Energy & Buildings 295 (2023) 113277

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Hybrid photovoltaic-thermal solar collector modelling with parameter identification using operation data

François Veynandt *, Peter Klanatsky, Helmut Plank, Christian Heschl

Burgenland University of Applied Sciences, Campus Pinkafeld, Steinamangerstraße 21, 7423 Pinkafeld, Austria

ARTICLE INFO

Keywords:
Hybrid photovoltaic-thermal (PVT) solar collector
Data-driven parameter identification
Grey-box model
Real operation measurement
Optimal control of buildings

ABSTRACT

Hybrid photovoltaic-thermal (PVT) solar collectors combine photovoltaic and thermal technologies to produce both electricity and thermal energy. To integrate efficiently PVT solar collectors, an optimal control approach is crucial. Accurate information about their thermal and electrical yields under real operating conditions is necessary for applying such an approach. Obtaining this information requires detailed testing and modelling of the PVT solar collectors. However, this is time consuming and does not take into account specific local characteristics. Consequently, a data-driven approach based on grey box modelling, automated parameter identification and numerical optimization becomes important. This approach can be particularly valuable in situations where traditional modelling methods do not fully capture the actual operating and installation conditions. Therefore, this paper presents several PVT grey-box model structures and analyses their accuracy and reliability using own experimental data. It also analyses the parameter identification performance and identifies the most appropriate model variants for optimal control application within a data-driven framework.

1. Introduction

1.1. PVT-control aspects

Photovoltaic-thermal (PVT) solar collectors are solar collectors that combine the functionality of a photovoltaic (PV) module, which produces electricity, and a solar thermal (T) collector, which produces heat from solar radiation. PVT collectors make the most of the available surface area by producing more energy –electricity and heat– than the separate technologies [1]. Depending on the technology, PVT collectors provide heat at different temperature levels. Uncovered collectors can be used as source for heat pumps [2] or for swimming pools. Covered collectors can provide domestic hot water and support space heating in buildings. They can also provide process heat in industry, up to high temperature levels above 150 °C with concentrating solar technologies. This wide range of applications shows that PVT collectors can play an efficient role in the necessary energy transition [3,4]. Several reviews point out the need of further developments to make the most of this promising technology [5,6], in particular for the optimal connection mode and integration of PVT collectors to the electrical grid [7].

At the same time, the increasing share of renewable energy requires more flexibility on the supply and demand side. Especially for the optimal operation of buildings, data-driven predictive control is strongly gaining interest to harness the flexibility potential on the demand side [8]. Prediction in the near future can provide useful information for an optimised control system, to maximise the collector energy yield and the energy efficiency of the associated building. By comparing the expected and actual performance, maintenance actions can also be identified. On the building-related supply side, the analysis of smart building-integrated photovoltaics (SBIPV) shows the great potential of data-driven approaches [9]. This review includes an overview of artificial intelligence methods applied to demand forecasting. It also emphasises that the topic goes beyond energy issues: it includes societal and political aspects.

The control strategy of a PVT system depends on several parameters: priority on heat or electricity production, temperature level of the useful heat, combination with a heat pump, etc. It must also take into account information on the state of the building, such as its current energy demand. The integration of PVT energy is particularly promising in combination with ground source heat pumps (GSHP), as recently reviewed in details [10]. The review points out the need of transient models to optimise the design as well as the operation of such hybrid systems, over their entire lifetime. The correct regeneration of the soil around the

* Corresponding author.
E-mail addresses: Francois.Veynandt@fh-burgenland.at (F. Veynandt), Peter.Klanatsky@fh-burgenland.at (P. Klanatsky), Helmut.Plank@fh-burgenland.at (H. Plank), Christian.Heschl@fh-burgenland.at (C. Heschl).

https://doi.org/10.1016/j.enbuild.2023.113277
Received 28 March 2023; Received in revised form 31 May 2023; Accepted 15 June 2023
Available online 28 June 2023

Nomenclature

List of symbols

IAM	Incidence Angle Modifier
MAE	Mean Absolute Error
MPPT	Maximum Power Point Tracking
PVT	Photovoltaic and Thermal (hybrid solar collector)
PV	Photovoltaic (solar technology)
RC	Resistance and Capacitance (model)
SBIPV	Smart Building Integrated Photovoltaic
WISC	wind and/or infrared sensitive collectors

ground heat exchanger is indeed important to ensure. Most studies use on–off or fuzzy logic controllers for the operation of hybrid PVT-GSHP systems. In [11], a machine-learning based predictive control of a PVT-GSHP has been used, predicting the room temperature, with detailed dynamic physical models implemented in TRNSYS. A second study demonstrated a model-based optimal control on a PVT system coupled with a GSHP [12]. Relying on simple simulations, it predicts 7 % energy savings and 5 % more power production than with a standard controller.

The results show that the efficient integration of PVT systems requires an intelligent control strategy. This requires an accurate knowledge of the expected electrical and thermal yields, which depend on weather conditions as well as fluid inlet temperature and mass flow rate. Only then can the optimal operating point for the overall system be identified. To achieve this, it is essential to have models that are highly accurate and yet simple enough to minimise the computational complexity of the optimisation task. In this context, grey-box models with automated parameter identification are very promising. Since they already contain physical information, they require significantly shorter training times compared to black-box models. Furthermore, with the known model structure, a more efficient implementation in the numerical optimiser (solver-based approach) can be realised while maintaining the data-driven approach with high scalability. Therefore, the investigation of suitable grey-box models, including automated parameter identification, represents a significant contribution to the provision of more efficient PVT control strategies.

1.2. Data-driven PVT modelling

Although PVT collectors are relatively young in the history of solar technology, they have been well studied [13]. The technology in question varies, with different designs and fluids, including mainly air- or water-cooled collectors. Many studies propose numerical models for PVT collectors, ranging from simple steady-state equations to dynamic models. The focus is on energy and often also exergy performance, sometimes including economic or environmental aspects. The models are mostly used to optimise the design and the sizing of the systems [14–17]. However, modelling for predictive control strategies is still little explored. To this aim, Xia et al., [12] used simple linear efficiency equations to calculate the energy performance of the PVT collector.

Considering the potential of PVT collectors and the efforts required to optimise the use of renewable energy, this paper proposes to explore more in details the modelling of the PVT production for data-driven predictive control. Simplified robust physical models are used for automated data-driven parameter identification. The calibration process is performed with recent past measurement data from real operation.

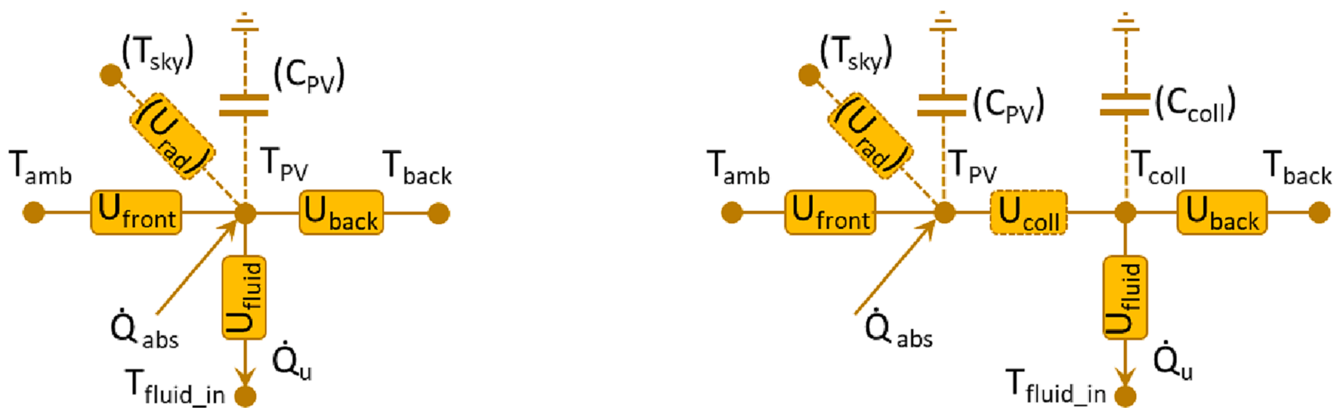

Fig. 1. Structure of the grey-box model for the thermal behaviour with one temperature node T_{PV} (left) and two temperature nodes T_{PV} and T_{coll} (right). Optional elements of the variants in dashed lines (– –): long wave radiation to the sky U_{rad}, capacitances C_{PV} and C_{coll}.

