PROCEEDINGS OF ECOS 2023 - THE 36TH INTERNATIONAL CONFERENCE ON
EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS
25-30 JUNE 2023, LAS PALMAS DE GRAN CANARIA, SPAIN

Combined Physics-Data Driven Modeling for Design and Operation Optimization of an Energy Concept

Rushit Kansara[a], Michael Lockan[b]

[a] *German Aerospace Center, Institute of Low-Carbon Industrial Processes, Simulation and Virtual Design, Walther-Pauer-Straße 5, Cottbus Germany, rushit.kansara@dlr.de*
[b] *German Aerospace Center, Institute of Low-Carbon Industrial Processes, Simulation and Virtual Design, Walther-Pauer-Straße 5, Cottbus Germany, michael.lockan@dlr.de*

Abstract:

The industrial sector accounts for a huge amount of energy- and process-related CO_2 emissions. One decarbonization strategy is to build an energy concept which provides electricity and heat for industrial processes using combination of different renewable energy sources such as photovoltaic, wind turbine, and solar thermal collector system combined with energy conversion power-to-heat components such as heat pump, electric boiler etc. The challenge for the industries is the economic aspect of the decarbonization, as industries require a cost-efficient solution. The total cost for an industrial energy concept includes investment and operating costs. This complex problem of minimizing cost and emission requires two major tasks: (I) modeling of components and (II) multi-objective coupled design and operation optimization of the energy concept. The optimal design and capacity of the components and optimal system operation depend majorly on the modeling of the components. The modeling of the components is either physics-driven or data-driven. The corresponding multi-objective coupled optimization is a complex problem with a large number of variables and constrains involved. This paper shows different types of physics- and data-driven modeling of energy components for the multi-objective coupled optimization for minimizing cost and emission of an industrial process as a case study. The optimization problem is solved as single-level problem and bi-level problem with different combinations of physics- and data-driven models. Different modeling techniques and their influence on the optimization are compared in terms of computational effort, solution accuracy and optimal capacity of components. The results show that the combination of physics and data-driven models have computational time reduction up to 37% with high accuracy compared to complete physics-driven models for the considered case study. Specific combination of physics-driven and polynomial regression models show the best trade-off between computational speed and accuracy.

Keywords:

energy concept, renewable energy sources, coupled optimization, data-driven modeling

1. Introduction

Sustainable development is one of the most pressing challenges for the industrial sector today. Industrial sector accounts for 34% of the end-energy-related CO_2 emissions in Europe [1]. A very crucial hurdle that needs to be overcome for sustainable future is the immense use of fossil fuels. The associated greenhouse gas emissions require urgent solutions to mitigate their effects on climate. Renewable energy sources (RES) such as wind and solar can be promising alternatives to fossil fuels as they are abundantly available and provide cleaner means of energy [2]. Energy transition concepts such as integrated energy systems (IES) combining RES, conversion components and fuel-based energy generation components could effectively improve the utilization of RES as well as promote the mitigation of CO_2 emissions. Energy-efficiency of such concepts plays a vital role in reducing CO_2 emissions. Efficient design and operation decisions of IES combine ecological and economic aspects; i.e., it does not only have potential for reducing CO_2 emissions, but it also supports significant cost savings as shown in various literature for industrial energy systems [3–5], district and urban energy systems [6–8] and building energy systems [9,10]. The design decisions are determined before and implemented during the development of the energy concept, such as the capacity of the energy components involved. The operation decisions are implemented after development of the energy concept, such as the physical conditions under which the system is operated [11]. However, complex energy component capacity configuration and various operation strategies make further development of such energy concepts for minimizing cost and emission difficult [12].

Multi-objective design and operation optimization is one of the most effective methods for solving such problems [13]. This optimization problem falls under coupled optimization category, where design and operation of the IES have to be optimized together to achieve the multi-objectives such as minimum CO_2 emissions and

https://doi.org/10.52202/069564-0120

minimum costs. Different literature have shown coupled optimization solved with different methods, such as bi-level and single-level optimization. Bi-level solution strategy is where an upper level problem decides the capacity of the components and a lower level decides the operation strategy based on the design decisions from upper level. Single-level solution strategy integrates design and operation optimization in a single mathematical problem. The authors in [7] showed a mixed integer linear program (MILP) of bi-level coupled optimization of district energy systems (DES) for minimizing overall cost as a single objective. In [14] a multi-objective, non-linear coupled optimization with bi-level problem formulation is presented. In [15] single-level MILP multi-objective coupled optimization for buildings is discussed. In [16] topology optimization of DES as MILP single-level formulation with scenario based operating conditions is presented. Above discussed researches have not majorly focused on industrial processes with multi-objective single-level coupled optimization. In this paper the multi-objective coupled design and operation optimization problem of an energy concept is solved using single-level as well as bi-level methods.

Modeling is an important aspect of energy system's design and operation optimization. Physics-driven modeling of energy components are majorly non-linear involving large number of variables and constraints, which makes the coupled optimization computationally expensive [16]. Data-driven models have opened new possibilities for energy system modeling [17]. Data-driven models could imitate the same physical relations hidden in the data sets without focusing on physical description of the process, which makes them quite flexible to use in an optimization problem [18]. The recent advancement in machine learning (ML) has a capability to handle high complexity of such energy system modeling arising from non-linearity of the physics [17]. In this paper data-driven models for solar thermal (ST) collector system and heat pump (HP) are used to formulate reduced order optimization problem and are compared with the physics-driven counterparts in terms of accuracy, computational efforts and optimal capacity of components. Comparison is carried out for both single-level and bi-level multi-objective optimization problem to show the optimization results' consistency.

