Advances in Applied Energy 2 (2021) 100020

Contents lists available at ScienceDirect

Advances in Applied Energy

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

Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties

Loris Amabile[a,b,*], Delphine Bresch-Pietri[b], Gilbert El Hajje[a], Sébastien Labbé[a], Nicolas Petit[b]

[a] EDF R&D, TREE, Avenue des Renardières - Ecuelles, Moret-Loing-et-Orvanne 77250 France
[b] Centre Automatique et Systèmes, MINES ParisTech, PSL, 60 bd Saint-Michel, Paris 75006 France

ARTICLE INFO

Keywords:
Intelligent control of power systems
Optimal operation and control of power systems
Smart grids
PV power production forecasts
Microgrid optimal control
Demand response

ABSTRACT

An energy management system optimizes the self-consumption of a residential photovoltaic installation, and the performance losses due to production uncertainties are evaluated. The specific case under study is an individual home equipped with photovoltaic (PV) panels where only an Electric Water Heater (EWH) is manipulated, and the rest of the appliances represent a fixed load. By formulating the problem of maximizing self-consumption as an unconstrained optimization problem, a novel and computationally efficient optimization algorithm has been proposed. The next step was to numerically evaluate the performance of this EWH management strategy under various PV power production scenarios, generated through a presented methodology. The reference baseline is a rule-based controller using a most likely forecast of PV production. Simulations performed in Dymola over 10 months demonstrate that, at a 30-minute timestep, the impact of a "perfect" PV production forecast is negligible compared with the impact of the choice of the control algorithm. Besides, a most likely forecast is good enough for the proposed algorithm to reach high self-consumption levels. Indeed, although the proposed optimization based on a most likely forecast yields an increase of 10 points of self-consumption compared to the baseline, only an additional 2 points of increase can be reached using "perfect" production information.

1. Introduction

1.1. Self-consumption in individual houses

The need for decarbonized energy as a response to the climate crisis has stimulated an international effort to increase the installed capacity of Renewable Energy Sources, which is reflected in various forms of governments support and unprecedented production cost reductions [1]. In many places, Distributed Energy Resources are expected to contain the costs and faults associated with increasing electrical consumption and extending electrical grids. Both these trends explain the rapid growth of worldwide installed capacities of solar power in particular, which is the cheapest and easiest-to-install distributed renewable energy source. Growing renewable generation induces higher variability in supply. Hence, production and consumption flexibilities become increasingly valuable for the transmission and distribution system operators as substitutes for substantial grid investments [2].

Consumption flexibility is the ability to temporally shift or to modify the power consumption of electrical appliances. Self-consumption is one regulatory framework intended to bolster these consumption flexibilities, by promoting local consumption over export. Self-consumption contributes to the distribution grid stability by avoiding voltage rise during peak photovoltaic (PV) generation periods and helps to reach higher shares of installed PV in the electricity mix [3]. As simplified administrative procedures [4] and new tariffs have made self-consumption of generated power a profitable option, these installations are increasingly encouraged and have become popular. They foster local smart management of electrical appliances because fully benefiting from the self-consumption financial regulatory incentives requires optimal management of household consumptions. Moreover, the value of these Energy Management Systems (EMSs) used to monitor, control, and optimize the household consumptions is enhanced by the increasing electricity bills of retail consumers. This trend is due to the electrification of energy consumptions traditionally supplied by fossil fuels.

1.2. Appliances and electric water heater control

To supply maximum flexibility, combined PV-battery setups are sometimes considered, but still represent a prohibitive overall additional cost both for individually owned [5,6] and for shared battery sys-

* Corresponding author.
E-mail addresses: loris.amabile@mines-paristech.fr (L. Amabile), delphine.bresch-pietri@mines-paristech.fr (D. Bresch-Pietri), gilbert.el-hajje@edf.fr (G. El Hajje), sebastien.labbe@edf.fr (S. Labbé), nicolas.petit@mines-paristech.fr (N. Petit).

https://doi.org/10.1016/j.adapen.2021.100020
Received 15 March 2021; Accepted 15 March 2021
Available online 18 March 2021

Table 1
Nomenclature.

