Combining Simulation with Machine Learning and Optimization to Assess Green Hydrogen Production via Offshore Wind in the Dutch North Sea

Justin Starreveld[1], Laurens Frowijn[2,*], Riccardo Travaglini[3], Renske van 't Veer[2], Alessandro Bianchini[3], Kenneth Bruininx[2], Dick den Hertog[1], and Zofia Lukszo[2]

[1] *Amsterdam Business School, University of Amsterdam, 1018 TV Amsterdam, Netherlands*
[2] *Faculty of Technology, Policy and Management, Delft University of Technology, 2600 GA Delft, Netherlands*
[3] *Dipartimento di Ingegneria Industriale, University of Florence, 50139 Firenze FI, Italy*
[*] *Corresponding author:* L.S.F.Frowijn@tudelft.nl.

August 25, 2025

Abstract

As countries seek to decarbonize their energy systems, green hydrogen has emerged as a promising energy carrier. This paper studies the production of green hydrogen from offshore wind in the Dutch North Sea, with particular emphasis on understanding how system design decisions and uncertain parameters affect key performance indicators. The analysis builds on a detailed techno-economic simulation modeling framework that is used to accurately simulate energy flows and estimate the levelized cost of hydrogen. To address the strong dependence of simulation outputs on input choices, and the computational infeasibility of evaluating the full input space, this study applies "optimization with constraint learning", where supervised machine learning models are trained using simulation data. This approach allows one to approximate the relevant input-output relationships and replace the computationally expensive simulation framework with surrogate machine learning models that are more tractable and can be embedded in mixed-integer optimization problem formulations. The methodology provides a fast and flexible decision-support tool, offering both descriptive insights and prescriptive guidance for decision-making. The results the effectiveness of such an approach for analyzing trade-offs and exploring stakeholder-defined objectives and constraints in a dynamic and computationally efficient manner.

Keywords: OR in energy; Hydrogen; Offshore wind; Optimization; Simulation.

1 Introduction

Hydrogen is expected to play a key role in the transition towards a sustainable energy system. Both the European Commission and the Dutch government have set ambitious targets for the deployment of green hydrogen [1, 2], which refers to hydrogen that is produced via water electrolysis powered by renewable electricity. The North Sea is one of the world's most active regions for offshore wind development and provides a favorable environment for the production of green hydrogen. While pilot projects are progressing [3, 4], the economic viability of large-scale offshore wind-to-hydrogen systems remains an open question.

To assess this viability of green hydrogen production, various contemporary studies utilize highly-detailed mathematical simulation models. Notable examples of such simulation-based studies are the following. Egeland-Eriksen and Sartori [5] investigate green hydrogen production via a 1.5 GW offshore wind farm close to the Norwegian island of Utsira. Hill et al. [6] assess the costs for offshore wind farms in the United Kingdom. Calado et al. [7] compare offshore and onshore system configurations, with a focus on assessing the influence of electricity market regulation, for the Iberian Peninsula. Singlitico et al. [8] examine the economic feasibility of hydrogen production from offshore wind power hubs for a 12 GW energy island located 380 km from the coast of Denmark. Finally, Travaglini et al. [9] analyze various large-scale offshore wind-based hydrogen production systems in the Dutch North Sea region, and their simulation framework serves as the foundation for this study.

Although such simulation-based studies can provide valuable insights into system design, costs, and operational performance, the outputs of these models are often highly dependent on user-provided inputs. An input may reflect either a decision or a parameter: a decision corresponds to a controllable choice (e.g., the sizing of an electrolyzer), whereas a parameter reflects exogenous information (e.g., the costs associated with installing offshore wind turbines). When modeling future hydrogen production systems, for which empirical data are limited and expert opinions can vary considerably, it is often unclear which decisions ought to be made or what the parameter settings of a simulation model should be.

A common technique for addressing the issue of input dependence is to perform a sensitivity (or Monte Carlo) analysis, which examines how changes to a model's inputs affect its outputs [10]. However, this type of analysis does not distinguish between decisions and parameters, limiting its prescriptive value. Furthermore, performing such an analysis in a comprehensive manner (e.g., a global sensitivity analysis) may require a computationally infeasible number of simulation runs.

Alternatively, one could address the issue via simulation–based optimization, which does distinguish between decisions and parameters. An overview of this broad class of methods is given by Amaran et al. [11]. While heuristic and metaheuristic algorithms (such as local search and simulated annealing) are popular and can be effective in certain settings [12], such algorithms require the objective and constraint(s) to be known a priori and are not designed to provide insight into the input-output relationships of a simulation model. Another approach within simulation-based optimization is to use surrogate modeling, where statistical or machine learning models are used to approximate the simulation outputs and guide the search for optimal solutions; see Cozad et al. [13] and Kleijnen [14] for overviews.

In this paper, we utilize a surrogate modeling approach based on recent advances in "optimization with constraint learning", see Fajemisin et al. [15] for a survey on the topic. Our approach, inspired by Maragno et al. [16], restricts itself to using surrogate machine learning (ML) models that can be directly embedded into mixed-integer optimization (MIO) problems. This class of models offers two key advantages over classical surrogate models, such as Kriging. First, modern ML models can be efficiently trained on large datasets, which enables the approach to be more effective in data-rich settings. Second, by ensuring that the surrogate models are "MIO-representable", one can directly optimize over them by leveraging commercial MIO solvers, thereby avoiding the need for custom optimization algorithms. This direct integration between ML and optimization allows the objective to be redefined, constraints to be added or removed, and the resulting problem(s) to be solved, with minimal additional computation time.

For an example of a more conventional surrogate modeling approach, we refer to Liu et al. [17], who also combine simulation with machine learning and optimization to study hydrogen production. Note that their approach considers a single fixed problem instance and involves a custom genetic algorithm to optimize over the surrogate models. Our approach differs, as we are able to consider multiple stakeholder-defined problem instances and solve these using a standard commercial MIO solver.

1.1 Contribution and novelty

We present an effective approach for optimizing system design decisions related to offshore wind-based green hydrogen production. A key advantage of our approach, with respect to alternative simulation-based optimization methods, is its flexibility. The user is able to modify the objective or the constraints without having to re-solve the problem from scratch. This is particularly beneficial in dynamic settings with multiple stakeholders and objectives. Our approach aligns with a "human-in-the-loop" approach to decision-making [18], where the provided solution(s) can be refined through interaction between stakeholders and the decision support tool. Additionally, our approach enables the user to quickly assess the cost of imposing a constraint, or to efficiently analyze specific situations and explore tradeoffs.

To the best of our knowledge, we are the first to apply optimization with constraint learning to this problem domain. While this approach has been successfully applied in other domains, such as radiotherapy [19], voltage regulation [20], molecular design [21], lot-sizing [22] and organ allocation [23], our application in optimizing green hydrogen production via offshore wind represents a novel contribution to the scientific literature.

1.2 Structure

The main body of the paper is organized as follows. In Section 2, we briefly describe the simulation framework developed by Travaglini et al. [9] and discuss how it is used herein. In Sections 3 and 4, we describe the problem we seek to address and the approach we take for tackling this problem. Finally, in Section 5, we present our results and in Sections 6 and 7, we discuss our findings and suggest some conclusions.

2 Simulation framework for green hydrogen production

The study builds on a techno-economic simulation framework that emulates green hydrogen production from dedicated offshore wind farms [9]. The simulation model [24] accurately captures the operational behavior of both alkaline (ALK) and proton exchange membrane (PEM) electrolyzers across a wide range of conditions. It supports the evaluation of electrolyzer types, as well as other system design choices, with a particular focus on the interactions between components and the intermittency of offshore wind power.

One of the primary goals of the framework is to estimate the "levelized cost of hydrogen" (LCOH). This metric is calculated by dividing the total sum of costs by the total amount of hydrogen that is produced by the system (within a certain time horizon).

The modeling framework consists of multiple steps, as shown in Figure 1. Starting from the estimation of the gross electrical output of the wind farm from real-world offshore wind speed and direction data, energy losses are subsequently estimated to determine the available power supplied to the electrolyzer at each timestep. A detailed electrolysis module is then employed to replicate the operational behavior of the electrolyzer stack and auxiliary systems, allowing for the calculation of the hydrogen output available for injection into an onshore hydrogen grid.

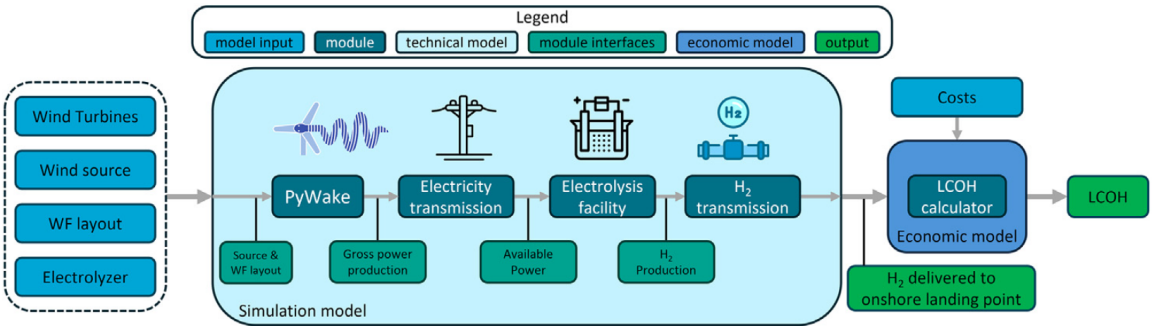

Figure 1: Overview of the techno-economic simulation framework developed by [9]

Highly detailed modules are used to simulate the wind farm and electrolyzer, see Appendix A for more information. Other components, such as pipelines, compressors, desalination units, and electrical infrastructure are emulated using simplified modules.