The most common modelling approach is to relate the PV cell temperature with a heat transfer coefficient to the mean fluid temperature on the thermal collector model side of the model. This is the case in the PVT model proposed by Florschuetz [18], which is an extension of the Hottel-Whillier model. This model is available in the TRNSYS reference software under the name Type 560 [19]. Other resistance and capacitance (RC) models, with one or two nodes, have been proposed for the parameter identification of solar thermal collectors. As Helmers und Kramer [20] showed, the quasi-dynamic method can be extended to better characterise the electrical and thermal performance under dynamic operating conditions. This is the method recommended by the IEA SHC Task 60 on the application of PVT collectors for parameter identification [21]. With the aim of establishing a standard model for PVT, the two-node, one-capacitance thermal model is preferred due to its compliance with the ISO 9806 standard. This standard defines reference methods for testing solar collectors, including a model of solar thermal collectors suitable for parameter identification, using a quasi-dynamic test method. This test method allows the use of measured data from operation over a whole day. In order to take into account the dynamic effects of the operating conditions, the quasi-dynamic model better represents the different positions of the sun (direct and diffuse share of solar radiation), the thermal inertia of the collector. In addition, the heat loss dependence on wind speed and long-wave irradiance is especially recommended for so-called wind and/or infrared sensitive collectors (WISC) [22]. According to Jonas et al., [21], the two-node, one-capacitance model allows successful parameter identification for all collector designs, except for integrated collector storage systems (high thermal inertia and long residence time). Jonas et al., [21] use Genopt for the parameter identification, with the Hooke-Jeeves algorithm and mean average error as the objective function. A simultaneous fit to the thermal and electrical power output is recommended.

The present work builds on this known experience to investigate data-driven parameter identification from monitoring under real operating conditions. The models investigated are grey-box models, in several variants. The models are developed using the MATLAB simulation environment [23], because of its strength in analysing large data sets. The thermal behaviour is modelled in several variants having one or two temperature-nodes, with or without thermal capacitances (R or RC). The models are tested on real data from an experiment, with a water-based PVT collector, with tubes and fins.

The model variants are tested in real operating conditions, using measured conditions as a training data set. The accuracy of this calibrated model may be limited in situations that differ significantly from the measured data. A sufficiently long data set can help to improve the variety of operating conditions used for training. The training can be updated throughout the lifetime of the collector.

2. Grey-box models used for PVT collectors

2.1. General model structure

To capture the main dynamics of the PVT collector, several variants of resistance–capacitance (RC) models are proposed. The structure of these RC model variants are shown in Fig. 1: with one temperature node (left) and with two temperature nodes (right). The first temperature node is T_{PV}, the PV cell temperature. The additional node is T_{coll}, the collector core temperature. The absorbed solar power is delivered at the T_{PV} node and the long-wave thermal radiation exchange with the sky (U_{rad}) is also applied to this node. The thermal fluxes (and the associated heat transfer coefficients) are split between the useful heat transferred to the fluid (U_{fluid}) and the losses to the front (U_{front}) and back (U_{back}) of the collector. The two temperature nodes are connected by a heat transfer coefficient U_{coll}. The thermal inertia of the collector can be modelled by a thermal capacitance applied to one or two temperature nodes in the collector. In Fig. 1, the dashed lines indicate the parts that are changed in the model variants. According to the variant analysis (see section 4.2), the best two-node option is the model with a thermal capacitance C_{coll} only at the T_{coll} node.

2.2. Initial node temperatures

In a dynamic thermal model, the initial temperature of the nodes must be determined to enable the calculation. The values of these state variables are not known from measurements, so they must be estimated.

The node temperatures are calculated at the initial time step $k = 0$, as in the case where no thermal capacity is considered, here with the option of radiative losses to the sky temperature:

$$T_{coll} = \frac{T_{fluid_{in}} U_{fluid}\left(\dot{m}_{fluid_{in}} > 0\right) + U_{back} T_{back} + U_{coll} T_{PV}}{U_{fluid}\left(\dot{m}_{fluid_{in}} > 0\right) + U_{back} + U_{coll}}$$

$$T_{PV} = \frac{\dot{Q}_{abs} + U_{front} T_{amb} + U_{coll} T_{coll} + T_{sky} U_{rad}}{U_{front} + U_{coll} + U_{rad}}$$

where $\dot{Q}_{abs}$ is the solar power absorbed as heat in the collector (W),

T_{amb} the ambient air temperature, which can be measured or provided by a weather station (°C),

T_{back} the temperature at the back of the collector: wall or indoor temperature for integrated collectors or close to the ambient temperature for free-standing collectors (°C),

T_{fluid_in} the temperature of the heat transfer fluid at the inlet of the collector field (°C),

$\dot{m}_{fluid_{in}}$ the mass flow of the heat transfer fluid flowing through the collector field (kg/s); $\left(\dot{m}_{fluid_{in}} > 0\right)$ is a Boolean: 1 if the mass flow is greater than zero, 0 if there is no mass flow; this reflects the fact that no thermal power is extracted when the mass flow is zero,

U_{back} the heat transfer coefficient of the losses to the back of the collector array (W/K),

U_{front} the heat transfer coefficient of losses to the front of the collector array (W/K),

U_{rad} the equivalent heat transfer coefficient for long wave radiation (W/K),

U_{fluid} the heat transfer coefficient from the collector to the fluid (W/K).

2.3. Irradiance available for the thermal side

Only part of the global solar irradiance is absorbed and the solar power converted into electricity must be subtracted, to obtain the absorbed thermal power available $\dot{Q}_{abs}$ (W). It can be written as:

$$\dot{Q}_{abs} = (1 - \eta_{PV}) F_{th} A_{PVT} I_{at}$$

where η_{PV} is the PV efficiency, defined in details in the following section 2.4,

F_{th} a thermal efficiency factor that adjusts the actual heat absorption in the collector from the irradiance that is not converted to electricity (expected to be close to 1),

A_{PVT} the surface area of the collector (m^2),

I_{at} the total irradiance absorbed in the tilted plane (W/m^2).

2.4. Electrical model: Temperature and irradiance dependency of PV conversion efficiency

The photovoltaic panel efficiency η_{PVref} is determined at standard conditions of irradiance and temperature. The effective PV efficiency η_{PV} is known to decrease with increasing temperature and to a lesser extent with increasing irradiance:

$$\eta_{PV} = \eta_{PVref} \cdot \left(1 + f_{corr\_T}\left(T_{PV} - T_{PVref}\right)\right) \cdot \left(1 + f_{corr\_I}\left(I_{gt} - I_{gtref}\right)\right)$$

where η_{PVref} the reference photovoltaic efficiency (-), at standard conditions T_{PVref} and I_{gtref},

T_{PV} the actual temperature of the photovoltaic cells (°C),

I_{gt} the actual global irradiance in the tilted collector plane (W/m^2),

f_{corr_T} the correction coefficient for temperature (1/°C),

f_{corr_I} the correction coefficient for irradiance (m^2/W).

The correction coefficients f_{corr_T} and f_{corr_I} are usually known from the technology characteristics, but they can also be identified.

2.5. Solar irradiance: Distinction of diffuse and direct parts

For a more reliable determination of the solar irradiance absorbed by the collector I_{at}, the diffuse and direct parts of the global irradiance can be considered separately. This is particularly important when the global irradiance at the collector plane is not measured directly. Diffuse irradiance has two components, one from the sky and one from the ground. Direct irradiance is absorbed differently depending on the angle of incidence at the collector plane. It can be written as follow:

$$I_{at} = \tau\alpha(\theta_{inc}) \cdot \left(I_{g_t} - I_{d\,sky} - I_{d\,gnd}\right) + \tau\alpha\left(\theta_{eff\,sky}\right) I_{d\,sky} + \tau\alpha\left(\theta_{eff\,gnd}\right) I_{d\,gnd}$$

where $\tau\alpha(\theta)$ is the transmittance and absorptance of the solar irradiance through and on the surface of the collector, as a function of the angle of incidence θ of the solar irradiance.

I_{g_t} the global incident radiation in the tilted collector plane (W/m^2),

$I_{dgnd} = \rho_{gnd} \frac{1 - \cos\beta}{2} I_{g_h}$ the incident diffuse radiation reflected from the ground (W/m^2),

Fig. 2. Photos of the PVT collector test rig: front (above left) and rear (above right) views. Nearby weather station (bottom centre).

$I_{dsky} = \frac{1+\cos\beta}{2} I_{dh}$ the incident diffuse radiation from the sky (W/m^2),

θ_{inc} the angle of incidence of the direct solar radiation, calculated from the date, time and collector position: tilt angle β, azimuth γ, latitude (0 at the equator, positive to the North) and longitude (0 at the Greenwich meridian, positive to the East) (°),

θ_{effsky} the effective or equivalent angle of incidence of the radiation from the sky (°),

θ_{effgnd} the effective or equivalent angle of incidence of the radiation from the ground (°),

For the diffuse irradiance, the effective angles of incidence are calculated using the classical relationships from Brandemuehl & Beckman [24]:

$$\begin{cases} \theta_{effsky} = 59.68 - 0.1388\beta + 0.001497\beta^2 \\ \theta_{effgnd} = 90 - 0.5788\beta + 0.002693\beta^2 \end{cases}$$

where β is the tilt angle of the collector (°).

The function $\tau\alpha(\theta)$ is defined by the model of Souka & Safwat [25] for angles up to 60°, completed with a linear segment up to 90°:

$$\tau\alpha(\theta) = \begin{cases} \forall 0 \leq \theta < 60^\circ & 1 - b_0 \cdot \left(\frac{1}{cos\theta} - 1\right) \\ \forall 60 \leq \theta \leq 90^\circ & (1 - b_0)\left(1 - \frac{\theta - 60}{30}\right) \end{cases} \quad \text{with } b_0 \text{ first order}$$

IAM.

