Article

Combined Physics- and Data-Driven Modeling for the Design and Operation Optimization of an Energy Concept Including a Storage System †

Rushit Kansara * , Michael Lockan and María Isabel Roldán Serrano

German Aerospace Center, Institute of Low-Carbon Industrial Processes, 03046 Cottbus, Germany;
michael.lockan@dlr.de (M.L.); maria.roldanserrano@dlr.de (M.I.R.S.)
* Correspondence: rushit.kansara@dlr.de
† This paper is an extended version of our paper published in ECOS 2023.

Abstract: The industrial sector accounts for a huge amount of energy- and process-related CO_2 emissions. One decarbonization measure is to build an energy concept that provides electricity and heat for industrial processes using a combination of different renewable energy sources, such as photovoltaic, wind turbine, and solar thermal collector systems, integrating also energy conversion power-to-heat components such as heat pumps, electric boilers, and thermal energy storage. The challenge for the industries is the economic aspect of the decarbonization, as industries require a cost-efficient solution. Minimizing cost and emissions together is a complex problem, which 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 component modeling, which is either physics-driven or data-driven. This paper shows different types of physics- and data-driven modeling of energy components for multi-objective coupled optimization in order to minimize costs and emissions of a specific industrial process as a case study. Several 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 has a computational time reduction of up to 37% for an energy concept without thermal energy storage and 29% for that with thermal energy storage, both with high-accuracy solutions compared to complete physics-driven models for the considered case study.

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. The 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 a 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. The 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 the potential for reducing CO_2 emissions,

but it also supports significant cost savings, as shown in various studies of 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 configurations 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 the coupled optimization category, where the design and operation of the IES have to be optimized together in order to minimize CO_2 emissions and costs. The authors of [14–16] showed coupled optimization solved with different methods, such as bi-level and single-level optimization. A bi-level solution strategy is where an upper-level problem decides the capacity of the components and a lower-level problem decides the operation strategy based on the design decisions from the upper level. A 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, nonlinear coupled optimization with bi-level problem formulation is presented. In [15], single-level MILP multi-objective coupled optimization for buildings is discussed. In [16], the topology optimization of DES as MILP single-level formulation with scenario-based operating conditions is presented. The above discussed studies have not majorly focused on industrial processes with multi-objective single-level coupled optimization. However, in this paper, the multi-objective coupled design and operation optimization problem of an energy concept with and without a storage system is solved using single-level as well as bi-level methods.

Modeling is an important aspect of an energy system's design and operation optimization. Physics-driven modeling of energy components is majorly nonlinear, involving a 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 the 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 the high complexity of such energy system modeling, arising from the nonlinearity of the physics [17]. The study in [19] modeled Greek long-term energy consumption using linear regression and support vector machines (SVM). The authors in [20] showed the use of artificial neural networks (ANN) and SVMs in building energy consumption modeling. Studies in the field of using neural networks for energy system modeling and control also include optimization of the thermal comfort control of HVAC systems and absorption chiller systems using ANN methods [21]. According to [22], data-driven modeling approaches to resource and energy efficiency are expected to have an 18% reduction in energy consumption. McKinsey & Company [23] estimates that data-driven approaches would improve productivity in the technical profession by 45 to 55%, which results eventually in energy savings. In this paper, data-driven models for solar thermal (ST) collector systems such as flat plate collectors and heat pumps (HP) are used to formulate a reduced-order optimization problem and are compared with the physics-driven counterparts in terms of accuracy, computational efforts and optimal capacity of components. ST collects heat by absorbing sunlight. ST converts solar energy into useful thermal energy. HP converts electricity into heat with high efficiency compared to conventional fossil-based technologies. A comparison is carried out for both single-level and bi-level multi-objective optimization problems to show the optimization results' consistency and to find the methodology that offers a lower time-consuming calculation with accurate results.

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 industry, Sanddorn GmbH in Brandenburg state of Germany, as a partner in European Union's Horizon Europe Project SINNOGENES. Figure 1 shows the proposed initial energy concept for the case study. It includes RESs such as photovoltaic (PV), wind turbine (WT) and ST RESs; energy conversion components such as HPS, gas boilers (GB) and electric boilers (EB); storage components such as thermal energy storage (TES); and electric grids (EG) and gas grids (GG) to suffice the consumption demand of the production. 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). Battery storage (BAT) is not considered in this paper for optimization, but can be integrated in future energy concepts for higher flexibility of operation.
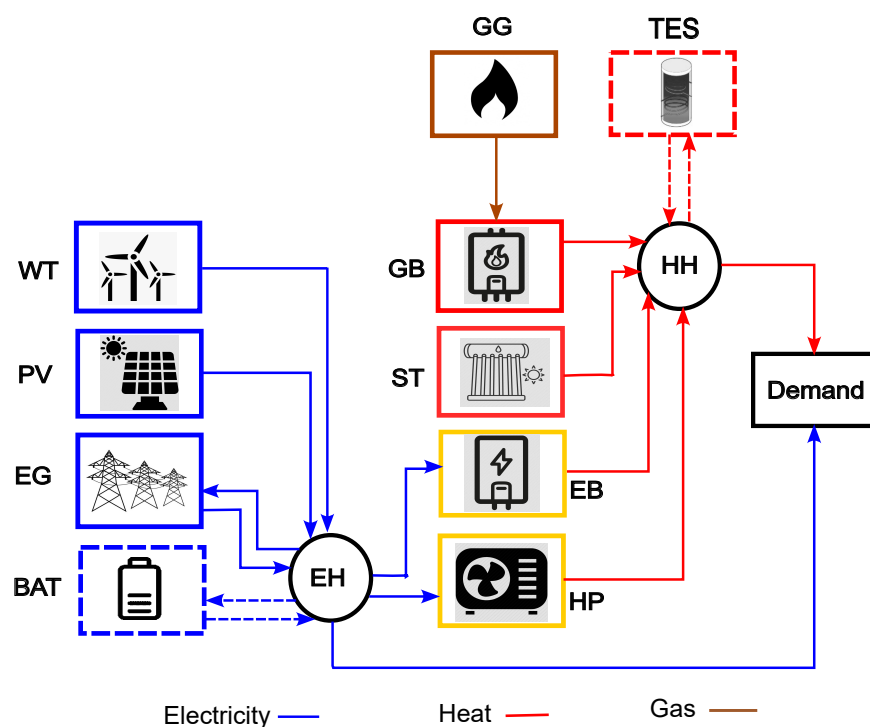