In this study, we use wind speed and direction data sampled at 10-minute intervals. This temporal resolution aligns with the typical averaging interval of wind farm power curves provided by manufacturers [25], and it enables reasonably accurate modeling of electrolyzer dynamics, including temperature variations, degradation behavior, and cold-start effects [9]. This is important as coarser time steps are likely to result in an overestimation of the hydrogen production potential [26].

3 Problem description

The techno-economic simulation framework described in the previous section is capable of accurately evaluating different configurations of offshore wind-based hydrogen production systems. However, the outputs of the simulation are highly dependent on the inputs, which consist of a

collection of decisions as well as various technical and economic parameters. To highlight this point, Travaglini et al. [9] perform a sensitivity analysis on their case study and find that the estimated LCOH can range anywhere from 3.0 to 10.5 euros per kilogram of hydrogen. The inputs and outputs of this simulation framework are visually illustrated in Figure 2.

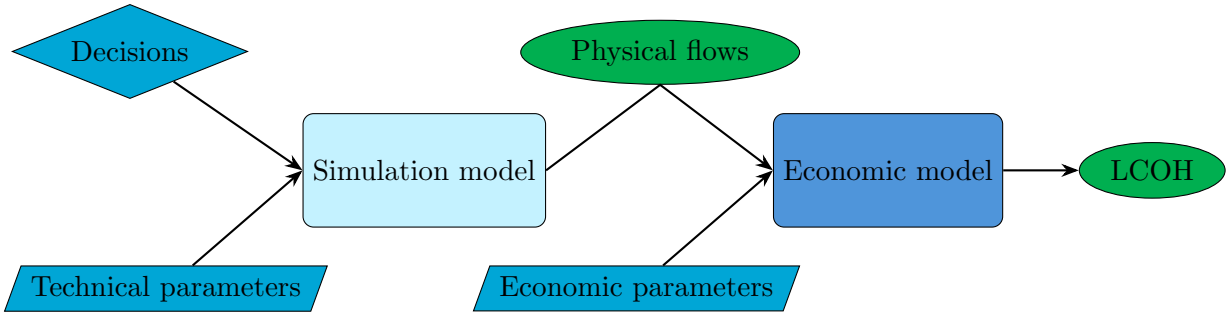

Figure 2: Visualization of the input-output relationships in the techno-economic simulation framework. The rectangles represent models, the diamond represents input decisions, the parallelograms represent input parameters and the ellipses represent model outputs. We use the same color scheme as in Figure 1.

The simulation model requires the user to specify a set of decisions, which define the configuration of the wind-to-hydrogen production system. These decisions span both continuous and discrete variables, including asset sizing, technology selection, and locational attributes. Table 1 provides an overview of the decision variables considered in this study, along with their respective domain and unit of measurement (UoM). These choices directly influence the technical performance of the system and indirectly influence the LCOH.

Table 1: Overview of the decisions under consideration

Decision	Type	Domain	UoM
Number of wind turbines	Integer	[34, 134]	[turbines]
Electrolyzer location	Discrete	{Onshore, Offshore}	—
Electrolyzer type	Discrete	{ALK, PEM}	—
Electrolyzer capacity	Integer	[5, 20]	[100 MW]
Distance to shore	Discrete	{50, 100, 200}	[km]
Cable type	Discrete	{AC, DC}	—

For each simulation model evaluation, technical input parameters are also required. Due to the relatively low technology readiness level of large-scale electrolyzer systems, the technical parameters considered most uncertain are the energy consumption and degradation rates of the ALK and PEM electrolyzers. For these uncertain parameters, we construct continuous interval domains using the lower and upper bounds from Travaglini et al. [9], which were determined using literature analysis and expert consultation. See Table 2 for an overview.

Table 2: Overview of the uncertain technical parameters

Parameter	Lower bound	Upper bound	UoM
Electrolyzer energy consumption (ALK)	47.1	55.6	[kWh/kg]
Electrolyzer energy consumption (PEM)	49.2	58.1	[kWh/kg]
Electrolyzer degradation rate (ALK)	4.00×10^{-6}	6.85×10^{-6}	[V/h]
Electrolyzer degradation rate (PEM)	4.29×10^{-6}	6.60×10^{-6}	[V/h]

After specifying the decisions and technical input parameters, the simulation model is able to simulate the physical flows through the system, and, given these flows, the economic model is able to estimate the LCOH. However, this last step requires specifying a number of economic input parameters and many of these parameters are considered to be uncertain [9]. Similarly to the uncertain technical parameters, we adopt the lower and upper bounds from Travaglini et al. [9] to construct interval domains. See Table 3 for an overview.

Table 3: Overview of the uncertain economic parameters

Parameter	Lower bound	Upper bound	UoM
CAPEX wind turbines	2,007,300	2,839,400	[€/MW]
O&M wind turbine	0.029	0.06	[% of CAPEX]
O&M reverse osmosis unit	0.02	0.025	[% of CAPEX]
CAPEX electrolyzer (ALK)	350,000	1,000,000	[€/MW]
O&M electrolyzer (ALK)	0.02	0.04	[% of CAPEX]
Replacement cost electrolyzer (ALK)	0.4	0.55	[% of CAPEX]
CAPEX electrolyzer (PEM)	500,000	1,400,000	[€/MW]
O&M electrolyzer (PEM)	0.015	0.04	[% of CAPEX]
Replacement cost electrolyzer (PEM)	0.4	0.55	[% of CAPEX]
CAPEX H_2 pipelines (onshore)	320,000	840,000	[€/km]
O&M H_2 pipelines (onshore)	0.008	0.07	[% of CAPEX]
CAPEX H_2 pipelines (offshore)	360,000	580,000	[€/km]
O&M H_2 pipelines (offshore)	0.008	0.07	[% of CAPEX]
CAPEX compressor (onshore)	800,000	4,000,000	[€/MW]
O&M compressor (onshore)	0.017	0.04	[% of CAPEX]
CAPEX compressor (offshore)	2,200,000	6,700,000	[€/MW]
O&M compressor (offshore)	0.017	0.04	[% of CAPEX]
CAPEX inter-array cables	194,000	500,000	[€/km]
O&M inter-array cables	0.002	0.022	[% of CAPEX]
CAPEX cables (AC)	3,000,000	6,840,000	[€/km]
O&M cables (AC)	0.005	0.025	[% of CAPEX]
CAPEX substructure (AC)	186,600	345,740	[€/MW]
O&M substructure (AC)	0.015	0.025	[% of CAPEX]
CAPEX cables (DC)	800,000	6,800,000	[€/km]
O&M cables (DC)	0.002	0.03	[% of CAPEX]
CAPEX substructure (DC)	565,000	807,750	[€/MW]
O&M substructure (DC)	0.015	0.023	[% of CAPEX]
CAPEX H_2 substation	141,000	318,000	[€/MW]
O&M H_2 substation	0.002	0.01	[% of CAPEX]

The LCOH, along with the quantity of hydrogen produced, are the two primary outputs of interest in this study. However, there are many more outputs that can be obtained from utilizing the simulation framework. In collaboration with industrial stakeholders and academic researchers, we construct a list of six key performance indicators for offshore wind-based hydrogen production. These are: amount of hydrogen produced, amount of brine produced as a byproduct, amount of electricity curtailed, electrolyzer capacity factor (CF), wind farm CF and LCOH. Note that this list extends beyond the techno-economic performance indicators that are commonly analyzed in contemporary studies, such as [6, 8, 27], as we also include brine waste due to its potential environmental impact [28].

Given this setup, we can now formalize our problem description. Let $y_1, \ldots, y_6$ represent the six outputs listed above, let the vector $\mathbf{x}$ represent the decisions under consideration and $\mathcal{X}$ the domain for these decisions. Furthermore, let $\mathbf{z}_t$ represent the uncertain technical parameters and let $\mathbf{z}_e$ represent the uncertain economic parameters. For each uncertain parameter, we assume that it resides within the lower and upper bounds specified in Tables 2 and 3. Each output y_k can be written as a function of the inputs, i.e., $y_k = F_k(\mathbf{x}, \mathbf{z}_t, \mathbf{z}_e)$, where F_k denotes utilizing the techno-economic simulation framework of Travaglini et al. [9] to evaluate output y_k, for $k = 1, \ldots, 6$.

