf{x}, \mathbf{y}_s, \boldsymbol{\theta}_s), \quad \text{s.t.} \quad \mathbf{g}(\mathbf{x}) \leq 0, \quad h(\mathbf{x}, \mathbf{y}_s, \boldsymbol{\theta}_s) \leq 0, \quad \forall s \in \mathcal{S}. \quad (36)$$

where  $\boldsymbol{\theta}_s$  denotes the parameters associated with subperiod  $s$ , e.g., the hourly capacity factors and electricity prices for a given representative day indexed by  $s$ . The objective is a weighted sum of the costs of the subperiods.  $h(\mathbf{x}, \mathbf{y}_s, \boldsymbol{\theta}_s)$  corresponds to the detailed operating constraints for the lower-level problem.

A single-scale model will ignore the detailed operating decisions during the subperiods. Instead, only the upper-level decisions  $\mathbf{x}$  are kept in the model. The nominal single-scale model (SM) is denoted as

$$\text{SM} = \min_{\mathbf{x}, \mathbf{z}} f(\mathbf{x}) + Q(\mathbf{x}, \mathbf{z}, \hat{\boldsymbol{\theta}}), \quad \text{s.t.} \quad \mathbf{g}(\mathbf{x}) \leq 0, \quad H(\mathbf{x}, \mathbf{z}, \hat{\boldsymbol{\theta}}) \leq 0. \quad (37)$$

where variables  $\mathbf{z}$  are introduced as a surrogate to the operating decisions.  $\hat{\boldsymbol{\theta}}$  is a nominal parameter that serves as a surrogate to the parameters  $\boldsymbol{\theta}_s$ . Typically, the dimensions of  $\mathbf{z}$  and  $\hat{\boldsymbol{\theta}}$  are lower than the dimensions of  $\mathbf{y}_s$  and  $\boldsymbol{\theta}_s$  because the detailed operations are ignored in the single-scale model. For example, in our model,  $\hat{\boldsymbol{\theta}}$  can be the average of the hourly electricity prices and capacity factors over the whole historical dataset. Similarly, the functions  $Q$ , and  $H$  are surrogates of the functions  $q$  and  $h$ . The reduction of the multi-scale model to the single-scale model is not a trivial task and usually depends on the engineering insight of the problem. By solving (37), we can obtain an upper-level solution denoted as  $\mathbf{x}^{SM}$ .

We further define a concept called *the Multi-scale Performance of the Single-scale Solution* (MPSS). The idea is to fix the upper-level decisions to  $\mathbf{x}^{SM}$  and solve the rest of the multi-scale problem to obtain the "true cost" of the solution  $\mathbf{x}^{SM}$ .

$$\text{MPSS} = \min_{\mathbf{y}_s} f(\mathbf{x}^{SM}) + \sum_{s \in \mathcal{S}} q(\mathbf{x}^{SM}, \mathbf{y}_s, \boldsymbol{\theta}_s), \quad \text{s.t.} \quad h(\mathbf{x}^{SM}, \mathbf{y}_s, \boldsymbol{\theta}_s) \leq 0 \quad \forall s \in \mathcal{S}. \quad (38)$$

It should be noted that  $\mathbf{x}^{SM}$  can be suboptimal or even infeasible when multi-scale decisions are considered. The difference between MPSS and MS indicates the additional economic savings we can get by solving the multi-scale model, which is defined as *the Value of the Multi-scale Model* (VMM).

$$\text{VMM} := \text{MPSS} - \text{MM}. \quad (39)$$

We will show in our case study the VMM of our multi-scale model compared with a *0 representative day model* that ignores the hourly operating decisions.

**Difference with VSS** In computing VSS, an *Expected Value* problem is solved for one nominal scenario corresponding to the expected value of  $\boldsymbol{\theta}_s$ , i.e.,  $\bar{\boldsymbol{\theta}} = \sum_{s \in \mathcal{S}} w_s \boldsymbol{\theta}_s$

$$\text{EV} = \min_{\mathbf{x}, \bar{\mathbf{y}}} f(\mathbf{x}) + q(\mathbf{x}, \bar{\mathbf{y}}, \bar{\boldsymbol{\theta}}), \quad \text{s.t.} \quad \mathbf{g}(\mathbf{x}) \leq 0, \quad h(\mathbf{x}, \bar{\mathbf{y}}, \bar{\boldsymbol{\theta}}) \leq 0 \quad (40)$$

The main difference between EV and the SM defined in (37) is that EV is still a multi-scale model but just one nominal

scenario is kept. In comparison, SM is a single-scale model that has a much smaller number of variables and constraints. For example, in our problem, EV is a *1 representative day model* while SM is a *0 representative day model*.

Let the solution to the EV be  $x^{EV}$ . We can evaluate the performance of  $x^{EV}$  over all the scenarios by computing the so-called *expected result of using the EV solution* (EEV) [53].

$$EEV = \min_{y_s} f(x^{EV}) + \sum_{s \in S} q(x^{EV}, y_s, \theta_s), \quad \text{s.t.} \quad h(x^{EV}, y_s, \theta_s) \leq 0 \quad \forall s \in S. \quad (41)$$

EEV and MPSS only differ in the way the upper-level decisions are obtained. VSS is defined by

$$VSS := EEV - MM. \quad (42)$$

Empirically, VMM is usually larger than VSS because the model complexity of the SM is lower than EV.