Figure 1. Concept of IES.

The coupled optimization problem of the energy concept shown in Figure 1 is complex due to a large number of continuous variables such as capacity of the components, power consumption of the components, etc., and binary variables such as existence of components in the energy concept are involved with nonlinear modeling of the energy components consisting of a large number of constraints. It falls under the general category of mixed-integer nonlinear programming (MINLP) nonconvex problems, which is computationally very expensive to solve [24]. 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 [25]. Solution accuracy and computational efforts of different models in optimization are compared later in Section 3.

The time horizon for 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 h. A total of 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 at Sanddorn

GmbH for a working day is shown in Figure 2b. 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 Figure 2a.
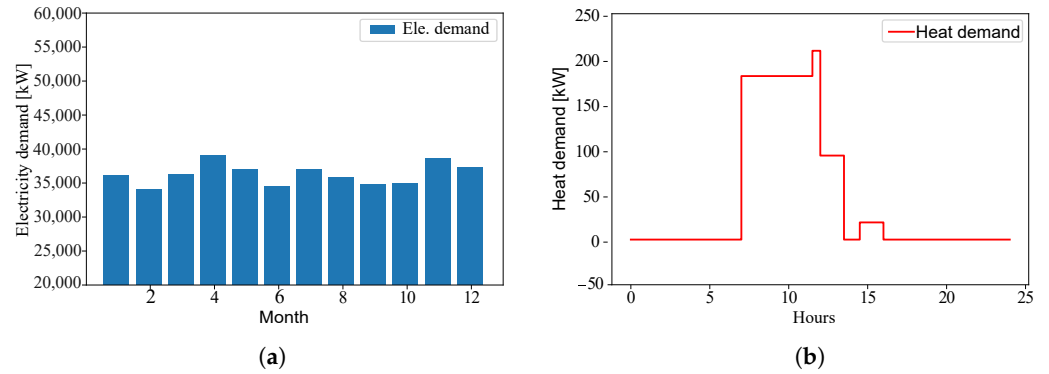

(**a**)   (**b**)

Figure 2. Monthly electricity and daily heat demand of the case study considered. (**a**) Electricity demand, (**b**) heat demand.

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; therefore, the investment cost for building the facility is excluded. GWI is the measurement of the CO_2 emission. Thus, the minimization problem is formulated as

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

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

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

where the operational costs depend on the net electricity and gas bought from the grids, which is shown as

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

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 [25]. $CAPEX$ 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 [25]. The second minimization objective GWI is presented as

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

where $M = [Jan, Feb...., Dec]$ and $S = \{PV, WT, SH, GB, EB, HP, TES\}$. M is the vector that includes the months of the year and S considers the components of the energy concept. $\mathbf{x} = [A^{PV}, P^{WT}_{nom}, Q^{GB}_{nom}, Q^{EB}_{nom}, Q^{HP}_{nom}, Q^{TES}_{nom}]$ 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 the 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 [25,26].

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

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

which indicates that the net heat generated from *ST*, *GB*, *EB*, *HP*, and *TES* should fulfil 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 [25].

Components	Reference Capacity x^0	$CAPEX^0$[€]	γ	α	β	$\tau[a]$	λ_{min}
PV	A^0 [m²]	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 [m²]	400	0.95	0.02	0.03	10	0
GB	Q^0_{nom} [kW]	2700	0.45	0.02	0.03	10	0.2
EB	Q^0_{nom} [kW]	70	0.95	0.01	0.03	10	0
HP	Q^0_{nom} [kW]	1655	0.66	0.02	0.03	10	0
TES	Q^0_{nom} [kWh]	200	0.86	0.01	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 model are explained. List of all variables, constant parameters, and input parameters for physics-driven component modeling is presented in Table 3.