Imagine that a stakeholder is interested in using the techno-economic simulation framework to determine optimal decisions for a specific problem instance, where a problem instance consists of the following: (i) a primary output y_j that he/she would like to minimize, (ii) a requirement b_k for each alternative output $k \neq j$ and (iii) a restricted decision domain $\mathcal{X}_P \subseteq \mathcal{X}$, where $\mathcal{X}_P$ may contain additional constraints or conditions imposed by the stakeholder or the specific problem context. This can be written as an optimization problem:

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & F_j(\mathbf{x}, \mathbf{z}_t, \mathbf{z}_e) \\ \text{subject to} \quad & F_k(\mathbf{x}, \mathbf{z}_t, \mathbf{z}_e) \leq b_k, \quad \forall k \neq j \\ & \mathbf{x} \in \mathcal{X}_P. \end{aligned} \tag{P}$$

Here we assume, without loss of generality, that lower values are preferable for all outputs, hence the minimization problem (note that one can flip the sign of the output if necessary).

There are two issues with solving Problem (P). First, the problem is ill-defined as $\mathbf{z}_t$ and $\mathbf{z}_e$ are uncertain. Second, the functions F_k are not assumed to have a known functional form, which implies that we are not able to derive any explicit gradients or structural properties that might aid in optimization. The functions may be queried for particular inputs and one could, in principle, attempt to solve (P) using brute-force enumeration, zero-order optimization techniques, or finite-difference gradient methods. However, due to the high dimensionality of $(\mathbf{x}, \mathbf{z}_t, \mathbf{z}_e)$ and the fact that each evaluation requires approximately 60 seconds, these approaches are rendered computationally prohibitive.

In the following section we describe an approach that enables us to approximately solve such stakeholder-defined problem instances in a computationally efficient manner (or provide reasonable alternative solutions if the problem instance is deemed infeasible).

4 Our approach

First, note that for the framework depicted in Figure 2, the simulation model requires approximately one minute per simulation, while the economic model is able to compute the LCOH within milliseconds. This implies that the simulation model is the primary computational bottleneck in the framework described in Section 2. Due to the ability to quickly compute the LCOH, we are able to "resolve" the uncertainty surrounding the economic parameters $\mathbf{z}_e$, by estimating appropriate risk measures, such as the expectation ($\mathbb{E}$) and conditional value-at-risk (CVaR), via Monte Carlo sampling [29]. As such, the methodology we present in this section focuses mainly on the inputs to the simulation model (the decisions $\mathbf{x}$ and the uncertain technical parameters $\mathbf{z}_t$).

Our methodology is as follows. First, we utilize a screening to determine which inputs are the most influential and use this information to exclude the least influential inputs from the rest of our analysis. Then, data is collected and used to construct surrogate ML models that are "MIO-representable" and computationally tractable. Finally, we use the surrogate models to approximately solve any arbitrary stakeholder-defined problem instance.

4.1 Input screening

The simulation model requires six decision variables ($x_1, \ldots, x_6$) and two uncertain technical parameters (z_1 and z_2) as inputs (see the description in Section 3 for more information). By identifying and retaining only the most influential inputs, we are able to reduce the dimensionality of the problem. This improves the signal-to-noise ratio, mitigates overfitting, and improves model interpretability [30]. Moreover, when our surrogate models are embedded into downstream MIO problems, a lower-dimensional input space is advantageous in terms of computational efficiency. As such, it is advisable to first perform an input screening and (potentially) discard the least impactful simulation inputs.

In settings with high experimental costs or a high number of inputs, it is advisable to use a screening design of experiments, such as the Plackett–Burman design [31]. However, in our setting we only consider eight inputs and can run a simulation within minutes. Therefore, we are able to utilize a 2-level full factorial design (where we treat the distance to shore as a 3-level input). This design prescribes 384 ($2 \times 2 \times 2 \times 3 \times 2 \times 2 \times 2 \times 2$) experiments/simulation runs that include all possible high/low permutations of the simulation input values. We refer to Montgomery [32] for more information on the topic of design of experiments.

For each of the 384 simulation runs, we evaluate the resulting LCOH across 10,000 economic scenarios. These scenarios are generated by independently sampling each uncertain economic parameter from a uniform distribution over its specified bounds (see Table 3), and are used to estimate risk measures for the LCOH.

We then standardize the input data and fit a linear regression model to the data by minimizing the sum of squared errors. The coefficients of the regression model provide rough estimates of the first-order (linear) effects of each simulation input, which we can then use to determine the relative influence of each input on each output of interest. Thus, for each simulation

run $i = 1, \ldots, 384$, we evaluate an input vector $(x_{1i}, \ldots, x_{6i}, z_{1i}, z_{2i})$ and observe each output y_{ki}. We use this data to fit the following model: $y_{ki} = \alpha + \beta_{k1}x_{1i} + \cdots + \beta_{k6}x_{6i} + \beta_{k7}z_{1i} + \beta_{k8}z_{2i} + \epsilon_i$, where β_{kj} provides an estimate of the effect of the j-th simulation input on output y_k.

To assess the significance of these effects, we examine the p-values p_{kj} associated with the null hypothesis $H_0 \colon \beta_{kj} = 0$ and identify the simulation inputs for which we find that $p_{kj} > 0.05, \forall k$. If this is the case, the corresponding simulation input is fixed to its nominal value (see Table 9 in Appendix A.3) and assumed to be constant during the rest of our analysis. Doing this allows us to potentially reduce the simulation input domain from $(\mathbf{x}, \mathbf{z}_t) \subseteq \mathbb{R}^8$ to a lower-dimensional subspace $(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t) \subseteq \mathbb{R}^d$, where $d \leq 8$ denotes the number of retained inputs post screening.

4.2 Data collection and machine learning

Having done an input screening and (potentially) fixed the inputs for which the estimated effect was not deemed sufficiently large, we collect data for the purpose of training ML models to predict our outputs of interest. We collect data from N simulation model evaluations and, for each simulation run, we also compute the resulting LCOH across 10,000 economic scenarios (which are randomly generated in the same manner as described in Section 4.1) to estimate risk measures for the LCOH. In our case, N is not fixed a priori, but increased incrementally until the ML models reach an acceptable degree of accuracy (based on out-of-sample testing).

We utilize the collected data to train and test machine learning models to predict the simulation outputs, using the simulation inputs as model features. The types of ML models we use in our approach must be MIO-representable, which implies that the input-output mapping of these models can be rewritten using continuous and integer variables, along with linear constraints [16]. This class of ML models includes: linear and logistic regressions (LR), support vector machines (SVM), classification and regression trees (CART), random forests (RF), extreme gradient boosting (XGB), gradient-boosting machines (GBM), multilayer perceptrons with a rectified linear unit activation function (MLP).

For each output y_k, we have access to a dataset $\{\hat{\mathbf{x}}_i, \hat{\mathbf{z}}_t, y_{ki}\}_{i=1}^N$ with N data points. These data can be used to construct a surrogate ML model f_k, where $f_k(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t) \approx y_k$. We do this as follows. First, we randomly splitting the data into training, validation, and testing sets, using an 80%-10%-10% split. Next, we perform hyperparameter tuning via grid search (see Table 10 in Appendix B), combined with 5-fold cross-validation on the training set. The best-performing model is selected based on its mean squared error on the validation data. The testing set is held out entirely during training and tuning, and is used exclusively to evaluate the final model's predictive accuracy. Thus, for each output y_k, a trained ML model f_k is saved and used in the following step.

4.3 Optimization using surrogate machine learning models

The final step in our approach is to embed our trained surrogate ML models into optimization problem formulations. The embedding of ML models into optimization problem formulations is commonly referred to as "constraint learning" (see Fajemisin et al. [15] for an overview of the topic). ML models can be embedded in various ways, see for example [33, 34]. For more

information on the specific embedding formulations used in this study, we refer to Section 2.2 of Maragno et al. [16].

Imagine we are interested in solving Problem (P), as described earlier in Section 3. Using our surrogate ML models (f), we can now approximate this problem and solve the following:

$$\begin{aligned} \min_{\hat{\mathbf{x}} \in \hat{\mathcal{X}}} \quad & f_j(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t) \\ \text{subject to} \quad & f_k(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t) \leq b_k, \quad \forall k \neq j \\ & \hat{\mathbf{x}} \in \hat{\mathcal{X}}_P. \end{aligned} \tag{$\hat{P}$}$$

Note that in this formulation, the difficult to optimize functions F_k are replaced by the more tractable surrogate models f_k. Additionally, the uncertain parameter vector $\mathbf{z}_t$ is replaced by $\hat{\mathbf{z}}_t$, which represents a deterministic parameter vector specified by the user.

Note that the uncertain economic parameters $\mathbf{z}_e$ do not appear in this formulation as our surrogate models for the LCOH-related outputs are trained to directly predict the expectation (or CVaR) of the LCOH with respect to the distribution of $\mathbf{z}_e$. We utilize Monte Carlo sampling to evaluate these risk measures, where we assume that each parameter is independently and uniformly distributed within its respective domain.

While MIO problems such as Problem $(\hat{P})$ are non-convex and, in general, "NP-hard", for many practical cases optimal (or near-optimal) solutions can be found in reasonable time using standard optimization solvers. As we demonstrate in Section 5.2, commercial solvers such as Gurobi are able to solve our problem instances within seconds.