2. Methods

The coupled design and operation optimization problem in this paper is formulated for a case study energy concept of a small- to medium-sized food and cosmetic company in Germany. Fig. 1 shows the proposed initial energy concept for the case study. It includes RES such as photovoltaic (PV), wind turbine (WT), ST; energy conversion components such as HP, gas boiler (GB), electric boiler (EB); electric grid (EG) and gas grid (GG) to suffice the consumption demand of the production. The heat is required for the steam generation, which is further used for heating and pasteurizing the products in batches. Electricity producing components and electricity consuming components are connected to an electricity hub (EH). In the same manner heat producing and consuming components are connected to a heat hub (HH). Energy storage components such as battery-storage (BAT) and thermal energy storage (TES) are not considered in this paper for the optimization, but can be integrated in future energy concepts for higher flexibility of operation. The coupled optimization

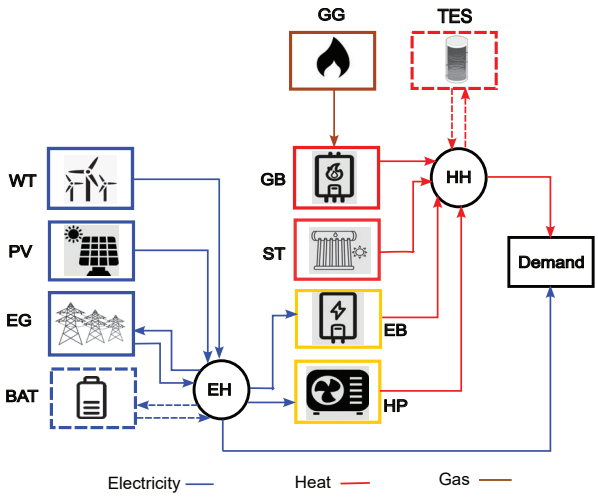
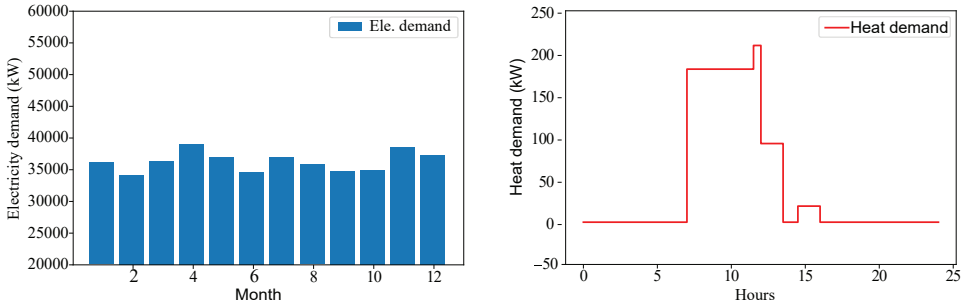


Figure 1: Concept of IES

problem of the energy concept shown in Fig. 1 is complex due to large number of continuous variables such as capacity of the components, power consumption of the components etc., and binary variables such as existence of component in the energy concept are involved with non-linear modeling of the energy components consisting large number of constraints. The total number of variables and constraints involved in this

https://doi.org/10.52202/069564-0120

optimization problem are 1741 and 2748, respectively. It falls under the general category of mixed-integer non-linear programming (MINLP) non-convex problems, which is computationally very expensive to solve [19]. Different types of physics- and data-driven models are integrated into multi-objective coupled optimization. Physics-driven modeling of the components is majorly followed as described in [20]. Solution accuracy and computational efforts of different models in optimization are compared later in section 3. The time horizon for



(a) Electricity demand (b) Heat demand

Figure 2: Monthly electricity and daily heat demand of the case study considered

the operation optimization is 1 year with an hourly time-step. The operation of the plant takes place for 5 days in a week daily for 8 hours. 21 days in a month are working days and 9 days are assumed to be off-days. The operation of the plant is assumed to be scenario-based for the optimization, where 1 working day and 1 off-day in each month are taken as the representative days. The heat demand for working day is shown in Fig. 2 (b). It can be seen that heat demand is divided into 4 operating scenarios for different operating hours on a working day for different batch processes. Off-days have base load requirement of 200 kWh of electricity. Monthly electricity demands are shown in Fig. 2 (a).

2.1. Problem formulation

This case study problem has two objectives to minimize: Total annualized cost (TAC) and global warming impact (GWI). TAC includes investment cost C and operational cost OC. The production facility of the case study is already built. The investment cost for building the facility is excluded. GWI is the measurement of the CO_2 emission. The minimization problem is formulated as

$$\min_{\mathbf{x},\mathbf{y}}[TAC(\mathbf{x},\mathbf{y}), GWI(\mathbf{x},\mathbf{y})], \tag{1}$$

with TAC and GWI as the minimization objectives. The first minimization objective TAC includes operational cost of the energy concept and investment cost of each component, which is defined as

$$TAC(\mathbf{x},\mathbf{y}) = OC(\mathbf{x},\mathbf{y}) + \sum_{i\in S} C^i(x^i), \tag{2}$$

where, the operational costs depends on the net electricity and gas bought from the grids. Which is shown as

$$OC(\mathbf{x},\mathbf{y}) = \sum_{m\in M}(p^{el}_{buy}\cdot E^{el}_{in,m} - p^{el}_{sell}\cdot E^{el}_{out,m}) + p^{gas}_{buy}\cdot E^{gas}_{in,m}), \tag{3}$$

and the investment cost C^i of each component is represented as

$$C^i = \left(\frac{(\beta+1)^\tau\cdot\beta}{(\beta+1)^\tau - 1} + \alpha\right)\cdot CAPEX, \tag{4}$$

which includes capital expenditure $CAPEX$, maintenance cost factor α, interest rate β and time horizon τ for financing cost [20]. $CAPEX$ is calculated based on reference capacity x^i of each component as

$$CAPEX = CAPEX^0\cdot\left(\frac{x^i}{x^0}\right)^\gamma, \tag{5}$$

where, γ represents the scaling exponent for the nominal capacity [20]. The second minimization objective GWI is presented as

$$GWI(\mathbf{x},\mathbf{y}) = \sum_{m\in M}(g^{el}\cdot(E^{el}_{in,m} - E^{el}_{out,m}) + g^{gas}\cdot E^{gas}_{in,m}), \tag{6}$$

where, $M = \{Jan, Feb...., Dec\}$ and $S = \{PV, WT, SH, GB, EB, HP\}$. $\mathbf{x} = [A^{PV}, P^{WT}_{nom}, Q^{GB}_{nom}, Q^{EB}_{nom}, Q^{HP}_{nom}]$ are the design variables and $\mathbf{y} = [E^{el}_{in,m}, E^{el}_{out,m}, E^{gas}_{in,m}]$ are the operational variables for the objective function. GWI is calculated based on the net buying of electricity and gas from the grid. Each net consumed unit of electricity $E^{el}_{in,m} - E^{el}_{out,m}$ and gas $E^{gas}_{in,m}$ has been given corresponding CO_2 factors g^{el} and g^{gas}, respectively, for the calculation of GWI. Parameters for the OC and GWI are shown in Table 1. Table 2 shows the parameters for calculating investment cost of each component and it also shows the minimum part load requirement λ_{min} for the components.