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 Parameters in Each Time-Step	Total Number of Variables
PV	A_{nom}^{PV}	P_{out}^{PV}	η^{PV}, P_{nom}^{PV}	I	3
WT	P_{nom}^{WT}	$\lambda^{WT}, P_{out}^{WT}$	-	v	4
GB	Q_{nom}^{GB}	$\lambda^{GB}, E_{in}^{gas}$	η^0	-	4
EB	P_{nom}^{EB}	$\lambda^{EB}, P_{in}^{EB}$	η^{EB}	-	4
ST	A_{nom}^{ST}	$T_{out}^w, T_{out}, T_{in}$	$I_{AM}, \eta_0, \dot{m}^{ST}$	$I, T_{amb}, \dot{m}^w$	5
HP	Q_{nom}^{HP}	$\lambda^{HP}, h_{out}^w, T_{out}^w, P_{in}^{HP}$	$a, b, c, d, T_{in}^c, T_{in}^w$	$\dot{m}^w$	6
TES	Q_{nom}^{TES}	$Q_{in}^{TES}, Q_{out}^{TES}$	$\eta_{out}^{TES}, \eta_{in}^{TES}, \tau_{in}^{TES}, \tau_{out}^{TES}$	-	4
Electric grid	-	$E_{in}^{el}, E_{out}^{el}$	-	-	2
Gas grid	-	E_{in}^{gas}	-	-	1

2.2.1. Photovoltaic

The electrical power P_{out}^{PV} generated by the PV unit is mostly dependent on the solar irradiance I, the efficiency of PV unit η^{PV} and total area A_{nom}^{PV} of the PV system. It is represented by

$$P_{out}^{PV} \leq A_{nom}^{PV} \cdot \eta^{PV} \cdot I. \tag{9}$$

The case study facility already has 55 kWp capacity 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

$$P_{out}^{PV} \leq A_{nom}^{PV} \cdot P_{nom}^{PV}, \tag{10}$$

where P_{nom}^{PV} is chosen as 0.171 kWm^{-2} [25] 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 majorly limited by the Betz limit, which in turn determines the part-load behavior of WT and its nominal power [27]. 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 $v_{ref} = 12$ m/s [25]). 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 [28]. ST model is based on the quadratic efficiency model developed by [29], 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 [30].

Complete hourly weather data for the location of the case study plant has been gathered from European commission photovoltaic geographical information system [31]. 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 [28]. 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 are 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 A_{nom}^{ST} 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 [28] compared this simplified physical model with detailed simulation model in TRNSYS. The parameters for the physical models are collected in [32], in which extensive research has been carried out for different ST flat collector manufacturer data in Germany and related parameters. Parameters chosen for the case study are shown in Table 4. Figure 3a 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 of 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}.

Table 4. Parameters for ST flat collector.

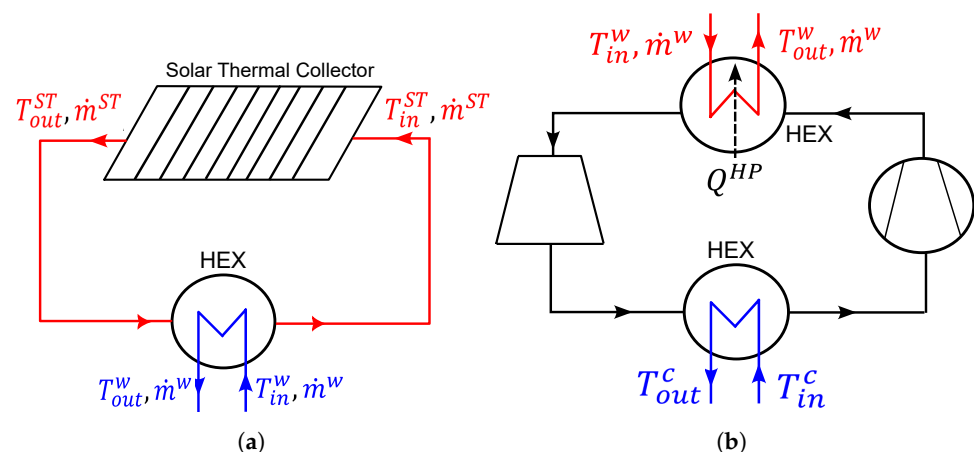| Collector	η_0	a_1 [W/(m^{2}K)]	a_2 [W/(m^{2}K)]	I_{AM}
Flat plate	0.79	4.03	0.0078	0.86

Figure 3. ST collector and HP schematic diagram. (**a**) Solar flat plate collector with HEX, (**b**) HP consisting of a compressor, 2 HEX, and a turbine.

2.2.6. Heat Pump

HP can use industrial waste heat as a thermal energy source as well as renewable electricity as power input to decarbonize industrial thermal processes, which makes it an essential technology [33]. In this case study, constant waste heat is considered. Renewable electricity is generated from PV and WT. There are many different physical models to

describe the performance of the HP. The authors in [33] investigated different case studies of different types of HPs and came up with different coefficient of performance (COP) regression models. Figure 3b shows a heat pump consisting of two HEXs (condenser and evaporator), a compressor, and a turbine. T_{in}^c and T_{out}^c represent input and output temperatures on the cold side, whereas T_{in}^w and T_{out}^w are the input and output temperatures on the hot side, respectively. The heat output of the HP depends on its nominal capacity Q_{nom}^{HP} and part-load λ^{HP}, which is shown as

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

The output temperature T_{out}^w on the hot side depends on the outlet pressure of water/steam p_{out}^w on hot side and enthalpy h_{out}^w of water/steam, which is a function of heat output Q_{out}^{HP} and inlet temperature T_{in}^w of water on the hot side, shown as

$$T_{out}^w = f(h_{out}^w(Q_{out}^{HP}, T_{in}^w), p_{out}^w). \tag{24}$$

The COP is calculated based on the model

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

suggested by [33]. 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_{out}^w < 160 °C). The heat output is a function of COP and the consumed electrical power P_{in}^{HP}, as well as the input and output enthalpy h_{in}^w and h_{out}^w of water/steam on hot side, which is shown as

$$Q_{out}^{HP} = COP \cdot P_{in}^{HP}, \qquad Q_{out}^{HP} = \dot{m}^w \cdot (h_{out}^w - h_{in}^w). \tag{26}$$