Notation	Description	Unit
E	Thermal energy of the water tank	Wh
E_{clear}	EWH energy level giving clearance to resume heating	Wh
E_{sat}	EWH saturation energy level	Wh
E_{PV}	Energy produced by the PV arrays	Wh
K	Number of stochastic PV scenarios	
k	Thermal loss coefficient of the water tank	h^{-1}
$2M$	Deadband around the setpoint temperature	°C
N	Index of last EWH clearance time resulting in an activation of the heating power	
P_{ewh}	EWH power consumption	W
$\bar{P}_{\text{ewh}}$	EWH power rating	W
P_{PV}	Surplus of produced PV power	W
Q	Instantaneous power drained from the EWH by hot-water consumption	W
T_{ewh}	Mean water temperature in the water tank	°C
T_{set}	Setpoint temperature	°C
t_{ewh}	EWH starting time of the heating command	h
t_{clear}	EWH clearance time	h
t_{sat}	EWH saturation time	h
t_{lim}	Latest acceptable starting time of the heating range	h
τ_{ewh}	EWH ending time	h
SC	Self-consumption over one period	Wh
S^j	PV production scenarios for day j. Indices: ML (most likely) or $k \in [0, K]$ (realization)	
$s_{\text{P}}(\cdot)$	EWH heating schedule computed by the proposed algorithm	
$s_{\text{R}}(\cdot)$	EWH heating schedule computed by the reference method	

tems [7]. As part of investigations into demand response (i.e., the short-term modification of consumer demand for energy according to external signals [8]), heating, ventilation and air-conditioning (HVAC) systems have attracted wide interest in the recent decades, both for residential buildings [9,10] and office buildings [11]. However, home HVAC control is not adopted in this study because the flexibility provided by these systems can be fairly low due to the need to maintain the desired indoor comfort. To further extend the scope of appliances subject to automatic control, efforts have been made to propose generic models of electrical appliances with several energy phases [12,13]. These models can be used for white goods control, but individually they represent low energy consumptions. With 12% of all European domestic electric consumption dedicated to water heating in 2018 [14], electric water heating appliances (with and without storage) represent a non-negligible share of residential electricity bills. A type of appliance of particular interest is Electric Water Heaters (EWHs), which provide storage capability in thermal form. With their low investment costs, EWHs have the benefit of being already widespread in various countries, with 57 million units installed in 2014 representing 23% of the total European primary water heater stock [15]. Hence, they represent a highly suitable appliance to control in the context of PV self-consumption optimization [16–21]. This paper focuses exclusively on the control of EWHs.

1.3. Production and consumption uncertainties

Numerous self-consumption optimization studies have reported impressive figures. However they suffer from requiring perfect knowledge of user behavior [18] and of the weather forecast (e.g. [17], which uses the same production scenario for the forecast, for the decision algorithm and for performance evaluation of the chosen control), which appears to be an important factor. Indeed, renewable generation is subject to weather variability, leading to a limit on production forecast precision [22]. Robust optimization [13,23] can be used to tackle these uncertainties, but these approaches often lead to considering the least favorable scenarios and therefore proposing too-conservative strategies.

Alternatively, Model Predictive Control (MPC) enables weather forecast updates at each prediction step, thus improving performance [13,17,18,24–26]. This improvement is even more significant when stochastic MPC [11,27] or Stochastic optimal control [21] is used, but comes at the price of potentially burdensome computation time.

1.4. Contribution

This paper focuses on PV self-consumption optimization for an individual home through sole control of the EWH. The reference situation is a rule-based basic heuristic representative of the current state-of-the-art controller in commercial products, which relies on a unique deterministic PV production forecast, imperfect but considered as the most likely prediction. Considering this class of EMS as a baseline, this paper investigates the performance gains provided by an efficient optimization strategy and the impact of imperfect PV production forecasts on an optimal control.

To conduct this investigation, a novel optimal EWH control algorithm was derived, and an original method was used to generate a vast class of PV power production scenarios.

First, the optimization problem was formulated in an unconstrained form involving a most likely PV production prediction. By modeling the EWH a an homogeneous-temperature tank equipped with a thermostat, the dynamics at stake can be explicitly resolved, leading to a simple and efficient optimal EWH control algorithm. This strategy is less general than MPC approaches, which could have been used in this context, but reveals less computationally intensive and requires no parameter tuning. Hence, because the computing time of the proposed algorithm is extremely low, various instances can be studied in a short time.

The next step is to assess the impact of 30-minute PV production forecast uncertainties on this optimization algorithm. Thorey et al. [28] provide a methodology to obtain PV load factor quantile forecasts. Building on these quantile forecasts, a novel method was developed in this study to generate PV production scenarios that collectively satisfy a given probabilistic distribution while individually presenting realistic intra-day variability, that is, a realistic correlation factor.