Table 1: Cost and emission parameters for grids taken from [20, 21]

Name	Parameter	Value
electricity buying price	p^{el}_{buy}	0.31 [€]
electricity selling price	p^{el}_{sell}	0.06 [€]
gas buying price	p^{gas}_{buy}	0.15 [€]
CO_2 factor for net consumed electricity	g^{el}	0.349 $[kg - CO_2eq/kWh]$
CO_2 factor for consumed gas	g^{gas}	0.244 $[kg - CO_2eq/kWh]$

The heat demand constraint is shown as

$$Q^{dem} - (Q^{ST} + Q^{GB} + Q^{EB} + Q^{HP}) \leq 0, \tag{7}$$

which indicates that the heat generated from ST, GB, EB and HP should fulfill the heat demand of the production in each time step. Constraint on the capacity of the components is shown as

$$x^i_{min} \cdot z^i \leq x^i \leq x^i_{max} \cdot z^i \qquad \forall i \in \mathbf{x} \quad and \quad z^i \in \{0, 1\}, \tag{8}$$

where, x^i_{min} and x^i_{max} are the lower and upper bounds of the capacity of each components. z is the binary variable, which is linked to the existence of the component in the concept. Equations (1)-(8) represent single-level multi-objective coupled design and operation optimization problem.

Table 2: Component parameters for investment costs and part-load constraints [20]

Components	reference capacity x^0	$CAPEX^0$[€]	γ	α	β	τ[a]	λ_{min}
PV	A^0 [m^2]	1400	0.95	0.01	0.03	10	0
WT	P^0_{nom} [kW]	5000	0.95	0.03	0.03	10	0.33
ST	A^0 [m^2]	400	0.95	0.02	0.03	10	0
GB	Q^0_{nom} [kWh]	2700	0.45	0.02	0.03	10	0.2
EB	Q^0_{nom} [kWh]	70	0.95	0.01	0.03	10	0
HP	Q^0_{nom} [kWh]	1655	0.66	0.02	0.03	10	0

2.2. Modeling of components

Modeling of the components is a crucial part of the optimization problem. In this subsection, for each component integrated in the energy concept either a physics-driven and/or data-driven models are explained. List of all variables, constant parameters and input parameters for physics-driven component modeling is presented in Table 3.

2.2.1. Photovoltaic

The electrical power P^{PV}_{out} generated by the PV unit is constrained by the solar irradiance I, the efficiency of PV unit η^{PV} and total area A^{PV}_{nom} of the PV system. It is represented by

$$P^{PV}_{out} \leq A^{PV}_{nom} \cdot \eta^{PV} \cdot I. \tag{9}$$

The case study facility already has some PV panels built on the terrace, with a tilt angle 5° for the PV model. Moreover, PV unit has the maximum output power limited to its nominal capacity, which is presented as

$$P^{PV}_{out} \leq A^{PV}_{nom} \cdot P^{PV}_{nom}, \tag{10}$$

https://doi.org/10.52202/069564-0120

Table 3: Component variables including binary variables, parameters and inputs for single-level problem

Components	Design variables **x**	Operational variables **y**	Constant parameters **c**	Input paramters in each timestep	Total number of variables
PV	A_{nom}^{PV}	P_{out}^{PV}	η^{PV},P_{nom}^{PV}	I	3
WT	P_{nom}^{WT}	λ^{WT},P_{out}^{WT}	-	v	4
GB	Q_{nom}^{GB}	λ^{GB}, E_{in}^{gas}	η^0	-	4
EB	P_{nom}^{EB}	λ^{EB}, P_{in}^{EB}	η^{EB}	-	4
ST	A_{nom}^{ST}	T_{out}^{w}, T_{out}, T_{in}	I_{AM}, η_0, $\dot{m}^{ST}$	I, T_{amb}, $\dot{m}^{w}$	5
HP	Q_{nom}^{HP}	λ^{HP}, h_{out}^{w}, T_{out}^{w}, P_{in}^{HP}	a, b, c, d, T_{in}^{c}, T_{in}^{w}	$\dot{m}^{w}$	6
Electric grid	-	E_{in}^{el}, E_{out}^{el}	-	-	2
Gas grid	-	E_{in}^{gas}	-	-	1

where, P_{nom}^{PV} is chosen as 0.171 kWm^{-2} [20] and efficiency η^{PV} is chosen to be 0.09 in order to meet the actual output data of PV panels built on the case study facility. P_{out}^{PV} is an operational variable and A_{nom}^{PV} of the unit is a design variable. Total number of variables in Table 3 shows number of design, operational and binary variables to be computed in each time step in optimization problem.

2.2.2. Wind turbine

The power output P_{out}^{WT} of WT is limited by the wind velocity, which in turn determines the part-load behavior of WT and its nominal power [22]. The output power of WT is given by

$$P_{out}^{WT} \leq \eta^{WT}(\lambda^{WT}) \cdot P_{nom}^{WT}, \tag{11}$$

where, P_{out}^{WT} is an operational variable and P_{nom}^{WT} of the wind turbine is a design variable. Operational variable λ^{WT} depends on the wind velocity ($\lambda^{WT} = v/v_{ref}$, where $v_{ref} = 12m/s$ [20]). The efficiency of wind turbine $\eta^{WT}(\lambda^{WT})$ is given by

$$\eta^{WT}(\lambda^{WT}) = \begin{cases} 0 & \text{if } \lambda^{WT} \leq 0.33 \\ 1.5393 \cdot \lambda^{WT} - 0.5091 & 0.33 \leq \lambda^{WT} \leq 1 \\ 1 & \lambda^{WT} \geq 1. \end{cases} \tag{12}$$