Equations (23)–(26) represent the physics-driven model of HP. A detailed list of design and operational variables of HP is 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, $Q_{out}^{HP}, COP, P_{in}^{HP}, \lambda^{HP}, T_{out}^w$; and one design variable, Q_{nom}^{HP}. Data-driven approach for HP is shown as

$$COP = f(\dot{m}^w, T_{in}^c, \lambda^{HP}, Q_{nom}^{HP}), \tag{27}$$

which includes 1 design variable Q_{nom}^{HP}, 1 operational variable λ^{HP}, and 2 input parameters $\dot{m}^w$ and T_{in}^c 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 are 22% and 18% lower 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 (without TES).

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

2.2.7. Thermal Energy Storage

TES plays a crucial role in the energy concept where renewable energy sources such solar energy are involved, because the time of the energy availability and the time of the

consumption of this energy differ frequently [34]. In this study, a lumped model of hot water storage is considered. The energy balance of such a TES model is represented as

$$\frac{dQ^{TES}}{dt} = \eta_{in}^{TES} \cdot Q_{in}^{TES} - \frac{1}{\eta_{out}^{TES}} \cdot Q_{out}^{TES} - \frac{1}{\tau_{loss}^{TES}} \cdot Q^{TES} \tag{28}$$

where Q^{TES} represents thermal energy stored in hot water storage; Q_{in}^{TES} and Q_{out}^{TES} represent charging and discharging heat transfer rates, respectively. η_{in}^{TES} and $\frac{1}{\eta_{out}^{TES}}$ are the charging and discharging efficiencies, which are considered to be constant at 95% [25]. The storage capacity is limited by the minimum and maximum capacities as shown in Equation (8). The heat transfer rates depend also on maximum charging and discharging rates, which are shown as

$$0 \le Q_{in}^{TES} \le \frac{1}{\tau_{in}^{TES}} \cdot Q_{nom}^{TES},$$

$$0 \le Q_{out}^{TES} \le \frac{1}{\tau_{out}^{TES}} \cdot Q_{nom}^{TES}. \tag{29}$$

3. Results

3.1. Data Analysis

As shown in Table 5, the data-driven models have fewer variables and constraints to compute in coupled optimization. In this paper, data-driven models such as linear regression (LR), polynomial regression (PR), and artificial neural networks (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 in each time step and all other constraints are eliminated. Table 6 shows the inputs, outputs, number of parameters, and number of data samples. These data-samples are divided into training and validation data. A total of 75% of the data samples are used for training the models and 25% of the data samples are 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 the coefficient of determination R^2 method [35].

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}	439,199
HP	$\dot{m}^w$, T_{in}^c, λ^{HP}, Q_{nom}^{HP}	COP	206,054

Three types of data-driven models: LR, PR, and ANN, are trained for ST and HP. In particular, PR models have two variants, such as second-degree PR (PR-1) and third-degree PR (PR-2) models. Furthermore, ANNs are feedforward neural networks with two different specifications: (I) two hidden layers, five neurons in each hidden layer (ANN-1); (II) three hidden layers, seven neurons in each hidden layer (ANN-2). The number of hidden layers and the number of neurons in each hidden layer are optimized using the hyperparameter tuning technique. ANNs are trained with the k-fold cross-validation method with k = 4 [36].

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 Figure 4b, where for ANN-2, the predicted output data do not match properly with the actual output data. In contrast, Figure 5 shows results for the two different ANN models of HP, where ANN-2 shows better fitting performance 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 a relatively higher R^2 score for LR and PR models. The training time for LR and PR models is shorter than their respective ANN models. The training time for

ANN is higher due to the large number of weights and biases involved in the ANN model. The 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.

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

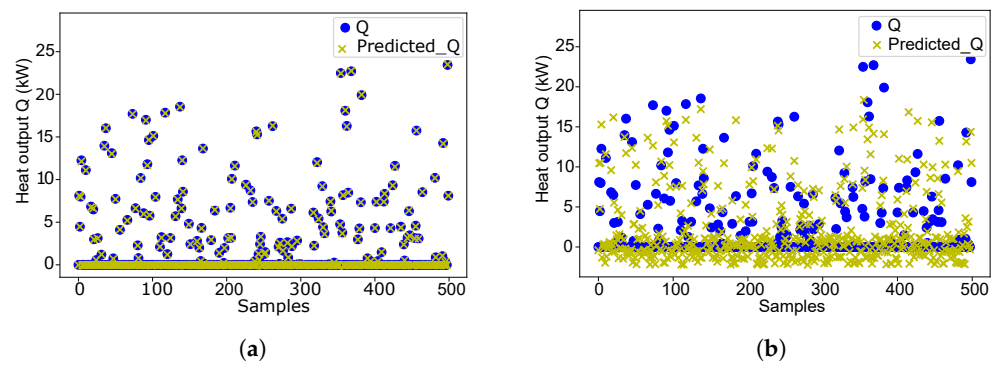| Model	Specification	Training Time for ST [s]	Training Time for HP [s]	R^2 for ST	R^2 for HP
LR	degree 1	2.12	1.82	0.96	0.45
PR-1	degree 2	5.33	4.31	0.972	0.68
PR-2	degree 3	11.26	9.25	0.986	0.825
ANN-1	2 hidden layers 5 neurons each	1254	1008	0.999	0.862
ANN-2	3 hidden layers 7 neurons each	1852	1369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