The IAM –short for Incidence Angle Modifier– characterises the optical behaviour of the collector surface for off-normal radiation.

2.6. Radiative losses to sky: Equivalent heat transfer coefficient for long wave radiation

The equivalent heat transfer coefficient from the collector to the sky, U_{rad}, is actually defined as a function of the equivalent radiative sky temperature, T_{sky}:

$$U_{rad} = A_{PVT}\, \varepsilon_{PVT}\, \sigma \cdot \left(T_{PV} + T_0 + T_{sky} + T_0\right) \cdot \left(\left(T_{PV} + T_0\right)^2 + \left(T_{sky} + T_0\right)^2\right)$$

where A_{PVT} the surface area of the PVT collectors (m^2),

ϵ_{PVT} the longwave emittance of the PVT collector surface (-),

$\sigma = 5.67 \times 10^{-8} W/m^2/K^4$ the Stefan-Boltzmann constant,

T_{PV} the PV cell temperature (°C),

T_{sky} the equivalent radiative temperature of the sky (°C),

$T_0 = 273.15^\circ C$ to convert temperatures from °C to K.

2.7. Mass flow dependency of thermal resistance from collector to heat transfer fluid

To model the thermal transfer from the collector to the heat transfer fluid in a simple way, the thermal resistance is calculated as a function of the mass flow as described for transient cases by Koschenz & Lehmann in [26]. The equation is adapted for parameter identification. To avoid infinite values when the mass flow is equal to zero, the inverse of the resistance U_{fluid} is defined (K/W):

$$U_{fluid} = \frac{1}{\frac{r_{1fluid}}{\dot{m}_{fluid\_in}} + r_{0fluid}}$$

where $\dot{m}_{fluid\_in}$ is the mass flow through the PVT collector (kg/s) and.

r_{1fluid} and r_{0fluid} are identified parameters, in (K/W).(kg/s) and in K/W respectively.

2.8. Wind influence

In the light of several tests comparing different models, a simple

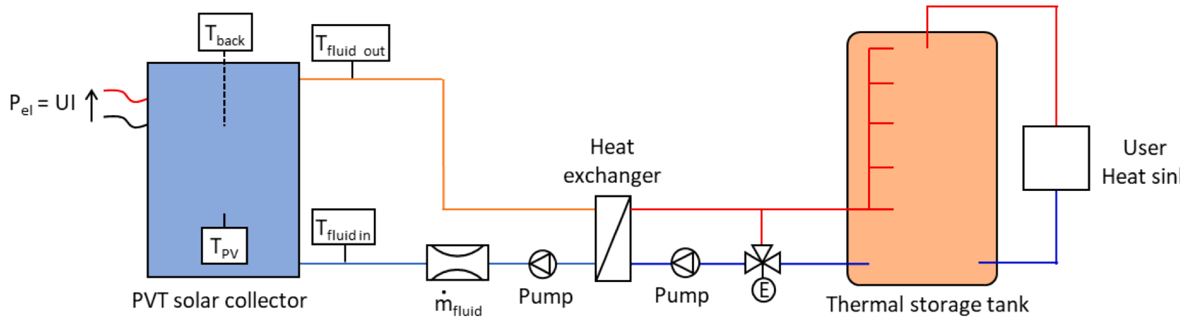

Fig. 3. Schematic of the hybrid PVT collector test bench with simplified hydraulic circuit.

Table 1
Description of measured variables, used in the model as input or as output: collector data.

Measured variable	Name	Unit	Details
Collector input variables for the model			
Inlet temperature of the heat transfer fluid	$T_{fluid\_in_m}$	°C	Measured at the collector inlet.
Mass flow rate of the heat transfer fluid	$\dot{m}_{fluid_{in}}$	kg/s	Measured as volumetric flow rate $\dot{V}_{fluid_m}$ (L/h) with a magnetic-inductive (MID) flow meter. The mass flow rate is: $\dot{m}_{fluid_{in}} = \frac{\rho_{fluid}\dot{V}_{fluid_m}}{3600 \times 1000}$ with ρ_{fluid} the fluid density (kg/m^3). The very small negative values (no flow) are corrected to zero.
Collector back surface temperature	T_{back_m}	°C	Measurement from one sensor on the backside of the collector.
Collector tilt	β_m	°	Measured from the horizontal. Converted from mA signal.
Azimuth of the collector	γ_m	°	Measured from the North, positive to the East. Converted from mA signal.
Collector output variables, used as calibration variables of the model			
Outlet temperature of the heat transfer fluid	$T_{fluid\_out_m}$	°C	Measured at collector outlet. Not directly used for calibration. See thermal power.
Useful thermal power	$\dot{Q}_u$	W	At each time step: $\dot{Q}_u = \dot{m}_{fluid_m} c_{p_{fluid}} (T_{fluid_{out_m}} - T_{fluid_{in_m}})$ with $c_{p_{fluid}}$ the specific thermal capacity (J/(kg.K)) of the heat transfer fluid.
Electrical power	P_{el_m}	W	Calculated from the measured current and voltage at the PV panel terminals.
PV cell temperature	T_{PV}	°C	Important parameter of the model: see temperature dependence of PV cell efficiency. It may be useful to check the fit between the calculated T_{PV} and the measured T_{PVm}.

model is proposed in which the front coefficient U is modified according to the wind speed. The thermal transfer coefficient U_{front_w} replaces U_{front} in the collector energy balance:

$$U_{front_w} = \frac{1}{\frac{1}{U_{front}} + r_{ref_{wind}} v_{wind}^{-n_{wind}}}$$

where U_{front} is the front heat transfer coefficient (identified parameter), characterising the thermal losses to the front of the collector, without wind,

v_{wind} wind speed (m/s).

$r_{ref_{wind}}$ and n_{wind} respectively reference coefficient and exponent, defining the wind speed resistance as a function of the wind speed.

Table 2
Description of measured variables, used as input or as output in the model: weather data.

Measured variable	Name	Unit	Details
Weather data at the collector			
Global irradiance in the tilted collector plane	I_{gt_m}	W/m^2	An occasional error occurs at night, with values above 5000 W/m^2: these are corrected to 0 W/m^2. From SMP10 pyranometer (manufacturer KIPP&ZONEN)
Ambient air temperature	T_{amb_m}	°C	Measured outside.
Wind speed in the collector plane	v_{wind_m}	m/s	"Actual wind speed": 10 s average of second measurements, from WS200 sensor (manufacturer Lufft)
Weather data from weather station (65 m away from the collector)			
Global solar irradiance in the horizontal plane	I_{gh_m}	W/m^2	From SMP21 pyranometer (manufacturer KIPP&ZONEN)
Diffuse solar irradiance in the horizontal plane	I_{dh_m}	W/m^2	From SMP21 pyranometer (manufacturer KIPP&ZONEN) with ball shadow
Equivalent radiative temperature of the sky	T_{sky_m}	°C	Calculated from $I_{ir_{sensor1}}$ downward longwave radiation, from SGR4 pyrgeometer (manufacturer KIPP&ZONEN)

2.9. Iterative definition of the temperature at nodes:

For the following time steps $k > 0$, the node temperature T_{coll} with thermal capacitance C_{coll} is calculated from the state of the previous time step:

$$T_{coll}^{k+1} = T_{coll}^{k} + \left(\left(T_{back}^{k} - T_{coll}^{k}\right) \cdot U_{back}^{k} + \left(T_{PV}^{k} - T_{coll}^{k}\right) \cdot U_{coll}^{k} + \left(T_{fluid_{in}}^{k} - T_{coll}^{k}\right) \cdot U_{fluid}^{k} \left(\dot{m}_{fluid}^{k} > 0\right) \right) \cdot \frac{dt}{C_{coll}}$$

The node temperature T_{PV}, here without thermal capacitance, is obtained by:

$$T_{PV}^{k+1} = \frac{\dot{Q}_{abs}^{k+1} + U_{front}^{k+1} T_{amb}^{k+1} + U_{coll}^{k+1} T_{coll}^{k+1} + T_{sky}^{k+1} U_{rad}^{k+1}}{U_{front}^{k+1} + U_{coll}^{k+1} + U_{rad}^{k+1}}$$

where:

- $$U_{rad}^{k+1} = A_{PVT}\,\epsilon_{PVT}\,\sigma \cdot \left(T_{PV}^{k+1} + T_0 + T_{sky}^{k+1} + T_0\right) \cdot \left(\left(T_{PV}^{k+1} + T_0\right)^2 + \left(T_{sky}^{k+1} + T_0\right)^2 \right)$$
- $\dot{Q}_{abs}^{k+1} = \left(1 - \eta_{PV}^{k+1}\right) F_{th} A_{PVT} I_{gt}^{k+1}$
- $\eta_{PV}^{k+1} = \eta_{PVref} \cdot \left(1 + f_{corr_T} \cdot \left(T_{PV}^{k+1} - T_{PVref}\right)\right) \cdot \left(1 + f_{corr_I} \cdot \left(I_{gt}^{k+1} - I_{gtref}\right)\right)$

Table 3
Values and range (for identification) of the (normally) fixed parameters (left) and of the identified parameters (right).