## 7 | CASE STUDY

The MILP model and the proposed algorithm are implemented in a case study in Western Texas with 20 candidate locations whose coordinates are known. We are given 15 1500 kW solar panels, 150 100kW wind turbines, modular chlorine plants operating in three modes, and 12 kV power lines. Additionally, the location of the point of common coupling and the location of four consumers and four suppliers are given. The demand for each consumer is generated randomly from a uniform distribution.

For the chemical plant, we consider a chloralkali process, an industrial process to produce chlorine and sodium hydroxide by the electrolysis of sodium chloride solutions. In this work, we use a chlorine plant that operates in 3 modes - standard cathodes (STCs/ST), oxygen-depolarized cathodes (ODCs/OD), and Shut down [26]. The reaction in ST mode producing  $H_2$  as by-product is:  $2NaCl + 2H_2O \rightleftharpoons 2NaOH + Cl_2 + H_2$ . The net reaction of the chloralkali electrolysis in the OD mode is  $2NaCl + H_2O + 0.5O_2 \rightleftharpoons 2NaOH + Cl_2$ . The energy requirement and the capacity of the chlorine plant are obtained from [26]. We get the capital costs of the plant and the capital costs and energy requirement of compressors and turbine from Aspen Plus and add a cost of \$160,000 for making the electrolyzer into bi-mode. A storage cost that is 5% of the total capital cost for each plant is further added to the capital cost. We assume that the plant remains in one mode for at least two hours after switching to it each day and has a negligible cleaning time while switching modes.

The power output for the wind turbines and solar panels at the candidate locations for three historical years (2011-2013), as well as the capital costs of these power-generating units, are obtained from the software SAM (System Advisor Model) [54] after obtaining the weather-related data from NREL (National Renewable Energy Laboratory) website (National Solar Radiation Database<sup>1</sup> and Wind Integration National Dataset Toolkit<sup>2</sup>). We assume that there are incentives for wind turbines. The electricity price for West Texas for three historical years (2011-2013) is obtained from the website Energy Online [55] and adjusted for inflation. The parameters of the power lines such as resistance and inductance, are obtained from DERCAM (Distributed Energy Resource Customer Adoption Model) [56]. Five representative days per month are chosen based on the data using k-means clustering. The power lines are used in bundles of size 20 while implementing the algorithm.

<sup>1</sup><https://nsrdb.nrel.gov/>

<sup>2</sup><https://www.nrel.gov/grid/wind-toolkit.html>

## 7.1 | Results

The proposed algorithm and the corresponding models are implemented in Julia/JuMP. All the MILPs are solved using Gurobi version 10.0 on a Linux cluster with 48 AMD EPYC 7643 2.3GHz CPUs and 1 TB RAM. The total annual profit obtained at the end of the algorithm is \$18.94 M and the total time taken for all the steps of the algorithm is 7.5 hours. An upper bound for the profit of the problem is obtained by solving the LP relaxation of the full-scale model using the barrier method without crossover and it is \$19.50 M. The original full space model has 10,111,090 constraints, 2,094,336 continuous variables, and 346,080 integer variables.

**The value of the multi-scale models:** In order to understand the value of the multi-scale model, it is compared with a model that does not consider temporal variations. To this end, a model ignoring hourly variations is formulated where the hourly operating variables are aggregated by month. The new model only has investment and monthly variables and thus has no representative day in each month. Therefore, it is called a *0 representative day model* and outputs the investment decisions. This investment decision can then be fixed in the multi-scale model to obtain the hourly operating variables and overall profit. Thus through this, the planning and scheduling are separated, and the planning does not take into account the hourly temporal variations. This profit is compared to the profit obtained from the algorithm. It has been found that the profit obtained using the *0 representative day model* is lesser than the profit obtained from the model and that the *0 representative day model* does not install any solar panels, as without the hourly variations taken into account, the solar panels are not profitable. The VMM is \$0.42 M<sup>3</sup>. Figures 6 and 7 show the optimal investment decisions obtained from the multi-scale model and the *0 representative day model*, respectively. The dots represent the candidate locations, the blue square represents the location of the point of common coupling, and the triangles represent consumers/suppliers.

The algorithm provides a configuration that favors distributed manufacturing. To further demonstrate the value of distributed manufacturing, it is compared with centralized manufacturing, where we calculated the profit by installing all the plants, solar panels, and wind turbines at each one of the 20 locations. The installed facilities are connected to the point of common coupling through sufficient number of power lines. The maximum profit obtained from centralized manufacturing across all 20 locations is around \$18.90 M, around \$0.04 M less than the profit obtained from the algorithm. However, land-based constraints may limit the number of plants or renewable resources that can be installed in a particular location. We have implemented the algorithm under such conditions, with further details and the associated results provided in Supplementary C. The profit obtained from the algorithm with these constraints, is \$18.76 M is lesser than the profit obtained without constraints.

Table 3 shows a summary of the algorithm's computing time per step and compares the profit obtained with the profit obtained from the model with 0 representative days and the average and maximum profit obtained from centralized production.

The breakdown of the profit is shown in Table 2.