5 Results

In this section, we present the results obtained by applying our methodology to the problem described in Section 3. First, in Section 5.1, we perform the input screening, summarize the collected simulation data and use this data to train (and evaluate the accuracy of) the surrogate ML models. In Section 5.2, we illustrate how our surrogate models can be used to obtain and evaluate solutions in a dynamic manner. In Section 5.3 we showcase how our surrogate models can be used to analyze specific situations and explore trade-offs. All computations are conducted on a 64-bit Windows machine equipped with a 2.80 GHz Intel Core i7 processor with 32 GB of RAM. All optimization problems are solved using Gurobi 12.0. The code is publicly available at: `https://github.com/JustinStarreveld/hydrogen-prod-simopt-via-ocl`.

5.1 Input screening, data and surrogate models

Table 4 presents the estimated first-order effects (β_{kj}) of each simulation input on the two primary outputs of interest: hydrogen production and the expected levelized cost of hydrogen. The estimated effects on the remaining outputs are provided in Table 11 in Appendix C. Statistically significant effects (with p-values below 0.05) are highlighted in bold.

We observe that the cable type and electrolyzer degradation rate exhibit no statistically significant effect on any of the considered outputs. Therefore, these two inputs are fixed to their

Table 4: Estimated effects (β_{kj}) of each simulation input on the two most important outputs of interest. Statistically significant effects (where the p-value is less than 0.05) are highlighted in bold.

Simulation input	Hydrogen production [kt/y]	$\mathbb{E}$(LCOH) [€/kg]
Number of wind turbines	**3.926**	**-1.195**
Electrolyzer location	-0.021	**0.186**
Electrolyzer type	0.232	**-1.423**
Electrolyzer capacity	**2.943**	-0.159
Distance to shore	-0.331	**0.739**
Cable type	0.006	0.135
Electrolyzer energy consumption	**-0.840**	**1.259**
Electrolyzer degradation rate	-0.024	0.035

nominal value and excluded from consideration in the following steps (data collection, surrogate modeling, and optimization). Specifically, the cable type is set to "AC", while the electrolyzer degradation rate is fixed at 5.04×10^{-6} for ALK and 4.85×10^{-6} for PEM.

This procedure results in a single remaining uncertain technical parameter: the electrolyzer power consumption (z_η), and five remaining decision variables: number of wind turbines (x_{WT}), electrolyzer location (x_{EL}), electrolyzer type (x_{ET}), electrolyzer capacity (x_{EC}) and distance to shore (x_{DTS}). Thus, we have:

$$\hat{\mathbf{x}} = (x_{WT}, x_{EL}, x_{ET}, x_{EC}, x_{DTS}), \text{ and } \hat{\mathbf{z}}_t = (z_\eta). \tag{1}$$

To characterize the boundaries of the reduced simulation input domain $(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t) \subseteq \mathbb{R}^6$, we construct a 2-level full factorial design of experiments comprising 96 ($2 \times 2 \times 2 \times 3 \times 2 \times 2$) simulation runs. In addition, we collect data from 4,000 simulation runs, where the input values are randomly and uniformly sampled from their respective domains. Collecting this data was done in parallel (using 4 cores) and took approximately 48 hours. A visual summary of the resulting 4,096 data points is provided in Figure 3.

Figure 3 showcases the substantial variability in the simulation framework's outputs that arises from variation in the simulation inputs. For example, the expected LCOH can be anywhere from 4 to 29 euros per kilogram of hydrogen, depending on $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}_t$.

In the next step, we use this data to train surrogate ML models to predict an output on the basis of the simulation inputs. Table 5 reports the best-performing model type for each output, together with its out-of-sample predictive accuracy.

The physical flow-related outputs (e.g., hydrogen production, brine waste, curtailment and the capacity factors) are predicted with high accuracy, as the mean absolute percentage errors (MAPE) is $\leq 3\%$ for these outputs. In comparison, the cost-related outputs are more challenging to predict, with mean absolute percentage errors of 11% and 10%. While less precise, these models can still function as effective guides in the optimization phase of our methodology, which we demonstrate in the following section.

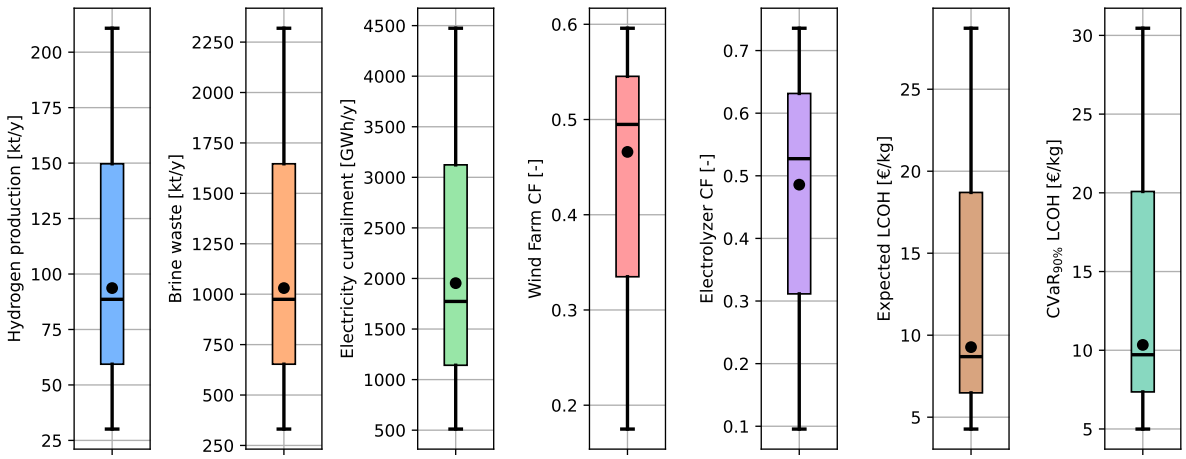

Figure 3: Box and whisker plots for each output in our dataset. The mean is depicted as a black circle.

Table 5: Results from fitting the ML models to our data. We report the best model, along with its coefficient of determination (R^2), mean absolute error (MAE) and mean absolute percentage error (MAPE), as evaluated on the testing data.

Output	UoM	Best model	R^2	MAE	MAPE
Hydrogen production	[kt/y]	XGB	0.99	2.35	0.03
Brine waste	[kt/y]	XGB	0.99	26	0.03
Electricity curtailment	[GWh/y]	MLP	0.99	39	0.02
Wind farm capacity factor	[-]	GBM	0.98	0.01	0.02
Electrolyzer capacity factor	[-]	MLP	0.99	0.01	0.02
$\mathbb{E}$(LCOH)	[€/kg]	GBM	0.73	1.05	0.11
$\text{CVaR}_{90\%}$(LCOH)	[€/kg]	MLP	0.74	1.06	0.10

5.2 Flexible exploration of solutions

This section illustrates how, once we have the trained ML models, we are able to provide a fast and flexible optimization tool, where the objective or constraints can be easily modified and a new solution can be obtained within seconds. Additionally, we illustrate how the obtained solutions can be evaluated using the simulation framework in order to verify the accuracy of the surrogate ML model predictions.

Suppose a stakeholder wishes to minimize the expected LCOH, under the assumption that the power consumption of the electrolyzer will be equal to the midpoint between its lower and upper bounds, i.e., $\hat{\mathbf{z}}_t^m = (z_\eta) = \left((z_{\eta,ET}^u - z_{\eta,ET}^l)/2\right)$, where $z_{\eta,ET}^u$ and $z_{\eta,ET}^l$ denote the upper and lower bounds of the electrolyzer energy consumption for electrolyzer type $ET \in \{\text{ALK, PEM}\}$. This can be written as follows:

$$\min_{\hat{\mathbf{x}} \in \hat{\mathcal{X}}} \; f_{\mathbb{E}(\text{LCOH})}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^m). \tag{P_1}$$

By embedding the best-performing ML model for the output $\mathbb{E}$(LCOH) into a mixed-integer

optimization problem formulation (see Section 4.3) and solving the MIO problem using Gurobi, we obtain a solution to Problem (P_1) within 0.3 seconds. Let this solution be denoted as $\hat{\mathbf{x}}^\star_{P_1}$. The decisions corresponding to $\hat{\mathbf{x}}^\star_{P_1}$ are shown in Table 6.

Table 6: Solution to Problem (P_1)

Decision	Value	UoM
Number of wind turbines	89	[turbines]
Electrolyzer location	Offshore	—
Electrolyzer type	PEM	—
Electrolyzer capacity	2000	[MW]
Distance to shore	50	[km]

After obtaining this solution, the stakeholder is able to evaluate the performance of the solution under variation in the electrolyzer energy consumption parameter (i.e., $z_\eta \in \{49.2, 53.7, 58.1\}$ kWh/kg). Using our surrogate ML models (f), we can provide a prediction for the performance within milliseconds. Additionally, we can utilize the simulation framework (F) to evaluate the accuracy of these predictions (this requires three simulation evaluations and takes approximately three minutes in total). The performance of solution $\hat{\mathbf{x}}^\star_{P_1}$ is shown in Figure 4.