Fixed parameters	Value and unit		Potential variation range	Identified parameters	Reference start value	Range for identification	Unit
$T_{PV_{ref}}$	25	°C	A_{PVT}	1.8	±50%	m^2	
$I_{gt_{ref}}$	1000	W/m^2	ϵ_{PVT}	0.15	±50%	–	
f_{corr_T}	−0.004	$°C^{-1}$	$\eta_{PV_{ref}}$	0.15	±50%	–	
f_{corr_I}	3.60E-5	m^2/W	F_{th}	1	[0.5 1.2]	–	
Latitude	47.358	°	U_{front}	10	$[10^{-3}\ 10^3]$	$W/(m^2K)$	
Longitude	16.128	°	U_{back}	10	$[10^{-3}\ 10^3]$	$W/(m^2K)$	
ρ_{gnd}	0.2		U_{coll}	10	$[10^{-3}\ 10^3]$	$W/(m^2K)$	±50%
b_0	0.05		r_{0fluid}	0.1	$[0\ 10^3]$	$W/(m^2K)$	±50%
β	20	°	r_{1fluid}	1	$[10^{-2}\ 10^2]$	$W/(m^2K)$	±10%
γ	180	°	C_{PV}	0	[0 0]	J/m^2K	±20%
$c_{p_{fluid}}$	3650	J/(kg.K)	C_{coll}	10^5	$[10^3\ 10^7]$	J/m^2K	±20%
ρ_{fluid}	1040	kg/m^3	$r_{ref_{wind}}$	1	$[10^{-2}\ 10^2]$	$\frac{m^2K}{W}/\left(\frac{m}{s}\right)^{n_{wind}}$	
			n_{wind}	0.6	[0.3 1.2]	–	

Table 4
Results of parameter identification for different training period durations in terms of MAE, averaged over the same 29 forecast days.

Training period duration (days)		7 days	14 days	21 days	28 days
Average calibration runtime (s)		79	126	236	272
Average Mean Absolute Error of collector output powers	$\overline{MAE\_P_{el}}(W/m^2)$	2.07	2. 02	2.02	2.01
	$\overline{MAE\_\dot{Q}_u}(W/m^2)$	19.5	17.2	17.6	17.7

2.10. Alternative with one node:

In the simplified RC model with a single temperature node T_{PV}, the absorbed solar power is delivered at this node and the thermal power is divided between the useful heat transferred to the fluid and the losses to the front and to the back side of the collector. The PV temperature is calculated as:

$$T_{PV} = \frac{\dot{Q}_{abs} + U_{back}T_{back} + U_{front}T_{amb} + T_{fluid_{in}}U_{fluid}\left(\dot{m}_{fluid_{in}} > 0\right)}{U_{back} + U_{front} + U_{fluid}\left(\dot{m}_{fluid_{in}} > 0\right)}$$

with the same definitions as in the two-node model.

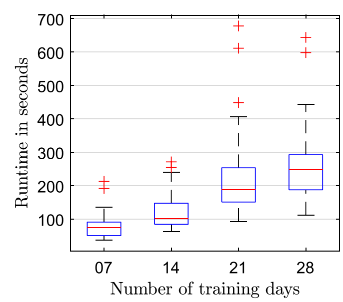

Fig. 5. Calibration runtime for different training durations: 7, 14, 21 and 28 days of training.

Table 5
Time step variation.

Time step of training data		5 s	1 min	5 min	10 min	15 min
Calibration runtime (s)		1403				13.7
Average Mean Absolute Error of the collector outputs	$\overline{MAE\_\dot{Q}_u}(W/m^2)$	20.0	16.9	17.7	20.1	22.7
	$\overline{MAE\_P_{el}}(W/m^2)$	2.14	2.09	2.07	2.05	2.01

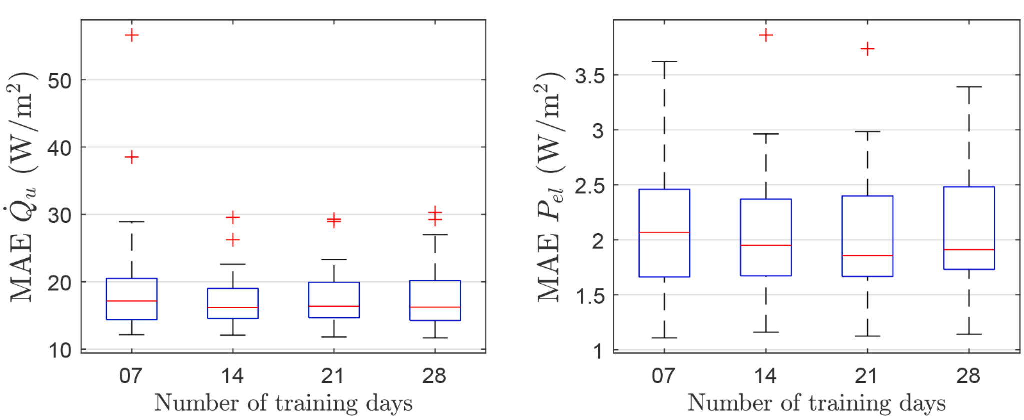

Fig. 4. Box plot including 29 simulations of a daily forecast: Mean Absolute Error of calculated versus measured electrical and thermal output of the collector, for different training durations: 7, 14, 21 or 28 days.

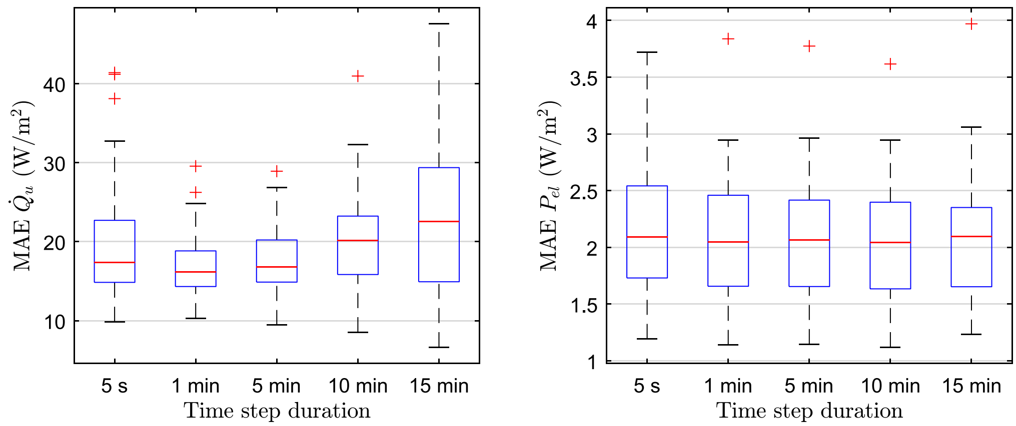

Fig. 6. Mean Absolute Error of electrical power P_{el} and thermal power $\dot{Q}_u$ for different time step durations.

3. Data and method for parameter identification

3.1. Experimental data from a PVT collector test bench

A PVT collector has been tested under real operating conditions. The photos in Fig. 2 show the collector test rig used and, at a distance of 65 m, the solar tracker of the weather station, all mounted on roofs. The facility is located in Austria, at 47° latitude and 400 m over sea level, in a temperate climate zone.

The experimental setup is presented in Fig. 3 with a schematic diagram of the hydraulic circuit that collects the thermal energy from the PVT collector. The thermal energy is transferred to a buffer tank by activating a pump. Different control strategies have been tested: the variable mass flow rate is used in this study. The heat transfer fluid is a mixture of demineralised water with 43 % mono-propylene-glycol by mass. The PVT collector produces the optimum electrical power through a Maximum Power Point Tracking (MPPT) controller.

The collector is mounted on a biaxial sun-tracker, on the roof of the laboratory hall. The position of the collector is defined by two angles: tilt (from horizontal) and the azimuth (from North, clockwise). The position can be easily varied or even tracked automatically throughout the day. To validate the model under the conditions of the most common applications, data with fixed collector position are used in the following.

The radiative equivalent sky temperature T_{sky_m} (°C) is derived from measurements (therefore the subscript m) of a SGR4 pyrgeometer (manufacturer KIPP&ZONEN). This equipment provides directly the downward longwave irradiance $I_{ir_{sensor1}}$ (W/m^2). The temperature T_{sky_m} is calculated as:

$$T_{sky_m} = \left(\frac{I_{ir_{sensor1}}}{\sigma}\right)^{\frac{1}{4}} - 273.15$$

where $\sigma = 5.67 \times 10^{-8} W/(m^2K^4)$ Stefan-Bolzmann constant.

All measured variables used for parameter identification are listed in Table 1 for collector data (see also Fig. 3) and Table 2 for weather data. The data acquisition system is automatically recording measurement points every 5 s, which are pre-processed by averaging into minute values. The sensors and variables of interest correspond, on the one hand, to the input variables of the model, and on the other hand, to the main output variables –calibration variables–, namely: the electrical power P_{el} and the thermal power $\dot{Q}_u$ of the collector. The temperature of PV absorber T_{PV} is an additional optional calibration variable. The model is calibrated by fitting the calculated values to the measured values. In this small installation, the time delay between inlet and outlet is neglected.