An important observation to note from the breakdown of the profit is that a major part of the revenue comes from the sales of chlorine and other chemicals, and the major cost is from the capital cost. The profit from the electricity makes a considerable contribution to the profit and thus the location of the plant and the power-generating units could influence the profit.

Figure 6 is a representation of the optimal configuration obtained from solving the MILP model using the algorithm proposed.

---

<sup>3</sup>Note that the mathuristic did not solve the model to global optimality, this calculated VMM is an approximation.

TABLE 2 Breakdown of profit

Breakdown of annual profit	Cost/Revenue in \$M
Net revenue from materials	28.22
Net transportation cost	0.23
Net profit from electricity	3.16
Total fixed operating cost	5.98
Total amortized capital cost	6.23

TABLE 3 Results Summary

Statistics	Value
Upper bound from LP relaxation (\$)	19.50 M
Total profit from proposed algorithm (\$)	18.94 M
Time for aggregation (h)	7.12
Time for disaggregation (h)	0.35
Time for solving operation decisions (h)	0.03
Profit obtained using the model with 0 representative days (\$)	18.52 M
VMM (\$)	0.42 M
Average profit from centralized production (\$)	18.72 M
Maximum profit from centralized production (\$)	18.90 M
Profit from proposed algorithm with constraints on investment decisions (described in Supplementary C) (\$)	18.76 M

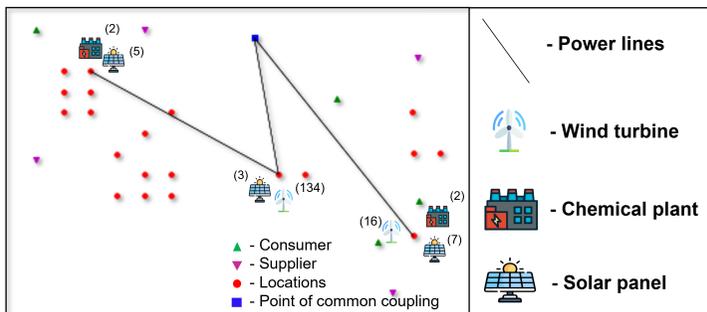
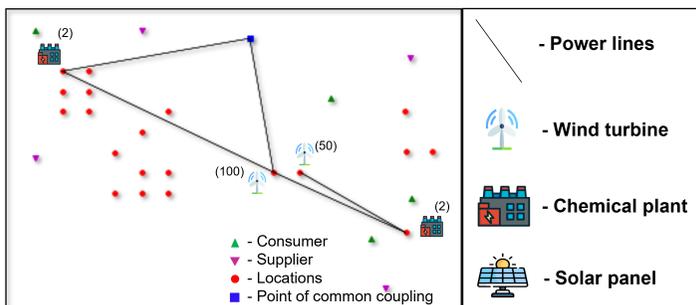
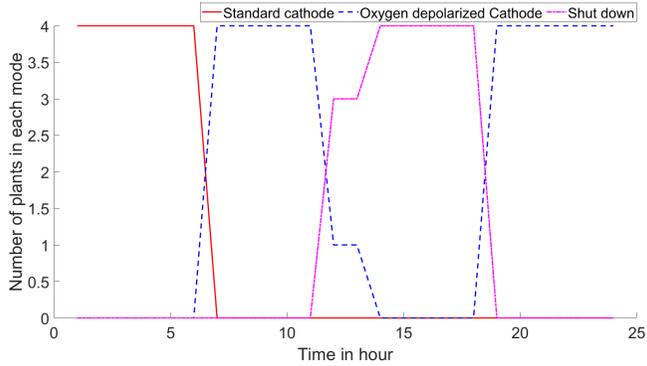


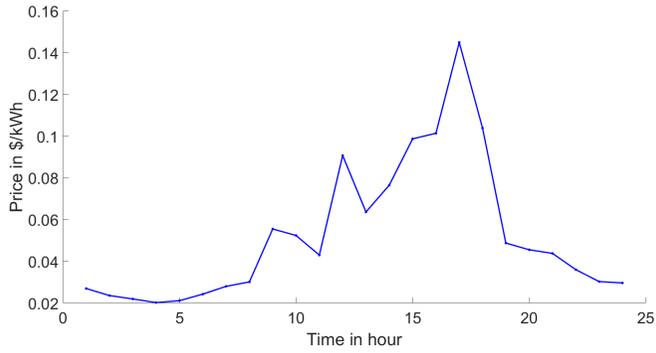
FIGURE 6 Solution Representation from the algorithm for the investment decisions