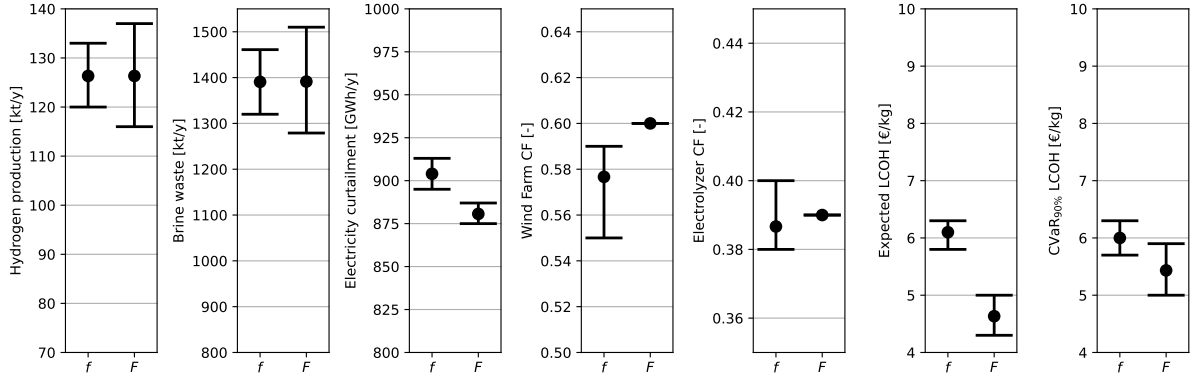

Figure 4: Performance evaluation of the solution described in Table 6. We show the surrogate model prediction (f) and the simulated performance (F) for each output, where the electrolyzer energy consumption parameter is set to its lower bound, midpoint or upper bound.

Now suppose that the stakeholder does not find this solution to be particularly desirable. For example, the stakeholder may wish to limit the amount of brine waste to a maximum of 1,000 kilotons per year. Additionally, the stakeholder may be reluctant to support the construction of a wind farm 50 kilometers from the shoreline, as such proximity may result in visual intrusion on the coastal landscape. Due to the flexibility of our approach, the stakeholder can simply add

these constraints to the optimization problem formulation:

$$\begin{aligned} \min_{\hat{\mathbf{x}} \in \hat{\mathcal{X}}} \quad & f_{\mathbb{E}(\text{LCOH})}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^m) \\ \text{s.t.} \quad & f_{\text{Brine waste}}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^m) \leq 1000 \text{ [kt/y]}, \\ & x_{DTS} \geq 100 \text{ [km]}, \end{aligned} \tag{P_2}$$

and solve Problem (P_2) to obtain a new solution $\hat{\mathbf{x}}^\star_{P_2}$. This solution is obtained within 0.4 seconds and is described in Table 7. We find that the electrolyzer location and type are unchanged, however, the number of wind turbines is reduced from 89 to 71, the electrolyzer capacity is reduced from 2000 MW to 1400 MW and the distance to shore is increased to 200 kilometers.

Table 7: Solution to Problem (P_2)

Decision	Value	UoM
Number of wind turbines	71	[turbines]
Electrolyzer location	Offshore	—
Electrolyzer type	PEM	—
Electrolyzer capacity	1400	[MW]
Distance to shore	200	[km]

Again, it is possible to evaluate the performance of this new solution using the surrogate ML models and the simulation framework, under variation w.r.t. the electrolyzer energy consumption parameter (i.e. $z_\eta \in \{49.2, 53.7, 58.1\}$ kWh/kg). These values are shown in Figure 5.

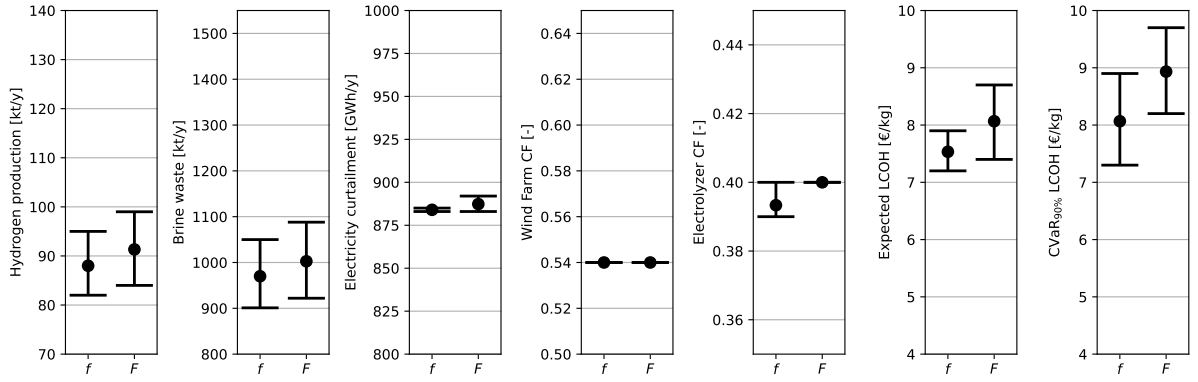

Figure 5: Performance evaluation of the solution described in Table 7. We show the surrogate model prediction (f) and the simulated performance (F) for each output, where the electrolyzer energy consumption parameter is set to its lower bound, midpoint or upper bound.

The results show that our new solution does not adhere to the brine waste restriction when the electrolyzer energy consumption is low (i.e., $z_\eta = 49.2$ kWh/kg). For this solution, the expected LCOH is higher, at 7.4, 8.1, or 8.7 euros per kilogram (depending on z_η). In addition, the CVaR at the 90% level can exceed 9.5 euros per kilogram, potentially raising concerns about the cost-efficiency of this solution under adverse scenarios.

Suppose that the stakeholder wishes to address such concerns and seeks a more robust solution. The stakeholder could set $\hat{\mathbf{z}}_t^u = \left(z_{\eta,ET}^u\right)$, and $\hat{\mathbf{z}}_t^l = \left(z_{\eta,ET}^l\right)$, for electrolyzer type $ET \in \{\text{ALK, PEM}\}$ and solve the following problem instance:

$$\begin{aligned} \min_{\hat{\mathbf{x}} \in \hat{\mathcal{X}}} \quad & f_{CVaR_{90\%}(\text{LCOH})}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^u) \\ \text{s.t.} \quad & f_{\text{Brine waste}}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^l) \leq 1000 \text{ [kt/y]}, \\ & x_{DTS} \geq 100 \text{ [km]}. \end{aligned} \tag{P_3}$$

Note that Problem (P_3) utilizes a different objective to Problem (P_2), where we minimize the conditional value at risk instead of the expectation. Additionally, in the objective we assume that the electrolyzer power consumption will equal its upper bound value, while in the brine waste constraint we assume that this parameter will equal its lower bound value. The solution $\hat{\mathbf{x}}_{P_3}^\star$ to Problem (P_3) is found within 0.3 seconds and is described in Table 8.

Table 8: Solution to Problem (P_3)

Decision	Value	UoM
Number of wind turbines	63	[turbines]
Electrolyzer location	Onshore	—
Electrolyzer type	PEM	—
Electrolyzer capacity	1200	[MW]
Distance to shore	200	[km]

This solution changes the location of the electrolyzer to be onshore instead of offshore, employs fewer wind turbines, and opts for a lower electrolyzer capacity. When the performance of $\hat{\mathbf{x}}_{P_3}^\star$ is evaluated using our surrogate models and the simulation framework (see Figure 6), we observe that this solution improves upon $\hat{\mathbf{x}}_{P_2}^\star$ in regards to the LCOH-related outputs and that our new solution $\hat{\mathbf{x}}_{P_3}^\star$ adheres to the brine waste restriction, even when the electrolyzer power consumption is at its lower bound.

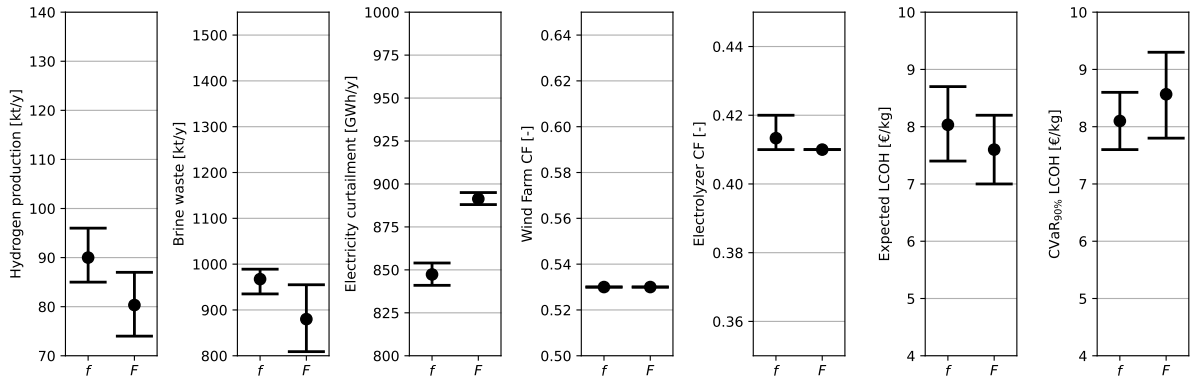

Figure 6: Performance evaluation of the solution described in Table 8. We show the surrogate model prediction (f) and the simulated performance (F) for each output, where the electrolyzer energy consumption parameter is set to its lower bound, midpoint or upper bound.