(**a**) (**b**)

Figure 4. Comparison of ANN models' output prediction to actual output of ST. (**a**) ANN-1 (2 layers, 5 neurons each); (**b**) ANN-2 (3 layers, 7 neurons each).
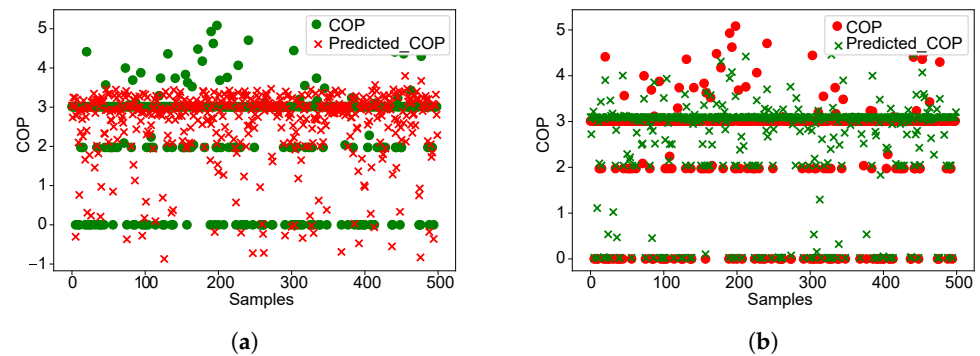

(**a**) (**b**)

Figure 5. Comparison of ANN models' output prediction to actual output of HP. (**a**) ANN-1 (2 layers, 5 neurons each); (**b**) ANN-2 (3 layers, 7 neurons each).

3.2. Optimization Results

3.2.1. Without Tes

Figure 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, Figure 6a 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 the 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.
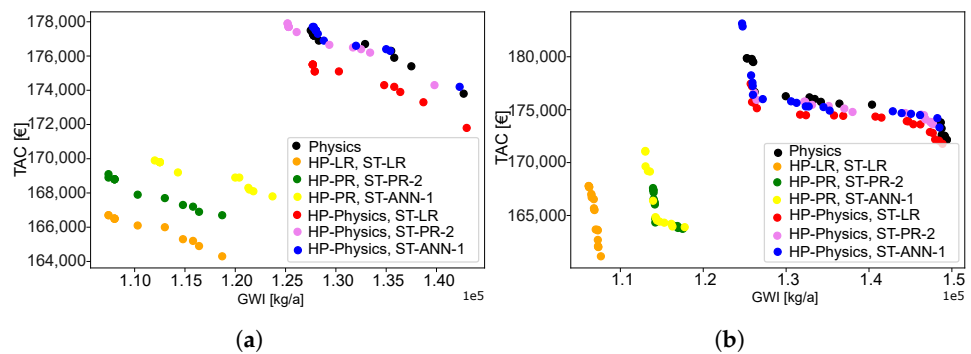
Figure 6. Comparison of the optimization results for different model combinations of ST and HP (without TES). (**a**) Single-level coupled optimization; (**b**) bi-level coupled optimization.

The original multi-objective MINLP optimization problem is linearized and converted into an MILP optimization problem. The MILP problem is solved with the GUROBI solver [37] on the PYOMO platform included in COMANDO [27]. More specifically, the single-level multi-objective Pareto-front is generated by the augmented ϵ-constraint method [38]. The bi-level problem is not linearized and solved on the PYMOO platform with the non-sorting genetic algorithm (NSGA-II) on the design level and differential evolution on the operation level [39]. Both optimization problems, single-level and bi-level, are solved on an 11th Gen Intel(R) Core(TM) i7-1185G7 with 16 GB RAM. It can be seen from Figure 6a,b that the physics-driven model of HP with the 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 the physics-driven model of HP and the PR-2 data-driven model of ST (violet dots). This trend applies to both single-level and bi-level optimization.

Figure 7 shows the comparison of computational time and solution accuracy of the coupled optimization results for combinations of different models. Accuracy is calculated based on the R^2 method, where a complete physics-driven model is reference. It can be seen that the best trade-off between computational time and accuracy is found for the HP physics-driven model and the 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 compared to complete physics-driven models. In contrast, the ST ANN-1 model with the 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 is very high due to a large number of weights and biases.
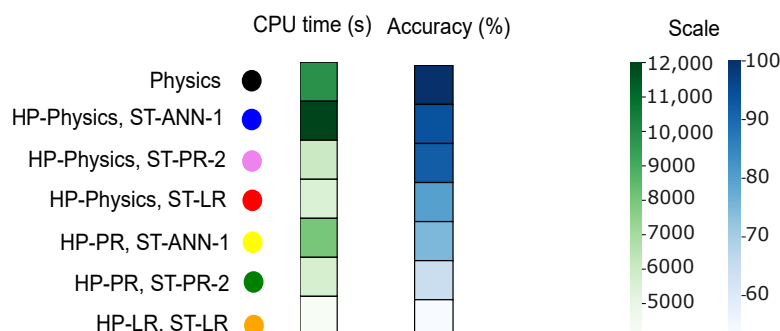

Figure 7. Time and accuracy comparison for different combination of models in single-level multi-objective optimization (without TES).