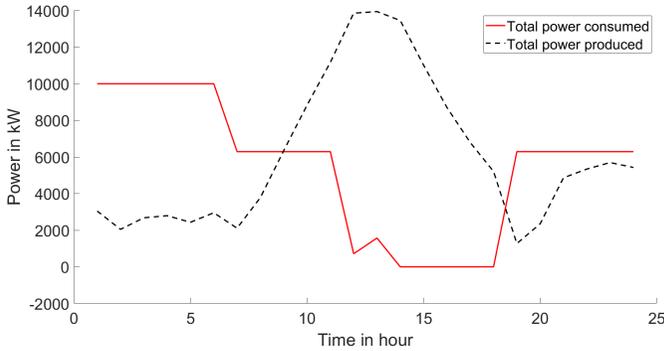
### 7.1.1 | Hourly Operating Decisions



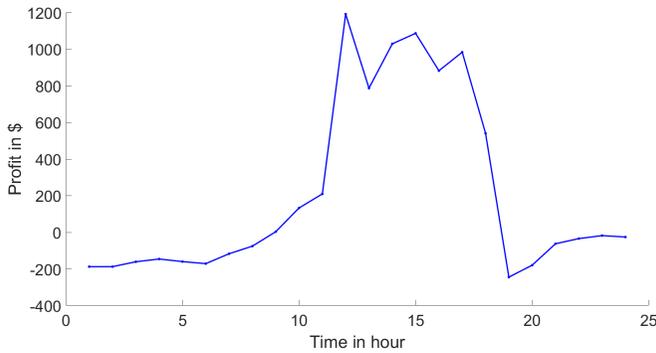
(a) Hourly mode decisions on representative day 1 of month 9



(b) Hourly electricity prices on representative day 1 of month 9



(c) Hourly power produced and consumed by the microgrid on representative day 1 of month 9



(d) Hourly profit from electricity on representative day 1 of month 9

FIGURE 8 Operating decisions on representative day 1 of month 9

In order to understand the trends on an hourly basis and the effect of the power produced and electricity price on hourly decisions, several hourly decisions and parameters of representative day 1 in month 9 are plotted and shown in Figures 8a-8d.

Figure 8a, shows the dynamics of the mode switching. When the electricity price is low, the STC mode is mostly preferred, when the price is moderately high the ODC mode is mostly preferred and when the price is extremely high the shutdown mode is mostly preferred. This is because the STC mode consumes high power but does not consume any oxygen and produces hydrogen, while the ODC mode consumes relatively lower power and consumes oxygen with no hydrogen produced. Thus, these two competing effects (consumption/production of oxygen/hydrogen and power required) help choose the mode for each hour.

Power is sold or bought from the electric utility depending on the net effect of the power consumed by the plant, the power generated by the solar panels and wind turbines, and the power loss through the power lines. It can be seen from Figure 8b that there is a peak in the electricity price in hour 17. The difference in power produced and power consumed in hour 17 is moderate, as seen in Figure 8c. Thus, all the plants run in the Shut down mode in hour 17, as seen in Figure 8a, and the profit from electricity is high, as seen in Figure 8d.

Another interesting observation is in hour 12 when the profit from electricity is at its peak for the day. The

electricity price is moderately high in hour 12, and hence the Shut down mode is chosen for 3 of the plants and OD mode for the other plant as seen in Figure 8a. In the first few hours, the power produced is negligible. However, the plants operate in the SC mode during the first few hours of the day. This is because the electricity price is very low.

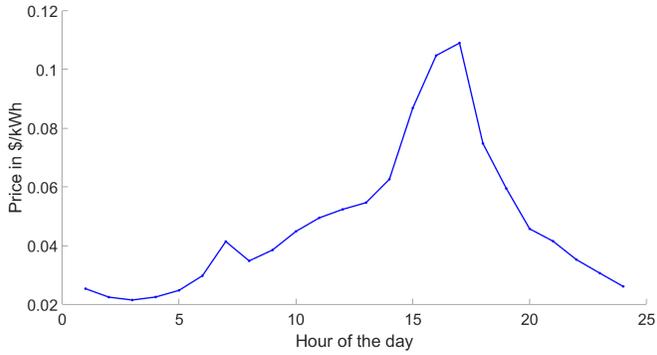


FIGURE 9 Average electricity price in each hour

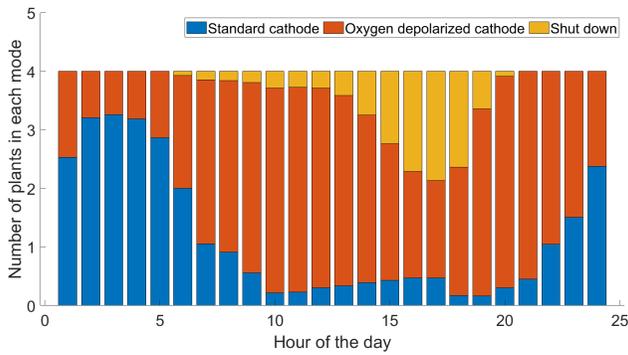


FIGURE 10 Average number of plants in each mode per hour

The weighted average of electricity price over all the representative days for each hour and the average number of plants in each mode for each hour is further plotted in Figures 9-10 and analyzed. An observation that can be made is that when the electricity price increases, the number of plants in the shutdown mode majorly increases. This is consistent with the trends.

### 7.1.2 | Monthly Trends

From the results obtained, an observation that can be made is that most of the demand is met by the plants and that the slack variable for each month is very small. Figure 11 shows the average electricity price for each month whereas Figure 12 depicts the average percentage of plants on each mode for each month. While the OD mode is predominantly chosen for plants on average, as seen in Figure 12, the percentage of shutdown mode mostly increases

as the price of electricity increases, which is consistent with the hourly trends as well.