The dataset used in this study contains 57 days, over summer months from July 11th to September 6th 2021. To have up to 28 training days, the forecast should start on August 9th. The test of a 24-hour forecast can then be applied to a row of 29 days. More details on the experiment and on the preparation of the dataset can be found via the repository, where the dataset itself is available [27].

3.2. Parameter identification procedure

3.2.1. Calibration algorithm

For the data-driven parameter identification of the model, the non-linear least squares data fitting method is used: *lsqnonlin* algorithm [23]. The following options are selected:

- *FunctionTolerance* $= 1.10^{-10}$
- *StepTolerance* $= 1.10^{-10}$
- *MaxFunctionEvaluations* $= 500$
- *MaxIterations* $= 500$

In most cases, the function tolerance is reached. In cases where the convergence is more difficult, the algorithm stops after reaching the maximum function evaluation limit. In such cases, simulations allowing 5000 evaluations (instead of 500) do not bring better results.

3.2.2. Objective function and quality criteria for validation

The quality of the calibrated model is assessed by the Mean Absolute Error *MAE*:

$$MAE = MAE\_P_{el} + MAE\_\dot{Q}_u = \overline{|P_{elc} - P_{elm}|} + \overline{\left|\dot{Q}_{uc} - \dot{Q}_{um}\right|}$$

where P_{el_c} and P_{el_m} are the calculated and measured electrical powers respectively,

$\dot{Q}_{u_c}$ and $\dot{Q}_{u_m}$ being the calculated and measured useful thermal power respectively.

As the algorithm used is a non-linear least-squares solver (*lsqnonlin*), it calculates the norm of the objective function output. To obtain the *MAE* as output, the square root of the difference is calculated in the objective function $F(X)$, which is an array of $N_{training}$ lines defined as follows:

$$F_i(X) = \left[\sqrt{|P_{elc}(t_i) - P_{elm}(t_i)|} \quad \sqrt{\left|\dot{Q}_{uc}(t_i) - \dot{Q}_{um}(t_i)\right|} \right]$$

where i is the index of time step t_i, from 1 to $N_{training}$, the number of time steps in the training data set, and X is the set of parameters,

The algorithm minimises the norm of the objective function $\|F(X)\|$:

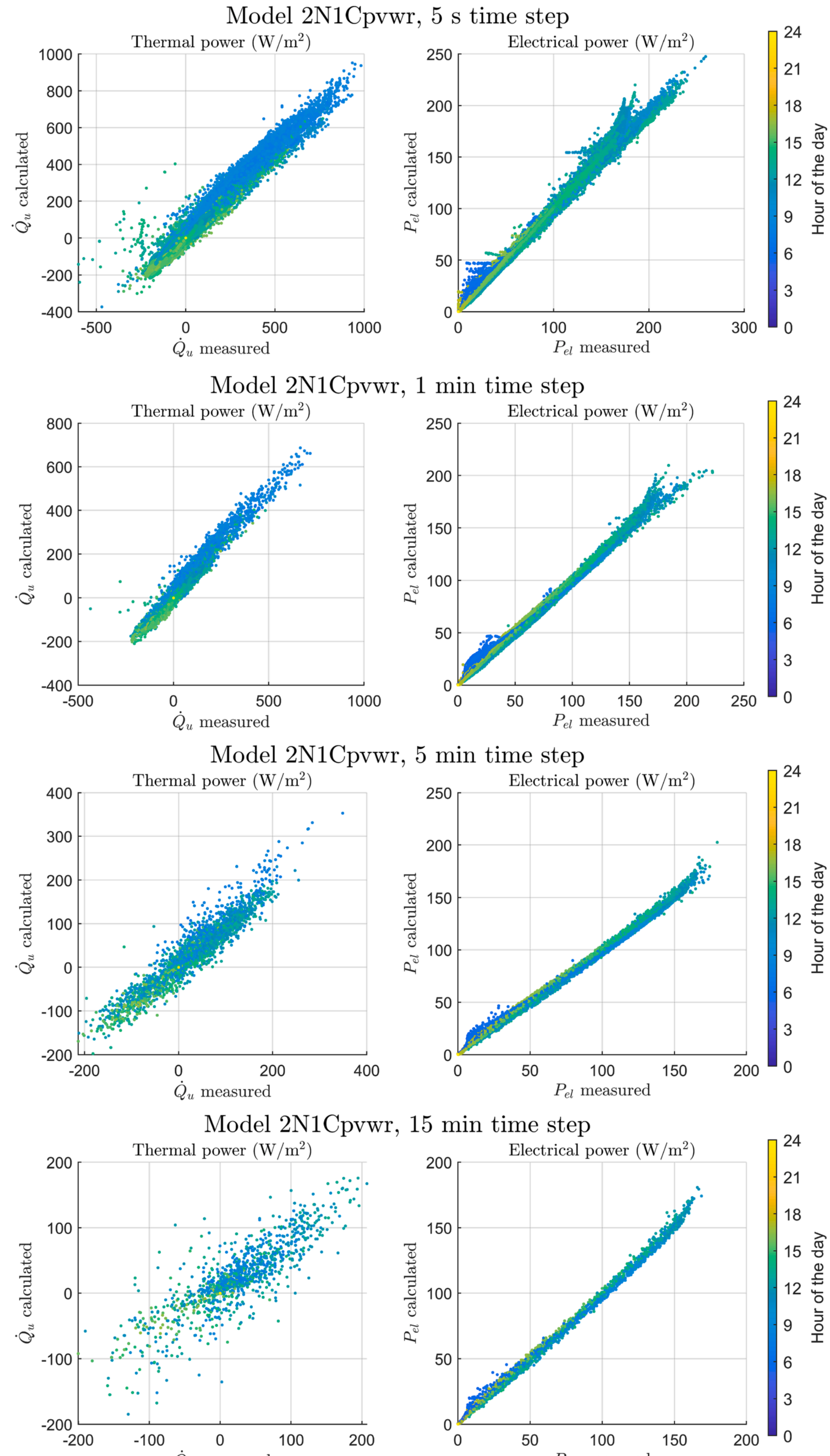

Fig. 7. Scatter distribution of calculated versus measured points, for different time steps: 5 s, 1 min, 5 min and 15 min. The colour scale represents the hour of the day.

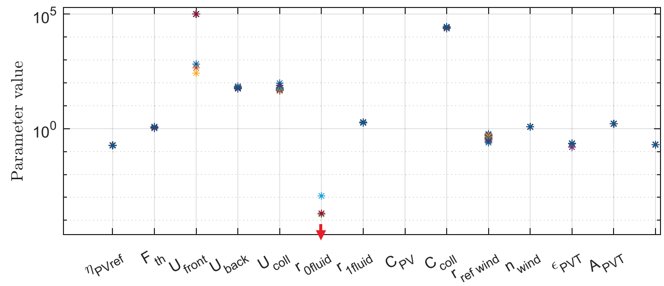

Fig. 8. Overview of identified parameters for model version 2N1C$_{col}$wr (further values of r_{0fluid} are down to 2.4×10^{-14}).

$$\|F(X)\| = \sum_{i=1}^{N}\left(\sqrt{|P_{elc}(t_i) - P_{elm}(t_i)|^2} + \sqrt{\left|\dot{Q}_{uc}(t_i) - \dot{Q}_{um}(t_i)\right|^2}\right)$$

$$= \sum_{i=1}^{N}\left(|P_{elc}(t) - P_{elm}(t)| + \left|\dot{Q}_{uc}(t) - \dot{Q}_{um}(t)\right|\right)$$

which is proportional to the *MAE*:

$$MAE = \frac{\|F(X)\|}{N}$$

To complement the analysis, a normalised Mean Absolute Error *nMAE* is defined to assess the error in relation to the order of magnitude of the respective variables P_{el} and $\dot{Q}_u$:

$$nMAE\_P_{el} = \frac{MAE\_P_{el}}{\overline{|P_{el}|}} \text{ and } nMAE\_\dot{Q}_u = \frac{MAE_{\dot{Q}_u}}{\overline{\left|\dot{Q}_u\right|}}$$

where $\overline{|P_{el}|}$ is the average of absolute values taken by the variable P_{el} over the forecast time period (24 h), similarly for $\dot{Q}_u$.

For a visual interpretation, the corresponding frequency distribution histograms and scatter plots are analysed. The runtime of the algorithm is also given.

3.2.3. List of parameters: Selection for identification

The RC models are tuned with the following parameters. The set of parameters consists of fixed and identified parameters, adjusted after parameter variation using the synthetic data set. The fixed parameters of the installation are summarised in Table 3 (left). They are defined from the knowledge of the collector under study. These parameters are physically well determined or they do not play a decisive role for the convergence of the model. For sensitivity analysis or validation purposes, these parameters can also be moved to the identified parameters.