2.2.3. Gas boiler

Heat output of the GB is determined by the part-load efficiency, which is given as

$$\eta^{GB}(\lambda^{GB}) = \frac{21.75378 \cdot \lambda^3 - 7.00130 \cdot \lambda^2 + 1.39731 \cdot \lambda - 0.07557}{20.66646 \cdot \lambda^3 - 5.34196 \cdot \lambda^2 + 0.67774 \cdot \lambda + 0.03487} \cdot \eta^0, \tag{13}$$

where η^0 is chosen to be 0.8 [1], which is called nominal efficiency. The heat output depends on the efficiency shown in (13), consumed gas power to heat up the incoming fluid P_{in}^{gas} and nominal capacity Q_{nom}^{GB}, which is presented as

$$Q_{out}^{GB} = \eta^{GB}(\lambda^{GB}) \cdot P_{in}^{gas}, \qquad Q_{out}^{GB} = \lambda^{GB} \cdot Q_{nom}^{GB}. \tag{14}$$

Here, heat output Q_{out}^{GB}, consumed power P_{in}^{gas} and the part-load λ^{GB} are operational variables and nominal capacity Q_{nom}^{GB} is the design variable. The integration of consumed power P_{in}^{gas} over the operation horizon delivers the total energy E_{in}^{gas} consumed from gas grid.

2.2.4. Electric boiler

EB is modeled in the same manner as GB. Efficiency η^{EB} for EB is assumed to be constant at 0.95. The operational variable heat output of the EB depends on operational variables such as consumed electric power P_{in}^{EB}, part-load λ^{EB} and design variable nominal capacity Q_{nom}^{EB}. It is shown as

$$Q_{out}^{EB} = \eta^{EB} \cdot P_{in}^{EB}, \qquad Q_{out}^{EB} = \lambda^{EB} \cdot Q_{nom}^{EB}. \tag{15}$$

2.2.5. Solar thermal collector

There are two kinds of ST collectors generally used in the market, flat plate and evacuated tube. For this case study, flat plate collectors with tilt angle of 40° are chosen to be integrated in the energy concept [23]. ST model is based on the quadratic efficiency model developed by [24], which is shown as

$$\eta^{ST} = \eta_0 - \frac{a_1 \cdot \Delta T}{I} - \frac{a_2 \cdot \Delta T^2}{I}, \tag{16}$$

where η_0 represents optical collector efficiency, a_1 and a_2 are loss coefficients related to linear and quadratic terms, ΔT represents temperature difference between collector fluid temperature and ambient temperature, and I as mentioned before is global solar irradiance on the collector surface. European EN 19275 standards prescribe the collector fluid temperature as average collector temperature of its inlet and outlet temperature [25].

Complete hourly weather data for the location of the case study plant has been gathered from European commission photovoltaic geographical information system [26]. The global solar irradiance and ambient temperatures are the important data for the ST collector model. Complete physical ST collector model is shown in (17) - (20). The ST collector efficiency is shown as

$$\eta^{ST} = \eta_0 \cdot I_{AM} - \frac{a_1 \cdot (T_m - T_{amb})}{I} - \frac{a_2 \cdot (T_m - T_{amb})^2}{I}, \tag{17}$$

where, I_{AM} represents incidence angle modifier, which corrects the optical efficiency for the irradiation not perpendicular to the surfaces [23]. T_m is the average collector fluid temperature, which depends on inlet and outlet temperature T_{in}^{ST} and T_{out}^{ST}, respectively. The optical collector gain is represented by

$$Q_0 = \eta_0 \cdot I_{AM} \cdot I \cdot A_{nom}^{ST}, \tag{18}$$

and thermal losses due to temperature difference between average fluid temperature and ambient temperature is modeled as

$$Q_L = \left(a_1 \cdot (T_m - T_{amb}) - a_2 \cdot (T_m - T_{amb})^2\right) \cdot A_{nom}^{ST}, \tag{19}$$

where, collector surface area is A_{nom}^{ST} of the collector surface is the design variable.

$$Q_u = \begin{cases} Q_0 - Q_L, & \text{if } Q_0 > Q_L \\ 0 & \text{if } Q_0 \leq Q_L \end{cases} \tag{20}$$

Equation (20) shows the useful solar gains with the condition of positive net solar gains. The authors in [23] compared this simplified physical model with detailed simulation model in TRNSYS. The parameters for the physical models are collected in [27], in which extensive research has been carried out for different ST flat collector manufacturers data in Germany and related parameters. Parameters chosen for the case study are shown in Table 4. Fig. 3(a) shows the ST collector with the heat-exchanger (HEX) used in the energy concept. The heat transfer over HEX is given by

$$Q^{ST} = \dot{m}^w \cdot c_p^w \cdot (T_{out}^w - T_{in}^w). \tag{21}$$

For calculating the operational variable heat transfer Q^{ST} over HEX in each time step, 5 different equations (17)-(21) must be solved in the physical model, which consist 3 additional operational variables such as collector inlet temperature T_{in}^{ST}, collector outlet temperature T_{out}^{ST}, water/steam outlet temperature T_{out}^w and one design variable, area A_{nom}^{ST} of the collector surface. Large number of variables and constraints make the optimization expensive. In order to reduce the number of variables and constraints, data-models are trained and used in the optimization.