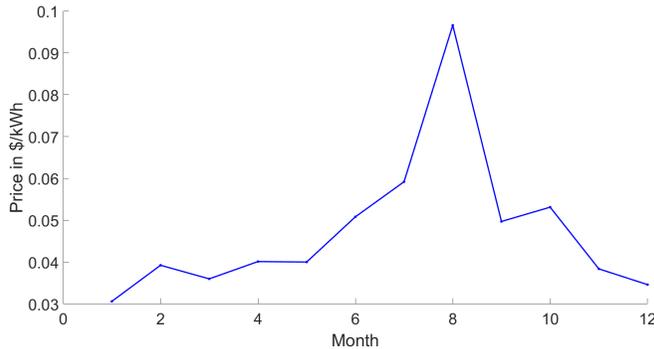


FIGURE 11 Average electricity price in each month

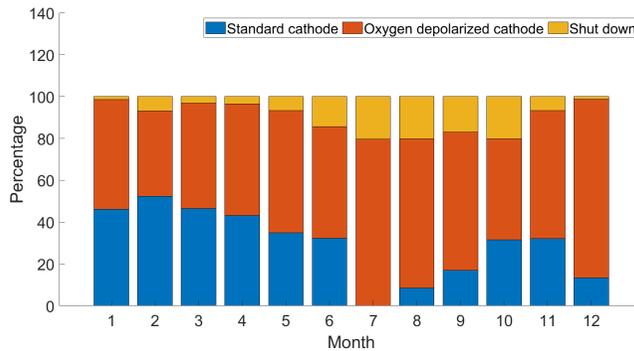


FIGURE 12 Mode decisions in a month

### 7.1.3 | Effect of more Renewable Electricity Generation

In the future, when the renewable electricity generation capacity becomes more abundant, electricity prices are expected to be depressed during hours of peak PV output and increase sharply at sundown [57]. Let us assume that the sundown time is from 6-7 pm. We reduce the price by 10% for all hours between 7 am to 5 pm and assign 10% more of the current highest price to the electricity price between hours 6 pm to 7 pm. For other hours, we keep the prices the same. This is done for each representative day of each month. Additionally, we expect that the costs of installing solar panels and wind turbines are going to reduce [58, 59]. We reduce the capital and operating costs of solar panels by 50% and wind turbines by 40% .

For this scenario, the MILP model and the proposed algorithm are implemented. The total profit obtained at the end of the algorithm is \$20.91 M. Figure 13 depicts the optimal configuration obtained from the MILP model for this scenario. The optimal configuration is different from the base case, with the plants, wind turbines, and solar panels relocated. The dots represent the candidate locations, the blue square represents the location of the point of common

coupling and the triangles represent consumers and suppliers.

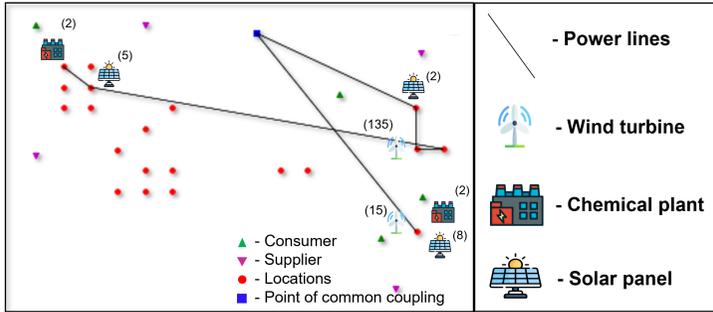


FIGURE 13 Solution Representation from the algorithm for investment decisions in Scenario 2

The split of the profit is shown in Table 4.

TABLE 4 Split up of profit for Scenario 2

Split up of annual profit	Cost/Revenue in \$M in Scenario 2	Cost/Revenue in \$M in Base Case
Net revenue from materials	28.13	28.22
Net transportation cost	0.23	0.23
Net profit from electricity	3.25	3.16
Total fixed operating cost	5.71	5.98
Total amortized capital cost	4.53	6.23

While the profit from electricity has increased in scenario 2 as compared to the base case, the capital cost and fixed operating cost in scenario 2 are less than in the base case. The net effect is that the profit in scenario 2 is more than the profit obtained in the base case.

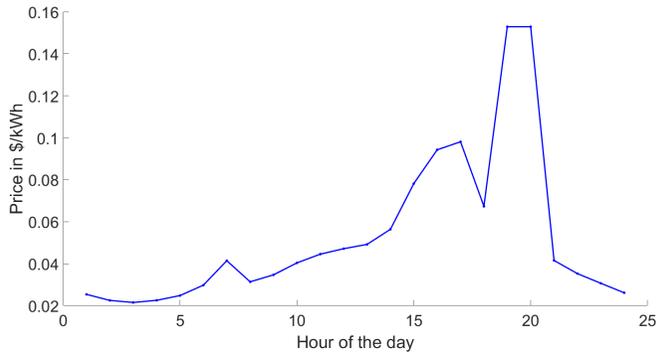


FIGURE 14 Average electricity price in each hour in scenario 2

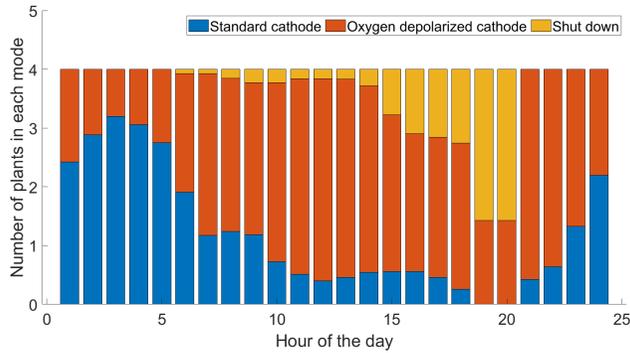


FIGURE 15 Average number of plants in each mode per hour in scenario 2

The plots of the average electricity price for each hour and the average number of plants in each mode for each hour are further analyzed in Figures 14-15. It is found that in the daytime, the average number of plants in the shutdown mode is lesser than that in the base case, while in the hours 19 and 20, which represent 6 pm and 7 pm, the average number of plants in the shutdown mode is more than that in the base case. This is because, in the daytime, the electricity price is lesser in this scenario, while at 6 pm and 7 pm, the electricity price is at its peak and is more than those in the base case.