Figure 7 provides a comparative overview of the performance of the three solutions ($\hat{\mathbf{x}}^\star_{P_1}$, $\hat{\mathbf{x}}^\star_{P_2}$ and $\hat{\mathbf{x}}^\star_{P_3}$) across the seven outputs of interest. By visualizing their simulated performance side by side, a stakeholder can quickly discern the relative strengths and weaknesses of each solution, as well as attain an understanding of the robustness of each solution with respect to variation in the uncertain parameters of the model (in this case the electrolyzer energy consumption).

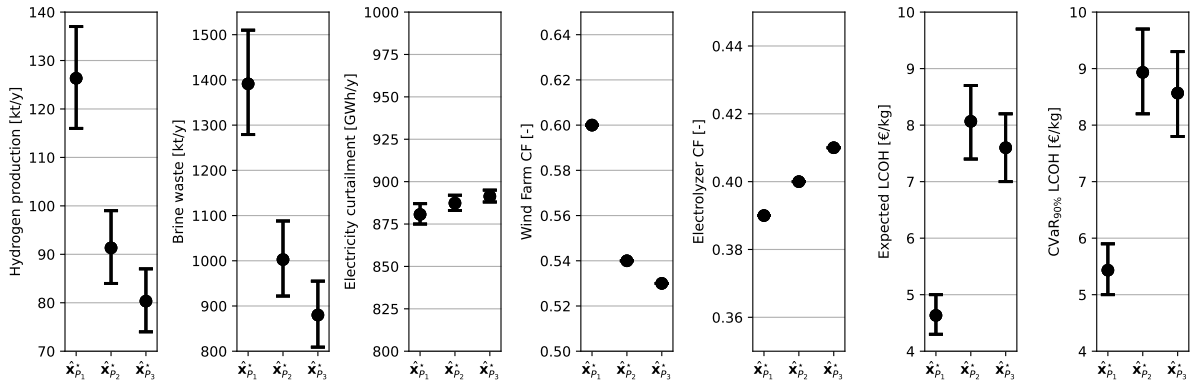

Figure 7: Comparison of simulation performance (F) for the three solutions ($\hat{\mathbf{x}}^\star_{P_1}$, $\hat{\mathbf{x}}^\star_{P_2}$ and $\hat{\mathbf{x}}^\star_{P_3}$), where the electrolyzer energy consumption parameter is set to its lower bound, midpoint or upper bound.

Finally, note that each evaluation performed using the simulation framework yields an additional data point that can be used to retrain and refine the surrogate ML models. Implementing such a continuous feedback loop can improve model accuracy and decision-making over time.

5.3 Analyzing specific situations and tradeoffs

This section demonstrates how the trained surrogate models can also be used to explore a variety of situations and trade-offs. By formulating and solving different optimization problems, we gain insights into how the decisions and parameters influence our outputs under various constraints and assumptions.

Suppose that a stakeholder is interested in knowing the maximum amount of hydrogen production, given a certain number of wind turbines and a fixed electrolyzer capacity of 1000 MW. We can (approximately) obtain this information by solving the following optimization problem:

$$\begin{aligned} \max_{\hat{\mathbf{x}}\in\hat{\mathcal{X}}} \quad & f_{\text{Hydrogen production}}(\hat{\mathbf{x}}, \hat{\mathbf{z}}^\alpha_t) \\ \text{s.t.} \quad & x_{EC} = 1000 \text{ [MW]}, \\ & x_{WT} = \theta \text{ [turbines]}, \end{aligned} \tag{2}$$

where the parameter θ is varied between $[34, 134]$ and the parameter vector $\hat{\mathbf{z}}^\alpha_t$ is set equal to the lower bound, midpoint, or upper bound for electrolyzer energy consumption. The results are displayed in Figure 8, where each marker in the graph represents an optimal solution to Problem (2).

Figure 8 shows that when the number of wind turbines is higher than 67 and the capacity of the wind farm exceeds the capacity of the electrolyzer (in this case 1000 MW), the increase

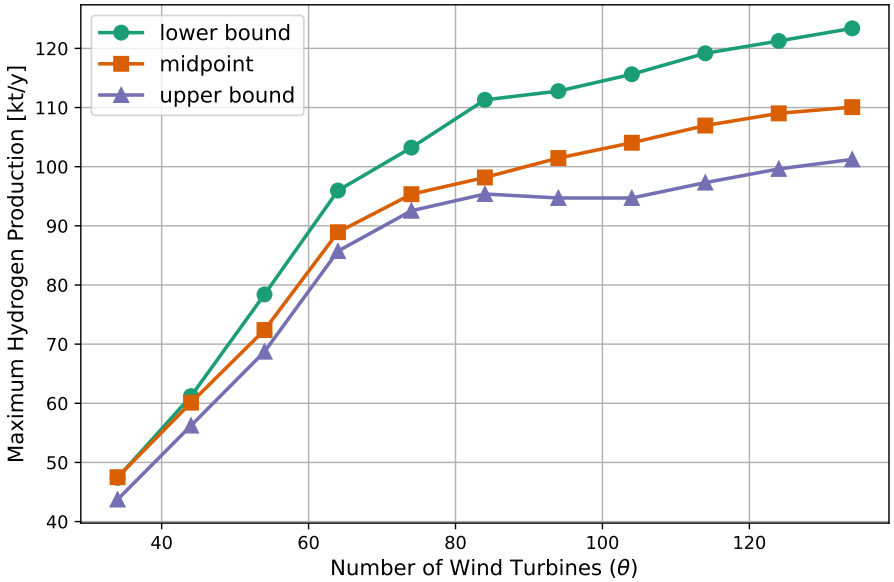

Figure 8: Maximum hydrogen production as a function of the number of wind turbines and the electrolyzer energy consumption, where we assume a fixed electrolyzer capacity of 1000 MW.

in hydrogen production tapers off. Furthermore, we observe that the effect of the electrolyzer energy consumption parameter is larger when the number of wind turbines is higher. For example, note that, when using more than 100 wind turbines, variation in electrolyzer efficiency can cause the maximum hydrogen production to vary by more than 20 kilotons per year.

In the following paragraphs we show how our methodology can also be used to efficiently explore tradeoffs. Suppose that a stakeholder is interested in the utilization of the wind farm and the electrolyzer, as measured by their respective capacity factors (CFs). A higher CF signals more effective utilization of the asset. However, improvements in one asset's utilization may come at the cost of the other. For example, oversizing the wind farm relative to the electrolyzer is expected to benefit the utilization of the electrolyzer, but is also likely to lead to increased electricity curtailment and a low wind farm CF. This tradeoff can be investigated in a quantitative manner by solving:

$$\begin{aligned} \max_{\hat{\mathbf{x}} \in \hat{\mathcal{X}}} \quad & \lambda \cdot f_{\text{Electrolyzer CF}}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^m) + (1-\lambda) \cdot f_{\text{Wind farm CF}}(\hat{\mathbf{x}}, \hat{\mathbf{z}}_t^m) \\ \text{s.t.} \quad & x_{ET} = \kappa, \end{aligned} \tag{3}$$

where the parameter λ is varied between $[0, 1]$, the technical parameter vector is set equal to the midpoint $\hat{\mathbf{z}}_t^m$ and κ is set to be either PEM or ALK.

The resulting Pareto-type frontiers are displayed in Figure 9, where each marker in the figure represents an optimal solution to Problem (3). We observe that the decisions under consideration ($\hat{\mathbf{x}}$) can lead to very different CFs for both assets, highlighting the importance of aligning the wind farm and electrolyzer capacities. As anticipated, the alkaline electrolyzer exhibits lower capacity factors than the PEM electrolyzer, reflecting its stricter minimum load requirements and reduced operational flexibility.

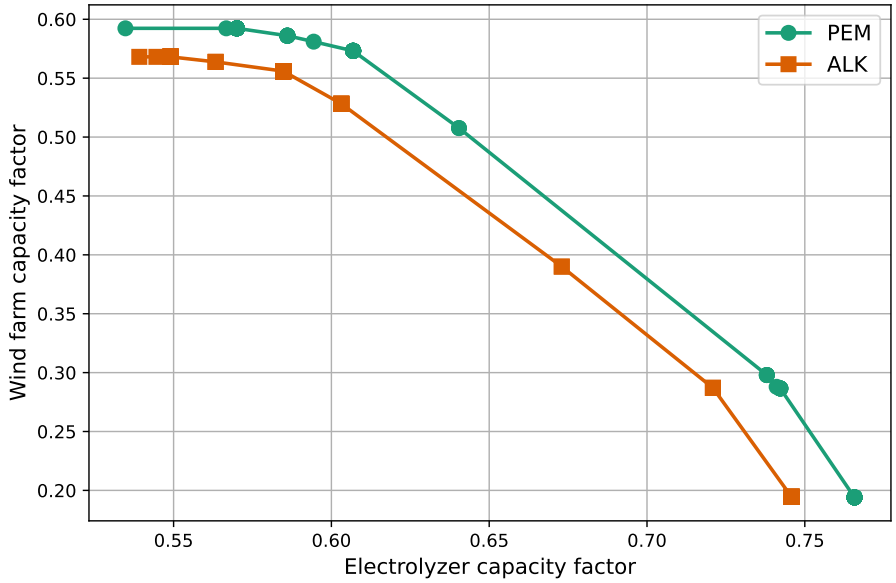

Figure 9: Visual representation of the tradeoff between electrolyzer and wind farm capacity factors for the two electrolyzer types.