The identified parameters are determined by the calibration algorithm. They are shown in Table 3 (right). These parameters are not known a priori and are specific to each system. They are given an initial value and a certain range in which to search for the actual value. These parameters are normalised as far as possible. This harmonizes the magnitude of the different parameters, for faster calibration. Some fixed parameters can be released as additional identified parameters: their potential variation range is indicated in the list of fixed parameters in Table 3 (left).

3.3. Evaluation of the results in a real application scenario

The intended application of the model is optimized control, using an adaptive model with data-driven parameter identification. The model is calibrated regularly, e.g. daily, based on current monitoring data. The period of training for the calibration can be chosen as a compromise between model accuracy and calibration effort (amount of data, parameter identification runtime).

To evaluate the quality of the approach and to compare different parameters and variants, the parameter identification is tested on a series of 29 days. For each day *d*, a parameter identification is performed using the last *n* days as training period. The *n* values tested are 7, 14, 21 and 28 days. The identified model is then used to predict the energy output of the collector on that day *d*, following the training period. The MAE is evaluated for each day and averaged over the 29 test days.

4. Results analysis

4.1. Preliminary analysis

These results are presented for the best model variant –two-node, one-capacitance, with wind and sky losses (short name: *2N1C$_{col}$wr*)– as determined by the analysis in section 4.2.

4.1.1. Variation of the training period

The following Table 4 shows the results obtained with different training data periods: 7, 14, 21 and 28 days. The same 29 forecast days are used, starting from 9 August 2021. While the table shows average values, Fig. 4 shows the box plots of the MAE on the electrical and thermal collector outputs. The results show that even a short training

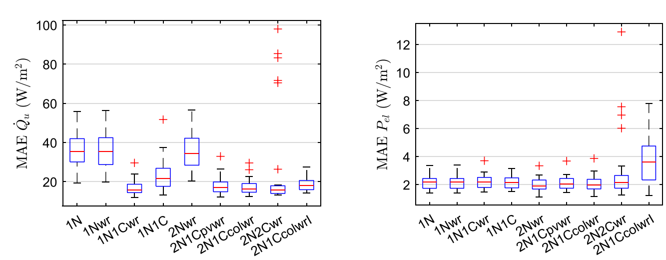

Fig. 9. Mean Absolute Error of calculated versus measured electrical and thermal collector output for the one-node and two node model variants.

Table 6
Parameter identification of the grey-box model for the thermal behaviour in one node 1 N and two nodes 2 N variants: Variant names and definitions. Main results: Mean Absolute Error and normalised Mean Absolute Error of collector energy production (heat and electricity), averaged over the 29 test days, and calibration runtime.

Model name	Variant characteristics						$\overline{MAE\_\dot{Q}_u}$(W/m^2)	$\overline{MAE\_P_{el}}$(W/m^2)	$\overline{nMAE\_\dot{Q}_u}$(%)	$\overline{nMAE\_P_{el}}$(%)	Runtime (s)
	Node	C_{pv}	C_{coll}	wind	sky rad	$I_b + I_d$					
1 N	1						35.8	2.17	38.9	3.55	35
1Nwr	1			✔	✔		35.6	2.17	38.6	3.55	26
1N1C	1	✔					23.0	2.16	25.4	3.56	19
1N1Cwr *	1	✔		✔	✔		17.0	2.19	18.3	3.60	30
2Nwr	1			✔	✔		35.5	1.98	38.4	3.18	93
2N1C$_{pv}$wr	2	✔		✔	✔		18.0	2.10	19.3	3.43	39
2N1C$_{col}$wr *	2		✔	✔	✔		17.2	2.04	18.5	3.28	75
2N1C$_{col}$wrI *	2		✔	✔	✔	✔	18.8	3.66	20.7	5.62	49
2N2Cwr	2	✔	✔	✔	✔		27.0	2.91	31.0	4.54	63

* these are the three best variants, selected for a more detailed analysis in section 4.3.

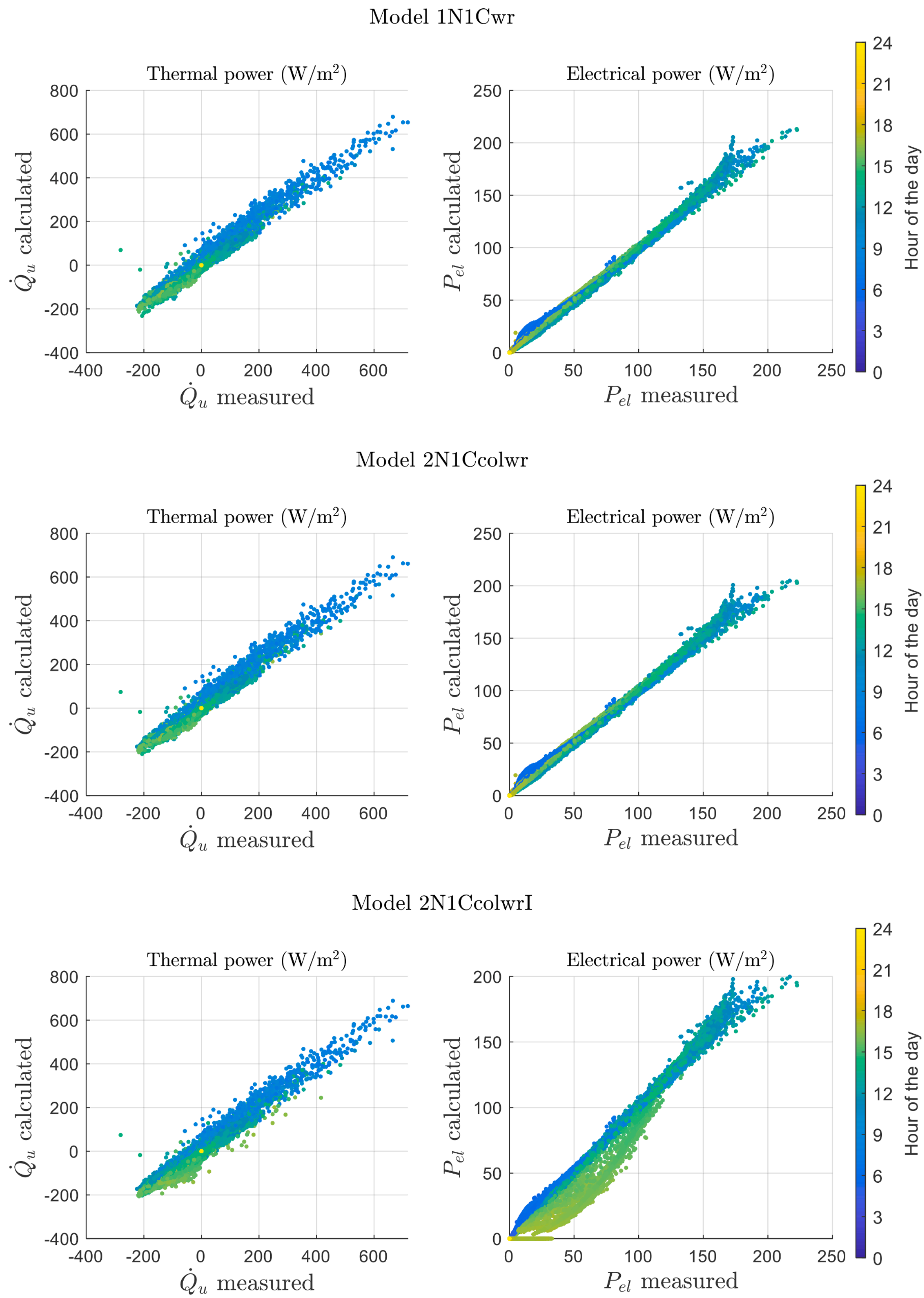

Fig. 10. Scatter plots of the calculated versus measured thermal and electrical powers for variants 1N1Cwr, $2N1C_{col}wr$ and $2N1C_{col}wrI$. The colour scale represents the hour of the day.

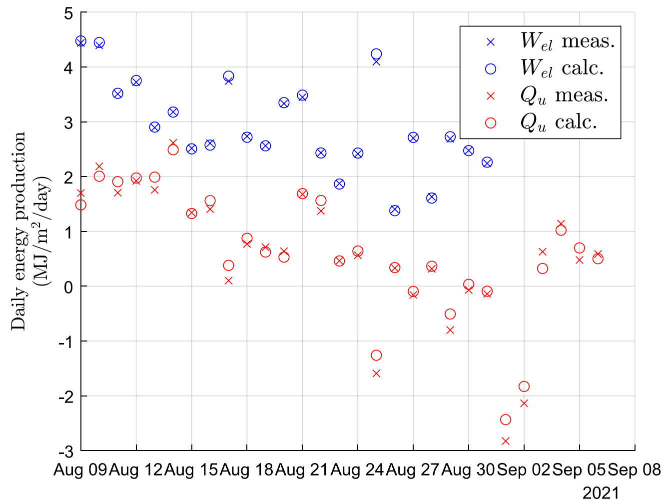

Fig. 11. Daily thermal and electrical energy production of the PVT collector. Comparison of measured values (x) and values calculated with the $2N1C_{col}wr$ model (o).

period of 7 days achieves a good parameter identification. The best results are obtained with a training period of 14 days, which improves the Mean Absolute Error on the thermal power $\overline{MAE\_\dot{Q}_u}$ by 10 %, compared to 7 days training. A longer training period does not further improve the results. The calibration runtime also increases almost linearly with the length of the training period (see Fig. 5). A training period of 14 days is chosen for the following work.