## 8 | CONCLUSIONS

Electrification is a solution being explored to tackle the problem of greenhouse gas emissions by the chemical industry. Electrification using renewable sources of energy includes spatial and temporal variations, which can affect plant production. An MILP model was proposed to determine a feasible near-optimal configuration of a network of modular plants, power-generating units, and power lines connected by a microgrid considering spatial and temporal variations in electricity price and weather conditions. The model takes decisions in three-time scales: one-time investment decisions, monthly decisions like transportation as well as hourly operating decisions. A spatial aggregation and disaggregation algorithm based on k-means clustering was developed to solve the model efficiently. We propose a new metric the value of the multi-scale model (VMM) to quantify the economic savings of the multi-scale model.

The model was tested on a case study with 20 candidate locations in Western Texas with data obtained from various sources. The mode of the plant was chosen depending on the power produced and the electricity price. The major contributor to the profit was found to be the materials (products and raw materials). The profit obtained from electricity was found to have a significant contribution to the total profit and thus, the spatial location can affect the profit. The mode switching and the profit from electricity were found to be consistent with the theoretical trends. The value of the multi-scale model was inferred by comparing it with the net annual profit obtained from a model with 0 representative days, and we found that the profit from the 0 representative day model was \$0.42 M lesser than the profit obtained from our model. The value of distributed manufacturing was inferred by comparing it with centralized manufacturing and it was found that the configuration with distributed manufacturing gave more profit. Further, we investigated the case study in the context of a scenario in the future, where there will be more renewable electricity generation, and found that the profit increases in this scenario.

Although the case study is on the chloralkali process, we expect that as technologies for green hydrogen, green

ammonia, and other chemicals which can be produced electrochemically, get more mature, the model can be used for multi-scale planning of these new technologies. We also expect the proposed approach can have a broader impact as the cost of renewable power-generating resources reduces, and microgrids become more widely adopted. In addition, the proposed heuristic algorithm is general and can be adapted to solve any multi-scale facility location problem, which has been understudied by the Operations Research community.

While bundling of power lines was used to speed up the algorithm, it adds inherent constraints to the number of power lines placed between 2 locations. Additionally, the number of lines in the bundle is heuristically determined based on data. In the future, we will develop novel methods that do not require bundling the power lines to speed up the algorithm.

## References

- [1] Tickner Joel, Geiser Ken, Baima Stephanie. Transitioning the Chemical Industry: The Case for Addressing the Climate, Toxics, and Plastics Crises *Environment: Science and Policy for Sustainable Development*. 2021;63:4-15.
- [2] Weeda Marcel, Delft Yvonne C, Kranenburg K J, Schols E, Gelevert Hugo, Kler Rajnish. Whitepaper Empowering the Chemical Industry: Opportunities for Electrification 2016.
- [3] Mallapragada Dharik S, Dvorkin Yury, Modestino Miguel A, et al. Decarbonization of the chemical industry through electrification: Barriers and opportunities *Joule*. 2023;7:23-41.
- [4] Orella Michael J, Román-Leshkov Yuriy, Brushett Fikile R. Emerging opportunities for electrochemical processing to enable sustainable chemical manufacturing *Current Opinion in Chemical Engineering*. 2018;20:159-167.
- [5] William J. Grieco . Advanced Manufacturing Progress: Distributed Manufacturing for the Process Industries 2019.
- [6] Sampat Apoorva M, Kumar Ranjeet, Pushpangatha Kurup Remya, Chiu Kawa, Saucedo Victor M, Zavala Victor M. Multisite supply planning for drug products under uncertainty *AIChE Journal*. 2021;67:e17069.
- [7] Lara Cristiana L., Grossmann Ignacio E.. Global Optimization for a Continuous Location-Allocation Model for Centralized and Distributed Manufacturing *Computer Aided Chemical Engineering*. 2016;38:1009-1014.
- [8] Palys Matthew J., Allman Andrew, Daoutidis Prodromos. Exploring the Benefits of Modular Renewable-Powered Ammonia Production: A Supply Chain Optimization Study *Industrial & Engineering Chemistry Research*. 2019;58:5898-5908.
- [9] Mahmoud M. S., Azher Hussain S., Abido M. A.. Modeling and control of microgrid: An overview *Journal of the Franklin Institute*. 2014;351:2822-2859.
- [10] The Role of Microgrids in Helping to Advance the Nation's Energy System
- [11] Ball Michael O. Heuristics based on mathematical programming *Surveys in Operations Research and Management Science*. 2011;16:21-38.
- [12] Fischetti Matteo, Lodi Andrea. Local branching *Mathematical Programming*. 2003;98:23-47.
- [13] Fischetti Matteo, Glover Fred, Lodi Andrea. The feasibility pump *Mathematical Programming*. 2005;104:91-104.