Data-driven approach to model ST collector could have some advantages over physical model. Proposed data-driven approach is presented as

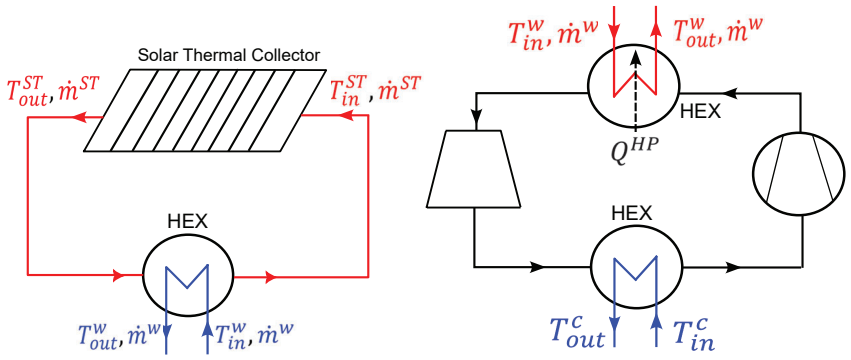
$$Q^{ST} = f(T_{in}^w, \dot{m}^w, I, A_{nom}^{ST}), \tag{22}$$

where, operational variable heat output of the whole ST collector system depends on 4 inputs. These inputs include only one design variable A_{nom}^{ST}, 3 input parameters T_{in}^w, $\dot{m}^w$, I and none of the operational variables shown before. f in (22) represents a generic data-driven model. This data-driven model represents the relationship between the given input parameters including the design variable area A_{nom}^{ST} with the output variable heat output Q^{ST}. In this paper data-driven models such as linear regression (LR), polynomial regression (PR) and artificial neural network (ANN) are considered. These data-driven models are trained with the input and output data generated from the physical model. In this manner, operational variables to be optimized are reduced to 1 in each time step and all other constraints are eliminated.

https://doi.org/10.52202/069564-0120

Table 4: Parameters for ST flat collector

Collector	η_0	a_1 $[W/(m^2K)]$	a_2 $[W/(m^2K)]$	IAM
Flat plate	0.79	4.03	0.0078	0.86



(a) Solar flat plate collector with HEX

(b) HP consisting a compressor, 2 HEX and a turbine

Figure 3: ST collector and HP schematic diagram

2.2.6. Heat pump

HP can use the industrial waste heat as a thermal energy source as well as renewable electricity as the power input to decarbonize industrial thermal processes, which makes it an essential technology [28]. In this case study, waste heat is considered as constant. Renewable electricity is generated from PV and WT. There are many different physical models to describe the performance of the HP. The authors in [28] have investigated different case studies of different types of HPs and came up with different coefficient of performance (COP) regression models. Fig. 3(b) shows a heat pump consisting of two HEXs (condenser and evaporator), a compressor and a turbine. T^c_{in} and T^c_{out} represent input and output temperatures on cold side, whereas T^w_{in} and T^w_{out} are the input and output temperatures on the hot side, respectively. The heat output of the HP depends on its nominal capacity Q^{HP}_{nom} and part-load λ^{HP}, which is shown as

$$Q^{HP}_{out} = \lambda^{HP} \cdot Q^{HP}_{nom}. \tag{23}$$

The output temperature T^w_{out} on the hot side depends on the outlet pressure of water/steam p^w_{out} on hot side and enthalpy h^w_{out} of water/steam, which is a function of heat output Q^{HP}_{out} and inlet temperature T^w_{in} of water on the hot side, shown as

$$T^w_{out} = f(h^w_{out}(Q^{HP}_{out}, T^w_{in}), p^w_{out}). \tag{24}$$

The COP is calculated based on the model

$$COP = a \cdot (\Delta T_{lift} + 2b)^c \cdot (T^w_{out} + b)^d, \tag{25}$$

suggested by [28]. The coefficients a,b,c and d of this model are chosen according to the suitable temperature ranges of the case study ($80\,°C < T^w_{out} < 160\,°C$). The heat output is a function of COP and the consumed electrical power P^{HP}_{in}, as well as the input and output enthalpy h^w_{in} and h^w_{out} of water/steam on hot side, which is shown as,

$$Q^{HP}_{out} = COP \cdot P^{HP}_{in}, \qquad Q^{HP}_{out} = \dot{m}^w \cdot (h^w_{out} - h^w_{in}). \tag{26}$$

Equation (23)-(26) represent the physics-driven model of HP. The detailed list of design and operational variables of HP are shown in Table 3. As seen from the physical model, 5 different equations need to be solved in each time step, which include 5 operational variables Q^{HP}_{out}, COP, P^{HP}_{in}, λ^{HP}, T^w_{out} and one design variable Q^{HP}_{nom}. Data-driven approach for HP is shown as

$$COP = f(\dot{m}^w, T^c_{in}, \lambda^{HP}, Q^{HP}_{nom}), \tag{27}$$

which includes 1 design variable Q^{HP}_{nom}, 1 operational variable λ^{HP} and 2 input parameters $\dot{m}^w$ and T^c_{in} as inputs to compute the output variable COP without any constraints. Table 5 shows the total number of variables and

constraints over the whole time-horizon for the single-level multi-objective coupled optimization. It is evident from Table 5 that the total number of variables and constraints involved in coupled optimization problem decreases with the use of data-driven models compared to the complete physics-driven models of ST and HP. Data-driven models for ST and HP reduce the number of variables by 388 and constraints by 504, which is 22% and 18% less than the complete physics-driven models.

Table 5: Total number of variables and constraints in single-level multi-objective coupled optimization problem for different combinations of physics-driven and data-driven models of HP and ST

ST	HP	Total number of variables	Total number of constraints
Physics-driven	Physics-driven	1741	2748
Data-driven	Data-driven	1353	2244
Data-driven	Physics-driven	1525	2316
Physics-driven	Data-driven	1669	2676

3. Results

3.1. Comparison of data-driven models

As shown in Table 5, the data-driven models have fewer variables and constraints to compute in coupled optimization. Table 6 shows the inputs, outputs, number of parameters and amount of data samples. This data samples are divided into training and validation data. 75% of the data samples is used for training the models and 25% of the data samples is used for validating the trained models. Table 7 shows the training time and comparison of accuracy between physical models (actual values) and different data-driven models (predicted values) based on coefficient of determination R^2 method [29]. Three types of data-driven models: LR, PR

Table 6: Data-driven models' input, output and number of samples

Component	Inputs	Output	Number of data samples
ST	T_{in}^{w}, $\dot{m}^{w}$, I, A	Q^{ST}	439199
HP	$\dot{m}^{w}$, T_{in}^{c}, λ^{HP}, Q_{nom}^{HP}	COP	206054