Figure 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 Figure 7. Both Figure 8a,b show that for minimum TAC, small HP and large GB are 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 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.
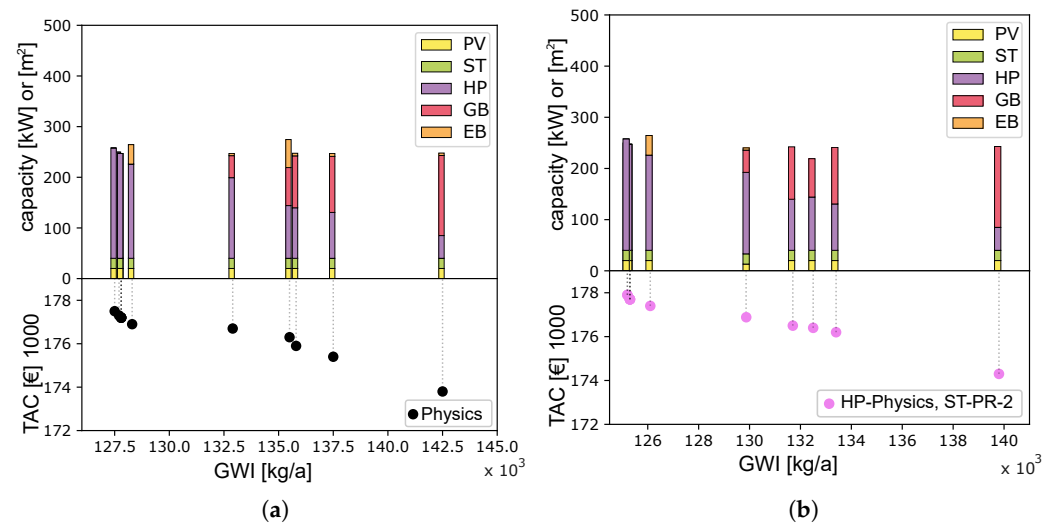

Figure 8. Capacity of components for reference solution and best trade-off solution for single-level coupled optimization. (**a**) Complete physics-driven model; (**b**) HP-Physics, ST-PR-2 model.

3.2.2. With TES

The best trade-off combination models of ST and HP from Section 3.2.1, which are ST-PR-2 and HP-Physics models, are considered to optimize the design and operation of the energy concept, including TES. As shown in Table 8, the number of variables and constraints are reduced by 10% and 13%, respectively, compared to the full physics-driven model of the energy concept. The inclusion of TES increases the overall computational time for coupled optimization as the number of decision variables and constraints increases compared to the energy concept without TES. Computational time with the use of combination of physics- and data-driven models is 29% less than the full physics-driven model.

Table 8. 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 (with TES).

ST	HP	Total Number of Variables	Total Number of Constraints	Computational Time [s]	Accuracy [%]
Physics-driven	Physics-driven	2175	3325	18,025	100
PR-2	Physics-driven	1959	2893	12,671	90.3

Figure 9 shows the comparison of coupled optimization results with the full physics-driven model and the combination of HP physics-driven and ST PR-2 data-driven model. The accuracy, as mentioned, is achieved up to 90%. Due to the inclusion of TES, GWI decreases as thermal energy generated from ST can be stored in TES for later use and the usage of gas and electricity from the grid can be reduced. But at the same time, the TAC increases because of the addition of a new component.
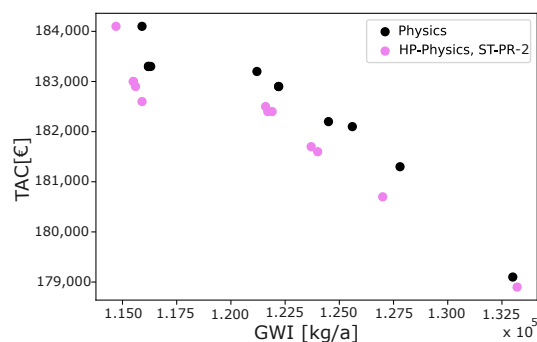
Figure 9. Comparison of the optimization results for different model combinations of ST and HP (with TES).

4. Conclusions and Outlook

This paper shows 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 the 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 ANN models of HP and ST showed a relatively higher accuracy, but the computational effort for the optimization was also relatively quite higher compared to the other combinations. The full physics-driven model of IES was also computationally very expensive to optimize. The results show that the combination of the data-driven PR model of ST and the physics-driven HP model is 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 the complete physics-driven model, and the computational time is reduced by 37% without TES. The integration of TES increased the computational time due to increased flexibility and degree of freedom. The accuracy has been achieved up to 90% and computational time is reduced by 29%.