4.1.2. Variation of the time step

The time step of the monitoring data, used for parameter identification, is important to correctly capture the dynamic behaviour of the system. Different time steps have been tested: 5 s, 1, 5, 10 and 15 min. Table 5 summarises the average results over the 29 test days. The best compromise is 1 min, where the Mean Absolute Error on the thermal side is the lowest. This is emphasised in Fig. 6 with the box plots of the MAE on the electrical and thermal powers for the different time steps. This can also be seen in Fig. 7, on the scatter plot of the calculated versus measured data points. With a longer time step, the dynamic behaviour of the model is not correctly captured. The 5 s time step (original raw data) does not improve the accuracy of the thermal and electrical power outputs. Some short-term deviations from the measurement system could explain the distortions that occur when using this short time step. This could also explain a slightly better fit of the electrical power for longer time steps, which are less affected by the short-term dynamics of the collector thermal behaviour. 1 min is therefore a good compromise in terms of model accuracy and sophistication of the monitoring system (reasonable data logging frequency and amount of data to store). The one-minute time step is used to compare the model variants in the following sections 4.2 and 4.3.

4.1.3. Values of the identified parameters

The values of the identified parameters are shown in the graph in Fig. 8, over the 29 test days. Most of the parameters are stable in the identification procedure: the variations of the parameters remain in the same order of magnitude. Only the heat transfer coefficient for the losses to the front of the collector U_{front} varies considerably ($2.6 \times 10^2 to 10^5$). The thermal resistance from the pipes to the fluid r_{0fluid} is always close to zero (1.2×10^{-3}, down to 2.4×10^{-14}).

4.2. Comparison of the models

All models have been tested using the same 29 days. For each day, a forecast is made using a model, trained on the 14 days prior to that day, using one-minute time step data.

Several variants of the one-node (*1 N*) and two-node (*2 N*) models have been tested. The grey-box model options are indicated by their *names in italics*:

r for the long-wave radiative transfer with the sky,
w for the influence of wind convection,
I for the model of irradiance split into direct and diffuse parts (instead of global irradiance in the tilted collector plane).
1C or *2C* for the single or double capacitance of the collector, respectively, applied to one or both nodes (*1C$_{pv}$* or *1C$_{col}$* to avoid ambiguity),

The energy balance of the collector is most important for an optimised control application. The results are presented in terms of normalised Mean Absolute Error *nMAE*, for both thermal and electrical energy production. Table 6 gives an overview of the different variants: description and main results. Fig. 9 shows the box plots of the MAE of the nine variants.

In the model variants, the importance of the thermal capacity of the collector is clearly visible: the three variants without thermal capacitance (*1 N*, *1Nwr* and *2Nwr*) have a high average MAE of 35 W/m^2 on the thermal power, twice as much as similar variants with thermal capacitance (*1N1Cwr*, *2N1C$_{pv}$wr* and *2N1C$_{col}$wr*). At the same time, the two-node, two-capacitance model (*2N2Cwr*) underperforms, because the error is abnormally high on some days (see Fig. 9).

In the two-node model, the variant with thermal capacitance C_{coll} (*2N1C$_{col}$wr*) gives slightly better results than the one with C_{pv} (*2N1C$_{pv}$wr*): $\overline{MAE\_P_{el}}$ -3 % and $\overline{MAE\_\dot{Q}_u}$ -5 %. The two-node model *2N1C$_{col}$wr* performs best overall with $\overline{MAE\_P_{el}} = 2.0\,\mathrm{W/m^2}$ and $\overline{MAE\_\dot{Q}_u} = 17\,\mathrm{W/m^2}$.

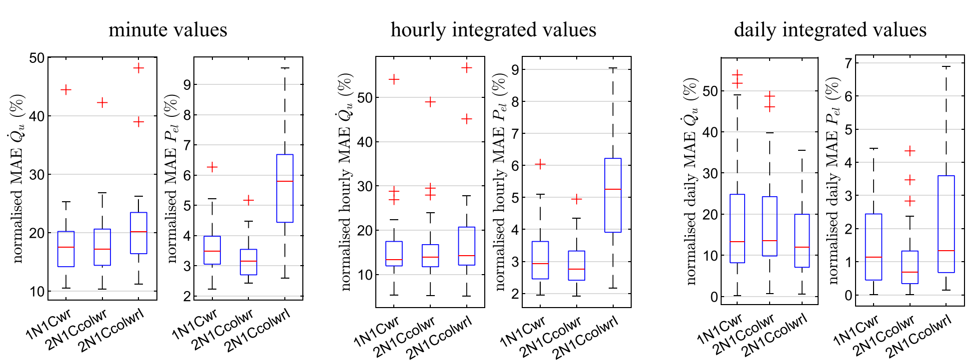

Fig. 12. Box plots of normalised Mean Absolute Error for thermal $\dot{Q}_u$ and electrical P_{el} powers with different time resolutions: 1 min (left), 1 h (centre) and 1 day (right). For the daily values, extreme points are not visible on the $MAE\_\dot{Q}_u$ diagram: these correspond to days with thermal energy outputs very close to zero, where the relative error is large (up to 300 %), but the absolute error remains small.

As the average runtime for the $2N1C_{col}wr$ is of 75 s, the one-node model *1N1Cwr* is also an alternative to consider: with even 1 % lower $\overline{MAE\_\dot{Q}_u}$ and only 7 % higher $\overline{MAE\_P_{el}}$.

It can be noted that the wind losses (*w*) and the radiative losses (*r*) improve the model, especially on the thermal side: 26 % lower MAE for *1N1Cwr* than for *1N1C*. Without C_{pv}, these two options (*wr*) do not make a difference in the quality of the results: *1 N* and *1Nwr* are similar.

If the global irradiance in the collector plane is not measured directly, it must be calculated from the global horizontal irradiance and the diffuse horizontal irradiance. The model that calculates the global irradiance from its direct and diffuse components (*I*) gives slightly lower quality in results: the variant $2N1C_{col}wrI$ has $\overline{MAE\_P_{el}} = 3.7\,\mathrm{W}/m^2$ and $\overline{MAE\_\dot{Q}_u} = 18.8\,\mathrm{W}/m^2$, respectively 80 % and 9 % more than $2N1C_{col}wr$. Although the electrical power error is almost twice as large, it remains small compared to the collector power output. The normalised MAE values confirm this: $\overline{nMAE\_\dot{Q}_u} = 20.7\,\%$ and $\overline{nMAE\_P_{el}} = 5.6\,\%$ for $2N1C_{col}wrI$ compared to 18.5 % and 3.3 % respectively for $2N1C_{col}wr$. Thus, even if the global irradiance in the collector plane were not available, the quality of the model would remain acceptable.

4.3. Best model variant

The following figures show the detailed results of the three best models: *1N1Cwr*, $2N1C_{col}wr$ and $2N1C_{col}wrI$. The relationship between measured and calculated variables is shown on the scatter plot in Fig. 10. These plots illustrate the correct modelling of thermal and electrical power. The irradiance model (*I*) in $2N1C_{col}wrI$ causes small biases, especially in the morning (green dots).

Fig. 11 shows the daily energy output of the collector for the 29 test days, comparing measured and calculated values with model variant $2N1C_{col}wr$. It can be seen, that the PVT collector under test, produces little thermal energy and even loses thermal energy on some days. This uncovered PVT collector is operated at its highest temperature limit, with inlet temperatures of 40 to 50 °C. As other collectors are connected to the same circuit, the pump of the solar circuit remains activated even at times when the collector cannot deliver thermal energy. This explains the negative thermal output on some days. In any situation of thermal output, the good fit can be seen in Fig. 11.

Fig. 12 gives a quantified assessment of the fit. To see if errors in the one-minute time step values lead to errors at the hourly and daily scales, the power measurements and calculations have been integrated. This shows how short-term errors in the output powers, propagate over longer time periods in the output power evaluation. The box plots therefore show the normalised MAE at different time resolutions: minute values, hourly and daily integrated values. The red lines represent the median values, which are more representative than the mean values, because of the rare events with high normalised MAE that occur on the thermal energy side, when the power is low compared to the absolute error.

The best overall performance is achieved by $2N1C_{col}wr$, with median normalised MAE of 17.2 % on $\dot{Q}_u$ and 3.2 % on P_{el} for minute values, 13.9 % and 2.8 % respectively for hourly integrated values and 13.6 % and 0.70 % respectively for daily integrated values. As expected, the electrical model is more accurate than the thermal model, because it is also simpler. The thermal model captures the thermal behaviour with reasonable accuracy, even at a one-minute time step. The thermal output at hourly and daily levels is therefore also well represented.