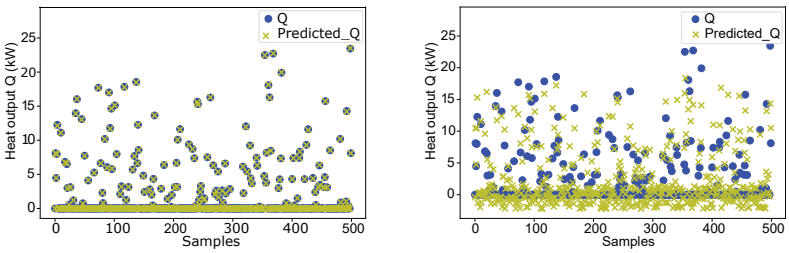
and ANN are trained for ST and HP. In particular, PR models have two variant such as 2^{nd} degree PR (PR-1) and 3^{rd} degree PR (PR-2) models. Furthermore, ANNs are feed forward neural networks with two different specifications: (I) 2 hidden layers, 5 neurons in each hidden layer (ANN-1) (II) 3 hidden layers, 7 neurons in each hidden layer (ANN-2). The number of hidden layers and number of neurons in each hidden layer are optimized using hyper-parameter tuning technique. ANNs are trained with k-fold cross validation method with k=4 [30].

It can be seen from Table 7 that the R^2 score of ST is less for ANN-2 compared to ANN-1. See also Fig. 4 (b), where for ANN-2 the predicted output data does not match properly with actual output data. In contrast, Fig. 5 shows results for the two different ANN models of HP, where ANN-2 showing better fitting performance

Table 7: Data-driven models' training time and accuracy

Model	Specification	Training time for ST [s]	Training time for HP [s]	R^2 for ST	R^2 for HP
LR	degree 1	2.12	1.82	0.96	0.45
PR-1	degree 2	5.33	4.31	0.972	0.68
PR-2	degree 3	11.26	9.25	0.986	0.825
ANN-1	2 hidden layers 5 neurons each	1254	1008	0.999	0.862
ANN-2	3 hidden layers 7 neurons each	1852	1369	0.845	0.982

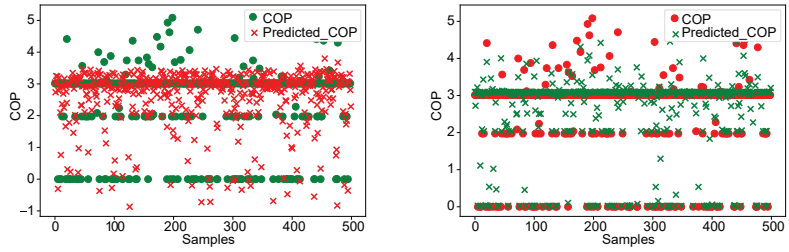
https://doi.org/10.52202/069564-0120



(a) ANN-1 (2 layers, 5 neurons each) (b) ANN-2 (3 layers, 7 neurons each)

Figure 4: Comparison of ANN models output prediction to actual output of ST

compared to ANN-1. Therefore, the optimized ANN for ST is ANN-1 and for HP is ANN-2. As expected, LR and PR models for HP have lower values of R^2 score compared to ANN models. On the other hand ST has quite higher R^2 score for LR and PR models. Training time for LR and PR models are lesser than their respective ANN models. Training time for ANN is higher due to large number of weights and biases involved in the ANN model. ANN-1 model and ANN-2 model have 35 and 112 such coefficients (and constants), respectively, which makes them time-expensive to compute when these ANN models are integrated into the coupled optimization.



(a) ANN-1 (2 layers, 5 neurons each) (b) ANN-2 (3 layers, 7 neurons each)

Figure 5: Comparison of ANN models output prediction to actual output of HP

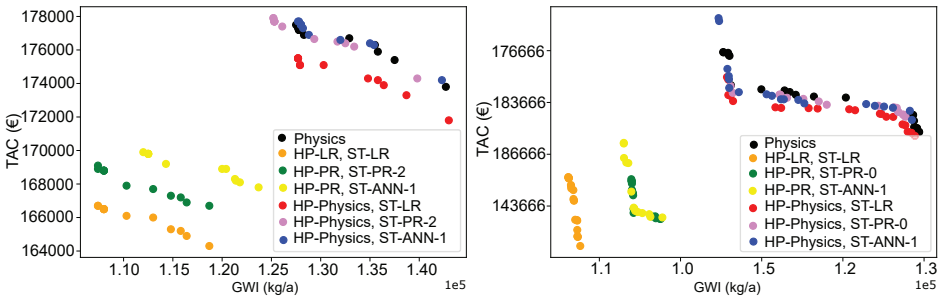
3.2. Comparison of optimization results

Fig. 6 shows several Pareto-fronts of TAC and GWI as results of multi-objective coupled design and operation optimization of the case study energy concept. These pareto-fronts are formed by various combinations of physics- and data-driven models. In particular, Fig. 6 (a) shows the single-level optimization results and (b) shows bi-level optimization results for different combinations of ST and HP models. The optimal Pareto-front of complete physics-driven model (black dots) of the IES is used as the reference solution to evaluate the accuracy and computational effort of the different model combinations.

The original multi-objective MINLP optimization problem is linearized and converted into a MILP optimization problem. The MILP problem is solved with GUROBI solver [31] on PYOMO platform included in COMANDO [22]. More specifically, the single level multi-objective pareto-front is generated by the augmented ϵ-constraint method [32]. Bi-level problem is not linearized and solved on PYMOO platform with non-sorting genetic algorithm (NSGA-II) on the design level and differential evolution on the operation level [33]. Both optimization problems, single-level and bi-level problem, are solved on 11th Gen Intel(R) Core(TM) i7-1185G7 with 16 GB RAM. It can be seen from Fig. 6 (a) and Fig. 6 (b) that the physics-driven model of HP with ANN-1 data-driven model of ST (blue dots) gives the most accurate results, which is close to the complete physics-driven model. The second most accurate result is provided by physics-driven model of HP and PR-2 data-driven model of ST (violet dots). This trend applies to both, single-level and bi-level optimization.