These results have specifically considered the constraints related to the use case selected for the food and cosmetic industry. The results might change under different constraints. Future work for this case study is to develop and optimize different scenarios such as retrofit designs, greenfield designs, cost-neutral solutions, and complete emission-free solutions for the given case study, and to provide different design operation solutions. The combination of physics–data models can be very useful in optimizing these scenarios, while requiring less computational effort and producing high accuracy. The modeling fidelity can be increased to obtain more accurate results and reduce emissions and costs accordingly. The use of high-performance computing and further reduced-order models can be utilized to achieve the above-mentioned objectives.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DES	District Energy Systems
IES	Integrated Energy Systems
RES	Renewable Energy Sources
TES	Thermal Energy Storage
PV	Photovoltaic
WT	Wind Turbine
ST	Solar Thermal
GB	Gas Boiler
GG	Gas Grid
EG	Gas Grid
EB	Electric Boiler
EH	Electric Hub
HH	Heat Hub
HP	Heat Pump
MINLP	Mixed-Integer Nonlinear Programming
MILP	Mixed-Integer Linear Programming
TAC	Total Annualized Cost
GWI	Gloabl Warming Impact
OC	Operational Cost
CAPEX	Capital Expenditure

Nomenclature

Letter symbols

$\mathbf{x}$	design variables
$\mathbf{y}$	operational variables
$\dot{m}$	massflow, kg/s
A	area, m^2
a	annum
C	investment cost
c	specific heat capacity, kJ/kgK
E	energy, kWh
g	CO_2 emission factor, g-CO_2eq/kWh
GWI	global warming impact, g/kWh
I	solar irradiance, kW/m^2
M	set of months
OC	operational cost
P	power, kW
p	price
Q	thermal capacity, kW
S	set of components
s	seconds
T	temperature, K
TAC	total annualized cost
z	binary variables

Greek symbols

α	maintenance cost factor
β	interest rate
ϵ	numerical limit
η	efficiency
γ	scaling exponent
λ	part load

τ	time horizon
Subscripts and superscripts	
0	reference
amb	ambient
buy	electric boiler
EG	electricity grid
el	electricity
gas	gas
GB	gas boiler
GG	gas grid
HP	heat pump
i	index for components
in	inlet
L	loss
nom	nominal
out	outlet
PV	photovoltaic
sell	selling
ST	solar thermal
u	useful
w	water
WT	wind turbine

References

1. Voll, P.; Klaffke, C.; Hennen, M.; Bardow, A. Automated superstructure-based synthesis and optimization of distributed energy supply systems. *Energy* **2013**, *50*, 374–388. [CrossRef]
2. Mayer, M.J.; Szilagyi, A.; Grof, G. Ecodesign of ground-mounted photovoltaic power plants: Economic and environmental multi-objective optimization. *J. Clean. Prod.* **2021**, *278*, 123934. [CrossRef]
3. Aguilar, O.; Perry, S.J.; Kim, J.-K.; Smith, R. Design and optimization of flexible utility systems subject to variable conditions. *Chem. Eng. Res. Des.* **2007**, *85*, 1149–1168. [CrossRef]
4. Velasco-Garcia, P.; Varbanov, P.S.; Arellano-Garcia, H.; Wozny, G. Utility systems operation: Optimisation based decision making. *Appl. Therm. Eng.* **2011**, *31*, 3196–3205. [CrossRef]
5. Tina, G.M.; Passarello, G. Short-term scheduling of industrial co-generation systems for annual revenue maximisation. *Energy* **2012**, *42*, 46–56. [CrossRef]
6. Bouvy, C.; Lucas, K. Multicriterial optimisation of communal energy supply concepts. *Energy Convers. Manag.* **2007**, *48*, 2827–2835. [CrossRef]
7. Weber, C.; Shah, N. Optimisation based design of a district energy system for an eco-town in the united kingdom. *Energy* **2011**, *36*, 1292–1308. [CrossRef]
8. Keirstead, J.; Shah, N. Calculating minimum energy urban layouts with mathematical programming and monte carlo analysis techniques. *Comput. Environ. Urban Syst.* **2011**, *35*, 368–377. [CrossRef]
9. Lozano, M.A.; Ramos, J.C.; Carvalho, M.; Serra, L.M. Structure optimization of energy supply systems intertiary sector buildings. *Energy Build.* **2009**, *41*, 1063–1075. [CrossRef]
10. Liu, P.; Pistikopoulos, E.N.; Li, Z. An energy systems engineering approach to the optimal design of energy systems in commercial buildings. *Energy Policy* **2010**, *38*, 4224–4231. [CrossRef]
11. Barton, P.; Li, X. Optimal design and operation of energy systems under uncertainty. In Proceedings of the IFAC International Symposium on Dynamics and Control of Process Systems, Mumbai, India, 18–20 December 2013; Volume 46, pp. 105–110.
12. Li, J.; Zhao, H. Multi-objective optimization and performance assessments of an integrated energy system based on fuel, wind and solar energies. *Entropy* **2021**, *23*, 431. [CrossRef] [PubMed]
13. Patwal, R.S.; Narang, N. Multi-objective generation scheduling of integrated energy system using fuzzy based surrogate worth trade-off approach. *Renew. Energy* **2020**, *156*, 864–882. [CrossRef]
14. Fazlollahi, S.; Becker, G.; Ashouri, A.; Marechal, F. Multi-objective, multi-period optimization of district energy systems: Iv–A case study. *Energy* **2015**, *84*, 365–381. [CrossRef]
15. Morvaj, B.; Evins, R.; Carmeliet, J. Optimising urban energy systems: Simultaneous system sizing, operationand district heating network layout. *Energy* **2016**, *116*, 619–636. [CrossRef]
16. Söderman, J.; Pettersson, F. Structural and operational optimisation of distributed energy systems. *Appl. Therm. Eng.* **2006**, *26*, 1400–1408. [CrossRef]
17. Zhang, C. Data Driven Modeling and Optimization of Energy Systems. Ph.D Thesis, Nanyang Technological University, Singapore, 2019.