It can be concluded that the model can be successfully used to assess the energy production of the collector. It can be used for optimised control, with a forecast horizon of 24 h, with an accuracy of 2.5 to 4.5 % for the electrical energy production and 10–27 % for the power output with a time step of one minute.

5. Conclusion and perspective

Several variants of a grey-box model of a PVT collector have been developed and validated with measured data from a real collector. The models perform well with one or two temperature nodes, as long as it includes a single thermal capacitance. This proves that the modelling of the collector thermal capacity is essential in real dynamic operating conditions. Convective wind losses and long-wave radiation to the sky are also important parts of the model. Over a 24-hour forecast horizon with a one-minute time step, the expected accuracy (MAE) is of 2 to 5 % on the electrical power and 10 to 27 % on the thermal power, depending on the day. The two-node model performs slightly better than the one-node model. However, the advantage of the one-node model is a 60 % shorter runtime. If solar irradiance at the collector plane cannot be measured or forecasted, an additional algorithm was used to convert the global and diffuse horizontal irradiance measurements to direct and diffuse irradiance at the collector plane. This conversion algorithm degrades the performance of the collector model, compared to direct measurement in the collector plane. The effect is especially on the electrical side: 2.5 to 9.5 % higher MAE on electrical power and 11 to 27 % higher MAE on thermal power, depending on the day.

The investigations show also that a relatively short training period of two weeks can provide satisfactory results for the grey-box model parameter identification. Furthermore, using a one-minute time step, derived from 5-second measurements, offers a favourable trade-off by addressing short-term deviations and improving the representation of dynamic thermal behaviour.

Since parameter identification is based on sensor data, the grey-box approach can be applied to PVT systems of any size and with any PVT configuration or design. The proposed model is not intended to improve directly the collector design. It rather focuses on representing the behaviour of any existing system. Therefore, the implementation of the best proposed model in a data-driven optimal control strategy promises to increase the overall energy efficiency for various PVT integration scenarios. One typical application, where model predictive control already revealed improvement potential, is for systems combining a PVT with a heat pump, especially ground source heat pump. However, further development and refinement, including robustness testing with diverse data sets and improved irradiance modelling, will be crucial for advancing the effectiveness and applicability of the model within predictive control strategies for buildings.

Funding

This work was elaborated within the PVadapt project and has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 818342.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] K. Kramer, Status Quo of PVT Characterization, IEA SHC Task 60, 2020. https://doi.org/10.18777/ieashc-task60-2020-0004.

[2] F. Hengel, C. Heschl, F. Inschlag, P. Klanatsky, System efficiency of pvt collector-driven heat pumps, Int. J. Thermofluids. 5–6 (2020), 100034, https://doi.org/10.1016/j.ijft.2020.100034.

[3] M. Lämmle, M. Herrando, G. Ryan, Basic concepts of PVT collector technologies, applications and markets IEA SHC Task 60/Report D5, 2021. https://doi.org/10.18777/ieashc-task60-2020-0002.

[4] A. Herez, H. El Hage, T. Lemenand, M. Ramadan, M. Khaled, Review on photovoltaic/thermal hybrid solar collectors: classifications, applications and new systems, Sol. Energy 207 (2020) 1321–1347, https://doi.org/10.1016/j.solener.2020.07.062.

[5] S.M. Sultan, M.N. Ervina Efzan, Review on recent photovoltaic/thermal (PV/T) technology advances and applications, Sol. Energy 173 (2018) 939–954, https://doi.org/10.1016/j.solener.2018.08.032.

[6] S. Kumar Verma, R. Kumar, M. Barthwal, D. Rakshit, A review on futuristic aspects of hybrid photo-voltaic thermal systems (PV/T) in solar energy utilization: engineering and Technological approaches, Sustain. Energy Technol. Assess. 53 (2022), 102463, https://doi.org/10.1016/j.seta.2022.102463.

[7] A. Kasaeian, G. Nouri, P. Ranjbaran, D. Wen, Solar collectors and photovoltaics as combined heat and power systems: a critical review, Energ. Conver. Manage. 156 (2018) 688–705, https://doi.org/10.1016/j.enconman.2017.11.064.

[8] A. Kathirgamanathan, M. De Rosa, E. Mangina, D.P. Finn, Data-driven predictive control for unlocking building energy flexibility: a review, Renew. Sustain. Energy Rev. 135 (2021), 110120, https://doi.org/10.1016/j.rser.2020.110120.

[9] Z. Liu, Z. Guo, Q. Chen, C. Song, W. Shang, M. Yuan, H. Zhang, A review of data-driven smart building-integrated photovoltaic systems: challenges and objectives, Energy 263 (2023), 126082, https://doi.org/10.1016/j.energy.2022.126082.

[10] T. You, W. Wu, H. Yang, J. Liu, X. Li, Hybrid photovoltaic/thermal and ground source heat pump: review and perspective, Renew. Sustain. Energy Rev. 151 (2021), 111569, https://doi.org/10.1016/j.rser.2021.111569.

[11] E. Entchev, L. Yang, M. Ghorab, A. Rosato, S. Sibilio, Energy, economic and environmental performance simulation of a hybrid renewable microgeneration system with neural network predictive control, Alex. Eng. J. 57 (2018) 455–473, https://doi.org/10.1016/j.aej.2016.09.001.

[12] L. Xia, Z. Ma, G. Kokogiannakis, S. Wang, X. Gong, A model-based optimal control strategy for ground source heat pump systems with integrated solar photovoltaic thermal collectors, Appl. Energy 228 (2018) 1399–1412, https://doi.org/10.1016/j.apenergy.2018.07.026.

[13] J.J. Michael, I. S, R. Goic, Flat plate solar photovoltaic–thermal (PV/T) systems: a reference guide, Renew. Sustain. Energy Rev. 51 (2015) 62–88.

[14] M. Dannemand, I. Sifnaios, Z. Tian, S. Furbo, Simulation and optimization of a hybrid unglazed solar photovoltaic-thermal collector and heat pump system with two storage tanks, Energ. Conver. Manage. 206 (2020), 112429, https://doi.org/10.1016/j.enconman.2019.112429.

[15] L. Ouhsiane, M. Siroux, M.E. Ganaoui, A. Mimet, Multi-Objective Optimization of Hybrid PVT Solar Panels, in: In: 2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE), 2018, pp. 1–5.

[16] S. Miglani, K. Orehounig, J. Carmeliet, Integrating a thermal model of ground source heat pumps and solar regeneration within building energy system optimization, Appl. Energy 218 (2018) 78–94, https://doi.org/10.1016/j.apenergy.2018.02.173.

[17] W. Fan, G. Kokogiannakis, Z. Ma, A multi-objective design optimisation strategy for hybrid photovoltaic thermal collector (PVT)-solar air heater (SAH) systems with fins, Sol. Energy 163 (2018) 315–328, https://doi.org/10.1016/j.solener.2018.02.014.

[18] L.W. Florschuetz, Extension of the Hottel-Whillier model to the analysis of combined photovoltaic/thermal flat plate collectors, Sol. Energy 22 (1979) 361–366, https://doi.org/10.1016/0038-092X(79)90190-7.

[19] TRNSYS, Welcome | TRNSYS : Transient System Simulation Tool, (2021). http://www.trnsys.com/ (accessed February 21, 2023).

[20] H. Helmers, K. Kramer, Multi-linear performance model for hybrid (C)PVT solar collectors, Sol. Energy 92 (2013) 313–322, https://doi.org/10.1016/j.solener.2013.03.003.

[21] D. Jonas, M. Lämmle, D. Theis, S. Schneider, G. Frey, Performance modeling of PVT collectors: implementation, validation and parameter identification approach using TRNSYS, Sol. Energy 193 (2019) 51–64, https://doi.org/10.1016/j.solener.2019.09.047.

[22] EN ISO 9806, ÖNORM EN ISO 9806 Edition: 2018-05-01 Solar energy: Solar thermal collectors: Test methods, (2018). https://shop.austrian-standards.at/action/en/public/details/631245/OENORM_EN_ISO_9806_2018_05_01.

[23] MATLAB, MATLAB - MathWorks, (2021). https://nl.mathworks.com/products/matlab.html (accessed November 29, 2022).

[24] M.J. Brandemuehl, W.A. Beckman, Transmission of diffuse radiation through CPC and flat plate collector glazings, Sol. Energy 24 (1980) 511–513, https://doi.org/10.1016/0038-092X(80)90320-5.

[25] A.F. Souka, H.H. Safwat, Determination of the optimum orientations for the double-exposure, flat-plate collector and its reflectors, Sol. Energy 10 (1966) 170–174, https://doi.org/10.1016/0038-092X(66)90004-1.

[26] M. Koschenz, B. Lehmann, Thermoaktive Bauteilsysteme tabs, Selbstverlag, Dübendorf (Schweiz), 2000 https://www.irb.fraunhofer.de/dokumentlieferdienst/?id=2004059018082.

[27] F. Veynandt, F. Inschlag, C. Seidl, C. Heschl, Measurement dataset from real operation of a hybrid photovoltaic-thermal solar collectors, used for the development of a data-driven model. https://doi.org/10.57739/3720.