Fig. 7 shows the comparison of computational time and solution accuracy of the coupled optimization results for combination of different models. Accuracy is calculated based on R^2 method, where complete physics-driven model is reference. It can be seen that the best trade-off between computational time and accuracy is found for HP physics-driven model and ST PR-2 model (violet dots). This combination of models reduces the coupled optimization computation time up to 37% with optimization results being approximately 90% accurate

https://doi.org/10.52202/069564-0120



(a) Single-level coupled optimization

(b) Bi-level coupled optimization

Figure 6: Comparison of the optimization results for different model combinations of ST and HP

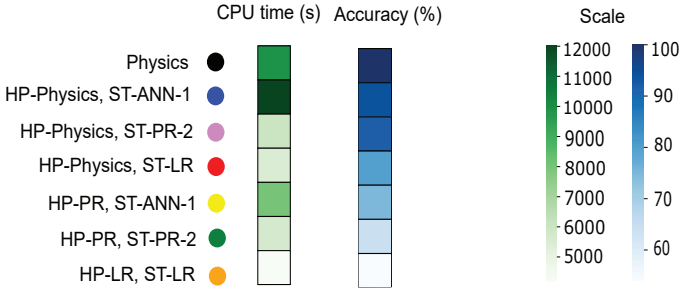
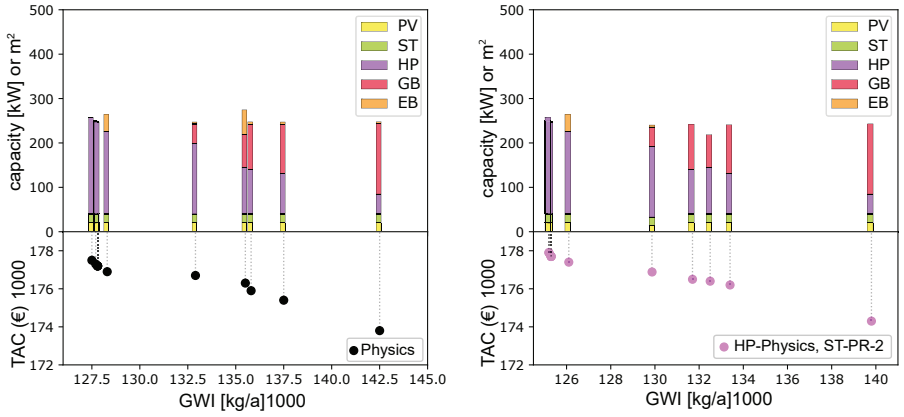


Figure 7: Time and accuracy comparison for different combination of models in single-level multi-objective optimization

compared to complete physics-driven models. In contrast, ST ANN-1 model with HP physics-driven (blue dots) model in coupled optimization is also very accurate, but the computational time for solving the coupled optimization problems with ANN-models are very high due to large number of weights and biases.



(a) Complete physics-driven model

(b) HP-Physics, ST-PR-2 model

Figure 8: Capacity of components for reference solution and best trade-off solution for single-level coupled optimization

Fig. 8 shows the optimal capacity of the components for the total physics-driven models and the best trade-off combination of physics- and data-driven models, which is chosen based on best trade-off between computational time and solution accuracy according to Fig. 7. Both Fig. 8 (a) and 8 (b) show that for minimum TAC small HP and large GB is required to meet the heating demand, while for minimum GWI large HP is required. The size of ST and PV remains almost constant in all of the Pareto-results. In some solutions, EB is needed

https://doi.org/10.52202/069564-0120

when HP and GB are not sufficient to provide the required heat. The optimal capacity of the components follow a similar trend for complete physics-driven and combined physics-data-driven models, demonstrating the credibility of the combined models.

4. Conclusion and outlook

This paper showed a multi-objective coupled design and operation optimization of an energy concept of a food and cosmetic industry as a case study. Single-level optimization for minimizing TAC and GWI as well as design and operation variables for each component involved is described. The aim of this paper is to compare different types of physics- and data-driven models and to integrate them into a coupled optimization problem to reduce the computational time while maintaining the accuracy of the optimization results. The results showed that the combination of data-driven PR model of ST and physics-driven HP model better than all other combinations in terms of computational time and solution accuracy. The accuracy of the optimization result is up to 90% compared to complete physics-driven model and the computational time is reduced by 37%.

Future work for this case study is to include energy storage models into the energy concept. Integration of thermal and electrical storage can increase the flexibility. Different scenarios such as retrofit designs, greenfield designs, cost neutral solutions and emission free solutions can be optimized for the given case-study and different design-operation solutions can be provided. Combination of physics-data models can be very useful in optimizing these scenarios at less computational efforts and high accuracy.

References

[1] P. Voll, C. Klaffke, M. Hennen, A. Bardow. *Automated superstructure-based synthesis and optimization of distributed energy supply systems*. Energy 2013; 50:374–388; doi:10.1016/j.energy.2012.10.045.

[2] M. J. Mayer, A. Szilagyi, G. Grof. *Ecodesign of ground-mounted photovoltaic power plants: Economic and environmental multi-objective optimization*. Journal of Cleaner Production 2021; 278: 123934; doi:10.1016/j.jclepro.2020.123934.

[3] O. Aguilar, S. J. Perry, J.-K. Kim, R. Smith. *Design and optimization of flexible utility systems subject to variable conditions*. Chemical Engineering Research and Design 2007; 85 (8):1149-1168; doi:10.1205/cherd06063.

[4] P. Velasco-Garcia, P. S. Varbanov, H. Arellano-Garcia, G. Wozny. *Utility systems operation: Optimisation based decision making*. Applied Thermal Engineering 2011; 31 (16):3196-3205; doi:10.1016/j.applthermaleng.2011.05.046.

[5] G. M. Tina, G. Passarello. *Short-term scheduling of industrial co-generation systems for annual revenue maximisation*. Energy 2012; 42(1):46-56; doi:10.1016/j.energy.2011.10.025.

[6] C. Bouvy, K. Lucas. *Multicriterial optimisation of communal energy supply concepts*. Energy Conversion and Management 2007; 48 (11):2827-2835; doi:10.1016/j.enconman.2007.06.046.

[7] C. Weber, N. Shah. *Optimisation based design of a district energy system for an eco-town in the united kingdom*. Energy 2011; 36 (2):1292-1308; doi:10.1016/j.energy.2010.11.014.