6 Discussion

From a practical perspective, a key advantage of our approach with respect to alternative simulation-based optimization methods is the ability to easily incorporate adjustments to the objective or constraints without having to start from scratch. This enables industrial stakeholders and policy makers to explore various objectives or introduce site-specific or preference-based constraints (such as limiting brine waste or excluding nearshore development) and obtain practically relevant solutions in real time.

From a scientific perspective, optimization with constraint learning can enhance traditional simulation-based studies by providing additional insights. The approach is complementary to a dedicated sensitivity analysis, as it not only reveals how inputs influence outputs but also offers prescriptive information by predicting "optimal" solutions to specific situations. This allows one to analyze the "cost" of imposing certain constraints and explore potential tradeoffs.

Nevertheless, our approach also has some limitations. First, the accuracy of our surrogate machine learning models is inherently limited by the quality and quantity of data that can be obtained, as well as the expressive power of the ML models to which we fit the data. Throughout this paper, we use the simulation model as our "ground truth". While highly detailed and grounded in engineering knowledge, it ultimately remains a model of reality and cannot fully capture all real-world complexities. Empirical data may offer an advantage in this regard. However, field data are not always available, and physical experiments are typically more costly and time-consuming than computer simulations.

Second, the surrogate models are trained (and tested) on data that are generated within a specific domain. If a stakeholder's interest shift outside this predefined domain, the model predictions may become unreliable. This limitation can be partially addressed by defining a

"trust region", as proposed in Maragno et al. [16], to ensure that surrogate model usage remains within certain boundaries. Future work could improve generalizability of the surrogate ML models by incorporating active learning or adaptive sampling to dynamically expand the dataset based on stakeholder needs, observed prediction errors or optimization results.

Third, we assume that the uncertain parameters are independent and uniformly distributed. In practice, these assumptions may not hold, as parameters could be correlated or non-uniformly distributed. As a result, risk measures such as expectation and conditional value-at-risk should be interpreted with caution, as they are sensitive to the underlying distributional assumptions. Nonetheless, our methodology is capable of incorporating more realistic probabilistic assumptions in regard to the uncertain parameters, but this would require additional data collection. An interesting direction for future research is to integrate robust or stochastic optimization techniques to handle parameter uncertainty in a more systematic way.

Fourth, our approach cannot guarantee global optimality with respect to the original techno-economic simulation framework. While we are able to validate the obtained solutions using the simulation framework, approximation errors may still cause superior solutions to be overlooked. In future work, the accuracy of the surrogate models could be quantified more rigorously, for example by deriving statistical probability guarantees. Although such guarantees would require additional assumptions, they may increase trust in the quality of the obtained solutions.

7 Conclusions

In the study described in this paper, we have used a techno-economic simulation framework, in combination with mixed-integer optimization with constraint learning, to assess green hydrogen production from offshore wind in the Dutch North Sea. Our data-driven approach not only provides insights into the estimated effects of decisions and parameters, but also provides a computationally efficient decision support tool, enabling fast and interactive exploration of various objectives and constraints. Furthermore, our approach accounts for uncertain parameters and facilitates the analysis of tradeoffs in a way that traditional simulation-based studies cannot.

Our results demonstrate the effectiveness of such a combined approach. While our surrogate machine learning models achieve high predictive accuracy for most outputs, the models are less accurate for the two cost-based outputs. Nevertheless, we show that these models remain valuable for guiding optimization, especially given that the obtained solutions can subsequently be validated using the simulation framework.

We hope this work inspires future research and applications, as we see significant potential for this approach to further support decision making in the complex and uncertain transition towards sustainable energy systems.

Acknowledgments

This work was supported by the Netherlands Organization for Scientific Research (NWO) under grant ESI.2019.003.

References

[1] European Commission and Directorate-General for Communication. *REPowerEU, joint European action for more affordable, secure and sustainable energy*. Publications Office of the European Union, 2022. DOI: doi/10.2775/076377.

[2] Ministry of Economic Affairs and Climate Policy. *Waterstof*. https://www.rijksoverheid.nl/onderwerpen/duurzame-energie/waterstof. Accessed: Nov. 12, 2024. [In Dutch]. 2024.

[3] Nexstep. *The Dutch government has granted PosHYdon, the world's first offshore green hydrogen pilot on a working platform, the DEI+ subsidy*. https://www.nexstep.nl/the-dutch-government-has-granted-poshydon-the-worlds-first-offshore-green-hydrogen-pilot-on-a-working-platform-the-dei-subsidy/. Accessed: Jul. 25, 2025. 2021.

[4] RWE Renewables GmbH. *RWE and Neptune Energy join forces to accelerate green hydrogen production in Dutch North Sea*. https://www.rwe.com/en/press/rwe-renewables/2022-02-15-rwe-and-neptune-energy-join-forces-to-accelerate-green-hydrogen-production-in-dutch-north-sea/. Accessed: Jul. 25, 2025. 2022.

[5] T. Egeland-Eriksen and S. Sartori. "Techno-economic analysis of the effect of a novel price-based control system on the hydrogen production for an offshore 1.5 GW wind-hydrogen system". In: *Energy Reports* 11 (2024), pp. 2633–2655.

[6] S. J. P. Hill, O. Bamisile, L. Hatton, I. Staffell, and M. Jansen. "The cost of clean hydrogen from offshore wind and electrolysis". In: *Journal of Cleaner Production* 445 (2024), p. 141162.

[7] G. Calado, R. Castro, A. J. Pires, and M. J. Marques. "Assessment of hydrogen-based solutions associated to offshore wind farms: The case of the Iberian Peninsula". In: *Renewable and Sustainable Energy Reviews* 192 (2024), p. 114268.

[8] A. Singlitico, J. Østergaard, and S. Chatzivasileiadis. "Onshore, offshore or in-turbine electrolysis? Techno-economic overview of alternative integration designs for green hydrogen production into Offshore Wind Power Hubs". In: *Renewable and Sustainable Energy Transition* 1 (2021), p. 100005.

[9] R. Travaglini, L. Frowijn, A. Bianchini, Z. Lukszo, and K. Bruninx. "Offshore or onshore hydrogen production? A critical analysis on costs and operational considerations for the Dutch North Sea". In: *Applied Energy* 397 (2025), p. 126290.

[10] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, et al. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Wiley Online Library, 2004.

[11] S. Amaran, N. V. Sahinidis, B. Sharda, and S. J. Bury. "Simulation optimization: A review of algorithms and applications". In: *Annals of Operations Research* 240 (2016), pp. 351–380.

[12] W. T. de Sousa Junior, J. A. B. Montevechi, R. de Carvalho Miranda, and A. T. Campos. "Discrete simulation-based optimization methods for industrial engineering problems: A systematic literature review". In: *Computers & Industrial Engineering* 128 (2019), pp. 526–540.

[13] A. Cozad, N. V. Sahinidis, and D. C. Miller. "Learning surrogate models for simulation-based optimization". In: *AIChE Journal* 60.6 (2014), pp. 2211–2227.

[14] J. P. Kleijnen. "Regression and Kriging metamodels with their experimental designs in simulation: A review". In: *European Journal of Operational Research* 256.1 (2017), pp. 1–16.

[15] A. O. Fajemisin, D. Maragno, and D. den Hertog. "Optimization with constraint learning: A framework and survey". In: *European Journal of Operational Research* 314.1 (2024), pp. 1–14.

[16] D. Maragno, H. Wiberg, D. Bertsimas, Ş. İ. Birbil, D. den Hertog, and A. O. Fajemisin. "Mixed-integer optimization with constraint learning". In: *Operations Research* 73.2 (2025), pp. 1011–1028.

[17] Z. Liu, Z. Cui, M. Wang, B. Liu, and W. Tian. "A machine learning proxy based multi-objective optimization method for low-carbon hydrogen production". In: *Journal of Cleaner Production* 445 (2024), p. 141377.

[18] F. Lombardi and S. Pfenninger. "Human-in-the-loop MGA to generate energy system design options matching stakeholder needs". In: *PLOS Climate* 4.2 (2025), e0000560.

[19] D. Maragno, G. Buti, Ş. İ. Birbil, Z. Liao, T. Bortfeld, D. den Hertog, and A. Ajdari. "Embedding machine learning based toxicity models within radiotherapy treatment plan optimization". In: *Physics in Medicine & Biology* 69.7 (2024), p. 075003.

[20] Y. Chen, Y. Shi, and B. Zhang. "Data-driven optimal voltage regulation using input convex neural networks". In: *Electric Power Systems Research* 189 (2020), p. 106741.

[21] T. McDonald, C. Tsay, A. M. Schweidtmann, and N. Yorke-Smith. "Mixed-integer optimisation of graph neural networks for computer-aided molecular design". In: *Computers & Chemical Engineering* 185 (2024), p. 108660.

[22] D. Tremblet, S. Thevenin, and A. Dolgui. "Constraint learning approaches to improve the approximation of the capacity consumption function in lot-sizing models". In: *European Journal of Operational Research* 322.2 (2025), pp. 679–692.

[23] T. Papalexopoulos, J. Alcorn, D. Bertsimas, R. Goff, D. Stewart, and N. Trichakis. "Reshaping national organ allocation policy". In: *Operations Research* 72.4 (2024), pp. 1475–1486.