18. Zhang, C.; Cao, L.; Romagnoli, A. On the feature engineering of building energy data mining. *Sustain. Soc.* **2018**, *39*, 508–518. [CrossRef]
19. Ekonomou, L. Greek long-term energy consumption prediction using artificial neural networks. *Energy* **2010**, *2*, 512–517. [CrossRef]
20. Ahmad, A.S.; Hassan, M.Y.; Abdullah, M.P.; Rahman, H.A.; Hussin, F.; Abdullah, H.; Saidur, R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renew. Sustain. Energy Rev.* **2014**, *33*, 102–109. [CrossRef]
21. Liang, J.; Du, R. Thermal comfort control based on neural network for HVAC application. In Proceedings of the 2005 IEEE Conference on Control Applications, CCA 2005, Toronto, ON, Canada , 28–31 August 2005; pp. 819–824. [CrossRef]
22. Koch, V.; Kuge, S.; Geissbauer, R.; Schrauf, S. *Industry 4.0-Opportunities and Challenges of the Industrial Internet*; Strategy Former Booz Company, PwC: Berlin, Germany, 2014; Volume 13, pp. 1–51.
23. Caylar, P.L.; Oliver, N.; Kedar, N. *Digital in Industry: From Buzzword to Value Creation*; McKinsey Digit: Berlin, Germany, 2016; pp. 1–9.
24. Dimopoulos, G.G.; Kougioufas, A.V.; Frangopoulos, C.A. Synthesis, design and operation optimization of a marine energy system. *Energy* **2008**, *33*, 180–188. [CrossRef]
25. Sass, S.; Faulwasser, T.; Hollermann, D.E.; Kappatou, C.D.; Sauer, D.; Schutz, T.; Shu, D.Y.; Bardow, A.; Groll, L.; Hagenmeyer, V.; et al. Model compendium, data, and optimization benchmarks for sector-coupled energy systems. *Comput. Chem. Eng.* **2020**, *135*, 106760. [CrossRef]
26. Umweltbundesamt Kohlendioxid-Emissionen. 2020. Available online: https://www.umweltbundesamt.de/daten/klima/treibhausgas-emissionen-in-deutschland/kohlendioxid-emissionen#herkunft-und-minderung-von-kohlendioxid-emissionen (accessed on 7 May 2023).
27. Langiu, M.; Shu, D.Y.; Baader, F.J.; Hering, D.; Bau, U.; Xhonneux, A.; Müller, D.; Bardow, A.; Mitsos, A.; Dahmen, M. Comando: A next-generation open-source framework for energy systems optimization. *Comput. Chem. Eng.* **2021**, *152*, 107366. [CrossRef]
28. Grahovac, M.; Liedl, P.; Frisch, J.; Tzscheutschler, P. Simplified Solar Collector Model: Hourly Simulation of Solar Boundary Condition for Multi-Energy Optimization. In Proceedings of the International Congress on Heating, Refrigerating and Air-Conditioning, Belgrade, Serbia, 1–3 December 2010; Volume 41, pp. 21–23.
29. Duffie, J.; Beckman, W.A. *Solar Engineering of Thermal Processes*; John Wiley and Sons Inc.: Hoboken, NJ, USA, 1991.
30. *DIN EN 12975*; Thermische Solaranlagen und Ihre Bauteile–Kollektoren–Teil 2: Prüfverfahren. DIN: Weimar, Germany, 2006.
31. EU-Science-Hub Photovoltaic Geographical Information System 2020. Available online: https://joint-research-centre.ec.europa.eu/pvgis-online-tool/pvgis-tools/hourly-radiation_en (accessed on 20 January 2023).
32. SPF Research. Research and Development for Sustainable Energy Systems: Flat Plate Collectors. Available online: https://www.ost.ch/ (accessed on 2 February 2023).
33. Schlosser, F.; Jesper, M.; Vogelsang, J.; Walmsley, T.G.; Arpagaus, C.; Hesselbach, J. Large-scale heat pumps: Applications, performance, economic feasibility and industrial integration. *Renew. Sustain. Energy Rev.* **2020**, *133*, 110219. [CrossRef]
34. Farkas, I.; Toth, J. Mathematical modelling of solar thermal collectorsand storages. *Acta Technol. Agric.* **2019**, *22*, 128–133. [CrossRef]
35. Model Selection Using R-Squared (R²) Measure. Available online: https://towardsdatascience.com/the-complete-guide-to-r-squared-adjusted-r-squared-and-pseudo-r-squared-4136650fc06c (accessed on 7 March 2023).
36. Ahmed, F.Y.H.; Ali, Y.H.; Shamsuddin, S.M. Using K-Fold Cross Validation Proposed Models for Spikeprop Learning Enhancements. *Int. J. Eng. Technol.* **2018**, *7*, 145–151. [CrossRef]
37. Non-Convex Quadratic Optimization. Available online: https://www.gurobi.com/events/non-convex-quadratic-optimization/ (accessed on 27 February 2023).
38. Mavrotas, G. Effective implementation of the e-constraint method in Multi-Objective Mathematical Programming problems. *Appl. Math. Comput.* **2009**, *213*, 455–465. [CrossRef]
39. Slowik, A.; Kwasnicka, H. Evolutionary algorithms and their applications to engineering problems. *Neural Comput. Appl.* **2020**, *2*, 12363–12379. [CrossRef]