[8] J. Keirstead, N. Shah. *Calculating minimum energy urban layouts with mathematical programming and monte carlo analysis techniques*. Computers, Environment and Urban Systems 2011; 35 (5):368-377; doi:10.1016/j.compenvurbsys.2010.12.005.

[9] M. A. Lozano, J. C. Ramos, M. Carvalho, L. M. Serra. *Structure optimization of energy supply systems intertiary sector buildings*. Energy and Buildings 2009; 41 (10):1063-1075; doi:10.1016/j.enbuild.2009.05.008.

[10] P. Liu, E. N. Pistikopoulos, Z. Li. *An energy systems engineering approach to the optimal design of energy systems in commercial buildings*.Energy Policy 2010; 38 (8):4224-4231; doi:10.1016/j.enpol.2010.03.051.

[11] P. Barton, X. Li. *Optimal design and operation of energy systems under uncertainty*. International Symposium on Dynamics and Control of Process Systems 2013; Mumbai, India.

[12] J. Li, H. Zhao. *Multi-objective optimization and performance assessments of an integrated energy system based on fuel, wind and solar energies*. Entropy 2021; 23 (4); doi:10.3390/e23040431.

[13] R. S. Patwal, N. Narang. *Multi-objective generation scheduling of integrated energy system using fuzzy based surrogate worth trade-off approach*.Renewable Energy 2020; 156:864-882; doi:10.1016/j.renene.2020.04.058.

[14] S. Fazlollahi, G. Becker, A. Ashouri, F. Marechal. *Multi-objective, multi-period optimization of district energy systems: Iv – a case study*. Energy 2015; 84:365-381; doi:10.1016/j.energy.2015.03.003.

[15] B. Morvaj, R. Evins, J. Carmeliet. *Optimising urban energy systems: Simultaneous system sizing, operationand district heating network layout*. Energy 2016; 116:619-636; doi:10.1016/j.energy.2016.09.139.

[16] J. Söderman, F. Pettersson. *Structural and operational optimisation of distributed energy systems*. Applied Thermal Engineering 2006; 26 (13) : 1400-1408; doi:10.1016/j.applthermaleng.2005.05.034.

[17] C. Zhang. *Data driven modeling and optimization of energy systems [dissertation]*. Singapore, Nanyang Technological University; 2019.

[18] C. Zhang, L. Cao, A. Romagnoli. *On the feature engineering of building energy data mining*. Sustainable Cities and Society 2018; 39:508-518; doi:10.1016/j.scs.2018.02.016.

[19] G. G. Dimopoulos, A. V. Kougioufas, C. A. Frangopoulos. *Synthesis, design and operation optimization of a marine energy system*. Energy 2008; 33 (2):180-188; doi:10.1016/j.energy.2007.09.004.

[20] S. Sass, T. Faulwasser, D. E. Hollermann, C. D. Kappatou, D. Sauer, T. Schutz, D. Y. Shu, A. Bardow, L. Groll, V. Hagenmeyer, D. Müller, A. Mitsos. *Model compendium, data, and optimization benchmarks for sector-coupled energy systems*. Computers and Chemical Engineering 2020; 135 : 106760; doi:10.1016/j.compchemeng.2020.106760.

[21] Umweltbundesamt *Kohlendioxid-Emissionen 2020*. Available at `https://www.umweltbundesamt.de/daten/klima/treibhausgas-emissionen-in-deutschland/kohlendioxid-emissionen#herkunft-und-minderung-von-kohlendioxid-emissionen`[accessed 07.05.2023]

[22] M. Langiu, D. Y. Shu, F. J. Baader, D. Hering, U. Bau, A. Xhonneux, D. Müller, A. Bardow, A. Mitsos, M. Dahmen. *Comando: A next-generation open-source framework for energy systems optimization*. Computers and Chemical Engineering 2021, 152:107366; doi:10.1016/j.compchemeng.2021.107366.

[23] M. Grahovac, P. Liedl, J. Frisch, P. Tzscheutschler. *Simplified Solar Collector Model: Hourly Simulation of Solar Boundary Condition for Multi-Energy Optimization*. International Congress on Heating, Refrigerating and Air-Conditioning 2011; Belgrade, Serbia.

[24] Duffie, J.,Beckman, W.A.. *Solar engineering of thermal processes*. John Wiley and Sons Inc. 1991.

[25] DIN EN 12975 *Thermische solaranlagen und ihre bauteile – kollektoren – teil 2 : prüfverfahren.*

[26] EU-Science-Hub *Photovoltaic geographical information system 2020*. Available at `https://joint-research-centre.ec.europa.eu/pvgis-online-tool/pvgis-tools/hourly-radiation_en`[accessed 20.01.2023]

[27] SPF Research. *Research and development for sustainable energy systems: Flat plate collectors*. Available at: `https://www.ost.ch/` [accessed 02.02.23]

[28] F. Schlosser, M. Jesper, J. Vogelsang, T. G. Walmsley, C. Arpagaus, J. Hesselbach. *Large-scale heat pumps: Applications, performance, economic feasibility and industrial integration*. Renewable and Sustainable Energy Reviews 2020; 133:110219; doi:10.1016/j.rser.2020.110219.

[29] *Model Selection using R-squared (R^2) Measure*. Available at `https://www.ml-concepts.com/2022/01/19/r-squared-r%C2%B2-measure-for-model-selection/`; [accessed 07.03.23]

[30] F.Y.H. Ahmed, Y. H. Ali, S. M. Shamsuddin. *Using K-Fold Cross Validation Proposed Models for Spikeprop Learning Enhancements*. International Journal of Engineering and Technology 2018; 7 (4.11):145-151.

[31] Non-convex quadratic optimization. *available at `https://www.gurobi.com/events/non-convex-quadratic-optimization/`*; [accessed 27.02.23]

[32] G. Mavrotas. *Effective implementation of the e-constraint method in Multi-Objective Mathematical Programming problems*. Applied Mathematics and Computation 2009; 213:455-465.

[33] A. Slowik, H. Kwasnicka. *Evolutionary algorithms and their applications to engineering problems*. Neural Computing and Applications 2020; 2:12363–12379; doi:10.1007/s00521-020-04832-8.

 https://doi.org/10.52202/069564-0120