- [14] Cplex I B M ILOG. V12. 1: User's Manual for CPLEX *International Business Machines Corporation*. 2009;46:157.
- [15] Bernal David E., Vigerske Stefan, Trespalacios Francisco, Grossmann Ignacio E.. Improving the performance of DICOPT in convex MINLP problems using a feasibility pump *Optimization Methods and Software*. 2020;35:171–190.
- [16] Lazouski Nikifar, Limaye Aditya, Bose Abhishek, Gala Michal L, Manthiram Karthish, Mallapragada Dharik S. Cost and Performance Targets for Fully Electrochemical Ammonia Production under Flexible Operation *ACS Energy Letters*. 2022;7:2627–2633.
- [17] Sánchez Antonio, Martín Mariano. Scale up and scale down issues of renewable ammonia plants: Towards modular design *Sustainable Production and Consumption*. 2018;16:176–192.
- [18] Cooper Nathaniel, Horend Christian, Röben Fritz, Bardow Andre, Shah Nilay. A framework for the design & operation of a large-scale wind-powered hydrogen electrolyzer hub *International Journal of Hydrogen Energy*. 2022;47:8671–8686.
- [19] Demirhan C Doga, Tso William W, Powell Joseph B, Pistikopoulos Efstratios N. Sustainable ammonia production through process synthesis and global optimization *AIChE Journal*. 2019;65:e16498.
- [20] Corengia Mariana, Torres Ana I.. Coupling time varying power sources to production of green-hydrogen: A superstructure based approach for technology selection and optimal design *Chemical Engineering Research and Design*. 2022;183:235–249.
- [21] Wang Hanchu, Daoutidis Prodromos, Zhang Qi. Harnessing the Wind Power of the Ocean with Green Offshore Ammonia *ACS Sustainable Chemistry & Engineering*. 2021;9:14605–14617.
- [22] Gençer Emre, Agrawal Rakesh. Strategy to synthesize integrated solar energy coproduction processes with optimal process intensification. Case study: Efficient solar thermal hydrogen production *Computers & Chemical Engineering*. 2017;105:328–347.
- [23] Bose Abhishek, Lazouski Nikifar, Gala Michal L, Manthiram Karthish, Mallapragada Dharik S. Spatial Variation in Cost of Electricity-Driven Continuous Ammonia Production in the United States *ACS Sustainable Chemistry & Engineering*. 2022;10:7862–7872.
- [24] Sánchez Antonio, Martín Mariano. Optimal renewable production of ammonia from water and air *Journal of Cleaner Production*. 2018;178:325–342.
- [25] Martín Mariano. Methodology for solar and wind energy chemical storage facilities design under uncertainty: Methanol production from CO<sub>2</sub> and hydrogen *Computers & Chemical Engineering*. 2016;92:43–54.
- [26] Brée Luisa C, Perrey Karen, Bulan Andreas, Mitsos Alexander. Demand side management and operational mode switching in chlorine production *AIChE Journal*. 2019;65:e16352.
- [27] Allman Andrew, Daoutidis Prodromos. Optimal scheduling for wind-powered ammonia generation: Effects of key design parameters *Chemical Engineering Research and Design*. 2018;131:5–15.
- [28] Bødal Espen Flo, Korpås Magnus. Value of hydro power flexibility for hydrogen production in constrained transmission grids *International Journal of Hydrogen Energy*. 2020;45:1255–1266.