[24] R. Travaglini and L. Frowijn. *Hydrogen production simulation tool*. Version 1.0.0. 2025. URL: `https://github.com/RTrava/H2_prod_sim`.

[25] V. Sohoni, S. C. Gupta, and R. K. Nema. "A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems". In: *Journal of Energy* 2016.1 (2016), e8519785.

[26] R. Travaglini, F. Superchi, F. Lanni, G. Manzini, L. Serri, and A. Bianchini. "Towards the development of offshore wind farms in the Mediterranean Sea: A techno-economic analysis including green hydrogen production during curtailments". In: *IET Renewable Power Generation* 18.15 (2024), pp. 3112–3126.

[27] A. Rogeau, J. Vieubled, M. de Coatpont, P. Affonso Nobrega, G. Erbs, and R. Girard. "Techno-economic evaluation and resource assessment of hydrogen production through offshore wind farms: A European perspective". In: *Renewable and Sustainable Energy Reviews* 187 (2023), p. 113699.

[28] O. S. Ibrahim, A. Singlitico, R. Proskovics, S. McDonagh, C. Desmond, and J. D. Murphy. "Dedicated large-scale floating offshore wind to hydrogen: Assessing design variables in proposed typologies". In: *Renewable and Sustainable Energy Reviews* 160 (2022), p. 112310.

[29] A. Shapiro. "Monte Carlo Sampling Methods". In: *Handbooks in Operations Research and Management Science* 10 (2003), pp. 353–425.

[30] D. C. Woods and S. M. Lewis. "Design of Experiments for Screening". In: *Handbook of Uncertainty Quantification*. Springer, 2017, pp. 1143–1185.

[31] R. Plackett and J. Burman. "The design of optimum multifactorial experiments". In: *Biometrika* 33.4 (1946), pp. 305–325.

[32] D. C. Montgomery. *Design and Analysis of Experiments*. John Wiley & Sons, 2017.

[33] R. Anderson, J. Huchette, W. Ma, C. Tjandraatmadja, and J. P. Vielma. "Strong mixed-integer programming formulations for trained neural networks". In: *Mathematical Programming* 183.1 (2020), pp. 3–39.

[34] B. L. Ammari, E. S. Johnson, G. Stinchfield, T. Kim, M. Bynum, W. E. Hart, J. Pulsipher, and C. D. Laird. "Linear model decision trees as surrogates in optimization of engineering applications". In: *Computers & Chemical Engineering* 178 (2023), p. 108347.

[35] M. M. Pedersen et al. "PyWake 2.5.0: An open-source wind farm simulation tool". In: *DTU* (2023). Publisher: DTU Wind, Technical University of Denmark. URL: `https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake`.

[36] I. Katic, J. Højstrup, and N. O. Jensen. "A simple model for cluster efficiency". In: *European Wind Energy Association Conference and Exhibition*. A. Raguzzi. 1987, pp. 407–410.

[37] T. Morten Lybech. *Introduction to wind turbine wake modelling and wake generated turbulence*. 2021. URL: `https://help.emd.dk/knowledgebase/content/ReferenceManual/Wake_Model.pdf` (visited on 04/19/2023).

[38] J. Tiktak. "Heat Management of PEM Electrolysis: A study on the potential of excess heat from medium- to large-scale PEM electrolysis and the performance analysis of a dedicated cooling system". PhD thesis. TU Delft, 2019. URL: `https://repository.tudelft.nl/islandora/object/uuid%3Ac046820a-72bc-4f05-b72d-e60a3ecb8c89` (visited on 06/20/2024).

[39] F. Superchi, F. Papi, A. Mannelli, F. Balduzzi, F. M. Ferro, and A. Bianchini. "Development of a reliable simulation framework for techno-economic analyses on green hydrogen production from wind farms using alkaline electrolyzers". In: *Renewable Energy* 207 (2023), pp. 731–742.

Appendix

A Additional information on simulation framework

A.1 Wind power production

To simulate the performance of the wind farm, the open-source Python package PyWake [35] has been adopted due to its computational efficiency and widespread use in applications of wind energy. PyWake enables the modeling of wake interactions under steady-state conditions and facilitates time-varying analysis of individual turbine output. Its modular structure allows for seamless integration with additional Python-based components, such as electrolysis systems and transmission models.

The simulation relies on PyWake's Wind Farm module, which comprises site and turbine objects. The site object specifies wind conditions based on wind turbine layout, reference speed, and direction, while the turbine object defines the power and thrust curves, the hub height, and the rotor diameter of aerogenerators.

To account for wake losses, the N.O.J. model [36] has been applied, which assumes a linearly expanding wake with a constant decay coefficient. Offshore-specific conditions, relevant to the Dutch North Sea context, have been modeled in accordance with DTU guidelines, omitting terrain-related influences such as surface roughness and orography [37].

A.2 Electrolyzer

To accurately capture hydrogen production under fluctuating power inputs, the lumped parameter electrolysis model presented in [9] has been adopted for the simulation of both alkaline (ALK) and proton exchange membrane (PEM) technologies. This model simulates the electrolyzer behavior relying on the polarization curves that describe the relationship between cell voltage and current density, accounting for key electrochemical losses: activation, ohmic, and concentration. The model takes into account the impact of thermal conditions, with heat generation and dissipation computed from geometric parameters, as well as degradation on the component efficiency. The impact of temperature variation on performance is modeled through voltage shifts, approximated at 0.5 mV/°C for ALK and 0.4 mV/°C for PEM [38, 39].

Hydrogen output is determined as a function of input power and a conversion coefficient, which is calculated at each timestep from the operating voltage, cell operating temperature, and aging. The model also incorporates degradation by assuming a linear increase in overvoltage with cumulative operating hours, allowing estimation of efficiency loss and stack replacements over the system's lifetime. A shutdown temperature of 30°C is imposed, with restart procedures based on cold-start durations of five minutes for PEM and twenty minutes for ALK, as informed by Singlitico et al. [8].

Regarding the control strategies, a homogeneous distribution of the available power to maximize the number of active stacks has been assumed. Finally, stack replacement is assumed upon exceeding critical voltage thresholds: 2.3 V for ALK and 2.23 V for PEM.

A.3 Nominal values for simulation inputs

Table 9: Nominal values for the simulation inputs

Decision	Nominal value	UoM
Number of wind turbines	67	[turbines]
Electrolyzer location	Onshore	—
Electrolyzer type	ALK	—
Electrolyzer capacity	10	[100 MW]
Distance to shore	50	[km]
Cable type	AC	—
Electrolyzer energy consumption (ALK)	51.8	[kWh/kg]
Electrolyzer energy consumption (PEM)	52.6	[kWh/kg]
Electrolyzer degradation rate (ALK)	5.04×10^{-6}	[V/h]
Electrolyzer degradation rate (PEM)	4.85×10^{-6}	[V/h]

B Surrogate modeling implementation details

We follow Maragno et al. [16] and use equivalent hyperparameter grids when training our surrogate models, these are shown in Table 10.

Table 10: Description of the hyperparameter grid used for fitting the surrogate models

Model type	Parameter grid
LR	`alpha`: [0.1, 1, 10, 100, 1000] `l1_ratio`: [0.1, 0.3, 0.5, 0.7, 0.9]
SVM	`C`: [0.1, 1, 10, 100]
CART	`max_depth`: [3, 4, 5, 6, 7, 8, 9, 10] `min_samples_leaf`: [0.02, 0.04, 0.06] `max_features`: [0.4, 0.6, 0.8, 1.0]
RF	`n_estimators`: [10, 25] `max_features`: [None] `max_depth`: [2, 3, 4]
XGB	`n_estimators`: [20] `max_depth`: [2, 3, 4] `min_child_weight`: [1,5,10] `gamma`: [0.5, 1, 2, 5, 10] `subsample`: [0.8,1] `colsample_bytree`: [0.8,1]
GBM	`n_estimators`: [20] `max_depth`: [2, 3, 4] `learning_rate`: [0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2]
MLP	`hidden_layer_sizes`: [(10,), (20,), (50,), (100,)]

C Additional results

Table 11: Estimated main effects of each simulation input on each output of interest. Statistically significant effects (where the p-value is less than 0.05) are highlighted in bold.

Simulation input	Brine [kt/y]	Curtailment [GWh/y]	Electrolyzer CF	Wind farm CF	$CVaR_{90\%}$(LCOH) [€/kg]
Number of wind turbines	**43.189**	**9.136**	**0.154**	**-0.064**	**-0.966**
Electrolyzer location	-0.227	-0.243	-0.001	-0.001	**0.376**
Electrolyzer type	2.553	**-88.646**	**0.010**	**0.025**	**-1.497**
Electrolyzer capacity	**32.378**	**33.302**	**-0.139**	**0.083**	**-0.622**
Distance to shore	-3.646	-2.268	**-0.017**	**-0.019**	**1.010**
Cable type	0.071	0.071	0.000	0.000	0.209
Electrolyzer energy consumption	**9.235**	0.520	-0.000	-0.000	**-1.501**
Electrolyzer degradation rate	-0.268	1.225	-0.002	0.001	0.047