- [29] Zheng Yi, You Shi, Li Ximei, Bindner Henrik W., Münster Marie. Data-driven robust optimization for optimal scheduling of power to methanol *Energy Conversion and Management*. 2022;256:115338.
- [30] Castro Pedro M, Dalle Ave Giancarlo, Engell Sebastian, Grossmann Ignacio E, Harjunkoski Iiro. Industrial Demand Side Management of a Steel Plant Considering Alternative Power Modes and Electrode Replacement *Industrial & Engineering Chemistry Research*. 2020;59:13642–13656.
- [31] Lal Apoorv, You Fengqi. Comparative Life Cycle Analysis and Optimization of Operating Conditions of Hydrogen Production Methods *Chemical Engineering Transactions*. 2022;94:517–522.
- [32] Kelley Morgan T, Do Tam T, Baldea Michael. Evaluating the demand response potential of ammonia plants *AIChE Journal*. 2022;68:e17552.
- [33] Roh K, Brée L C, Schäfer P, Strohmeier D, Mitsos A. Flexible operation of modular electrochemical CO<sub>2</sub> reduction processes *IFAC-PapersOnLine*. 2022;55:298–303.
- [34] Guerra Omar J., Eichman Joshua, Kurtz Jennifer, Hodge Bri Mathias. Cost Competitiveness of Electrolytic Hydrogen *Joule*. 2019;3:2425–2443.
- [35] He G, Mallapragada D S, Bose A, Heuberger C F, Gençer E. Hydrogen Supply Chain Planning With Flexible Transmission and Storage Scheduling *IEEE Transactions on Sustainable Energy*. 2021;12:1730–1740.
- [36] Welder Lara, Ryberg D, Severin, Kotzur Leander, Grube Thomas, Robinius Martin, Stolten Detlef. Spatio-temporal optimization of a future energy system for power-to-hydrogen applications in Germany *Energy*. 2018;158:1130–1149.
- [37] Li Jiarong, Lin Jin, Heuser Philipp-Matthias, et al. Co-Planning of Regional Wind Resources-based Ammonia Industry and the Electric Network: A Case Study of Inner Mongolia *IEEE Transactions on Power Systems*. 2022;37:65–80.
- [38] Allman Andrew, Daoutidis Prodomos, Tiffany Douglas, Kelley Stephen. A framework for ammonia supply chain optimization incorporating conventional and renewable generation *AIChE Journal*. 2017;63:4390–4402.
- [39] Bødal Espen Flo, Mallapragada Dharik, Botterud Audun, Korpås Magnus. Decarbonization synergies from joint planning of electricity and hydrogen production: A Texas case study *International Journal of Hydrogen Energy*. 2020;45:32899–32915.
- [40] Hong Xiaodong, Garud Sushant S., Thaore Vaishali B., et al. Hydrogen Economy Assessment & Resource Tool (HEART): A python-based tool for ASEAN H<sub>2</sub> roadmap study *International Journal of Hydrogen Energy*. 2022;47:21897–21907.
- [41] Wang Hanchu, Daoutidis Prodomos, Zhang Qi. Green Ammonia Supply Chain Design for Maritime Transportation *Computer Aided Chemical Engineering*. 2022;49:589–594.
- [42] Ogumerem Gerald S, Tso William W, Demirhan C Doga, Lee Seung Y, Song Han E, Pistikopoulos Efstratios N. Toward the Optimization of Hydrogen, Ammonia, and Methanol Supply Chains \*\*The authors declare no competing interests and acknowledge financial support from the Texas A&M Energy Institute and Shell Oil Company *IFAC-PapersOnLine*. 2019;52:844–849.

- [43] Potrč Sanja, Čuček Lidija, Martin Mariano, Kravanja Zdravko. Sustainable renewable energy supply networks optimization – The gradual transition to a renewable energy system within the European Union by 2050 *Renewable and Sustainable Energy Reviews*. 2021;146:111186.
- [44] Reinert Christiane, Schellhas Lars, Mannhardt Jacob, et al. SecMOD: An Open-Source Modular Framework Combining Multi-Sector System Optimization and Life-Cycle Assessment *Frontiers in Energy Research*. 2022;10.
- [45] Reinert Christiane, Nilges Benedikt, Baumgärtner Nils, Bardow André. This is SpArta: Rigorous Optimization of Regionally Resolved Energy Systems by Spatial Aggregation and Decomposition *arXiv e-prints*. 2023:arXiv:2302.05222.
- [46] Maravelias Christos T, Sung Charles. Integration of production planning and scheduling: Overview, challenges and opportunities *Computers & Chemical Engineering*. 2009;33:1919–1930.
- [47] Li Can, Conejo Antonio J., Liu Peng, Omell Benjamin P., Sirola John D., Grossmann Ignacio E.. Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems *European Journal of Operational Research*. 2022;297:1071–1082.
- [48] Allen R. Cory, Iseri Funda, Demirhan C. Doga, Pappas Iosif, Pistikopoulos Efstratios N.. Improvements for decomposition based methods utilized in the development of multi-scale energy systems *Computers & Chemical Engineering*. 2023;170:108135.
- [49] Li Can, Conejo Antonio J., Sirola John D., Grossmann Ignacio E.. On representative day selection for capacity expansion planning of power systems under extreme operating conditions *International Journal of Electrical Power & Energy Systems*. 2022;137:107697.
- [50] Jung Johannes, Postels Sarah, Bardow André. Cleaner chlorine production using oxygen depolarized cathodes? A life cycle assessment *Journal of Cleaner Production*. 2014;80:46–56.
- [51] Frank Stephen, Rebennack Steffen. An introduction to optimal power flow: Theory, formulation, and examples *IIE Transactions*. 2016;48:1172–1197.
- [52] McCormick Garth P. Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems *Mathematical Programming*. 1976;10:147–175.
- [53] Birge John R., Louveaux François. *Introduction to Stochastic Programming*. New York, NY: Springer New York 2011.
- [54] System Advisor Model 2021.12.2, (SAM 2021.12.2)
- [55] ERCOT: Real-time price - LCG consulting
- [56] Deforest Nicholas, Cardoso Goncalo, Brouhard Thomas, USDOE . Distributed Energy Resources Customer Adoption Model (DER-CAM) v5.9 2018.
- [57] Mallapragada Dharik Sanchan, Gençer Emre, Insinger Patrick, Keith David William, O’Sullivan Francis Martin. Can Industrial-Scale Solar Hydrogen Supplied from Commodity Technologies Be Cost Competitive by 2030? *Cell Reports Physical Science*. 2020;1:100174.

- 
- [58] Collins Leigh. Solar to be world's largest power source by 2050 as cost halves 2019.
- [59] Wiser Ryan, Rand Joseph, Seel Joachim, et al. Expert elicitation survey predicts 37% to 49% declines in wind energy costs by 2050 *Nature Energy*. 2021;6:555–565.