It was shown in this study that, beyond a given level of forecast accuracy, the impact of knowing in advance the exact PV production forecast is negligible compared with the impact of the choice of the EWH control algorithm. This is the main contribution of the paper.

1.5. Paper organization

Section 2 details the modeling of appliances in the setup considered here, the optimization problem at hand, and the core of the proposed algorithm. Section 3 presents a methodology to generate a set of correlated realizations of a stochastic process, the PV production scenarios,

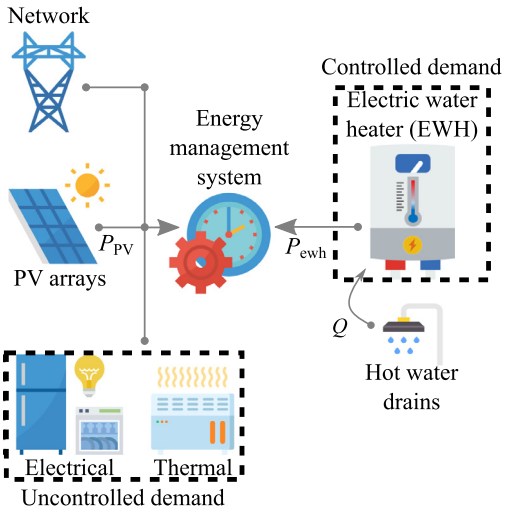

Fig. 1. Setup under consideration: the Energy Management System regulates the EWH.

which are representative of the estimated uncertainties of most likely PV production forecasts. Section 4 presents the numerical experiments conducted to validate the overall methodology and then assesses the comparative values of the "perfect" PV production forecast and the optimal control algorithm. Section 5 reports conclusions and perspectives.

2. EWH energy management system

2.1. Setup

The general setup under consideration is depicted in Fig. 1 and consists of an individual house connected to the electrical grid. It is equipped with PV arrays, heating and cooling equipment, an EWH, and other electrical appliances such as lights, a fridge, and a dishwasher.

The heating and cooling systems are assumed to be entirely dedicated to maintaining house thermal equilibrium. Therefore, these appliances are considered as uncontrolled and are not associated with any decision variables in the remainder of this work.

In this paper, the only controllable appliance to be considered is the EWH.

To solve the optimization problem, the proposed algorithm considers to have an exact prediction of the following elements:

- PV production curve;
- Electrical consumption profile of uncontrollable loads;
- Hot water drains from the EWH.

Note that in the final evaluation of the proposed EMS performance, the imperfect nature of PV production forecasts will be taken into account.

2.2. Electric water heater modeling and control

The EWH is modeled as a homogeneous temperature volume. Leaving aside the modeling of the temperature stratification inside the tank implies a loss of precision, but simplifies the control design [29]. Considering an initial energy state at time t_0, thermal losses, the power input from resistive heating, and the hot-water consumption during the given period, the energy stored at a time t_2 in the EWH, knowing the energy E_1 stored at a time t_1, is given by

$$E(t_2) = E_1 e^{-k(t_2-t_1)} + \int_{t_1}^{t_2} e^{-k(t_2-t)}\big[P_{\text{ewh}}(s,t) - Q(t)\big]dt \tag{1}$$

where

- $E(t)$ is the energy stored in the EWH at time t;
- k is the thermal loss coefficient [29];
- $P_{\text{ewh}}(s,t)$ is the electrical power input of the EWH at time t, following heating strategy s;
- $Q(t)$ is the energy used at time t by a hot-water drain.

In this work, the EWH is assumed to be equipped with an internal thermostat that is adjusted on a setpoint temperature T_{set} and a deadband $2M$. Hence, the internal thermostat stops any heating command when the measured temperature reaches $T_{\text{set}} + M$, and enables heating to resume when the temperature drops below $T_{\text{set}} - M$. Because the temperature inside the EWH is considered homogeneous, these specific temperatures $T_{\text{set}} + M$ and $T_{\text{set}} - M$ correspond to energy levels, denoted as E_{sat} and E_{clear}.

Correspondingly, the power input of the EWH can be written as

$$P_{\text{ewh}}(s,t) = \big(1 - \delta_{\text{sat}}(t)\big)\overset{\circ}{P}_{\text{ewh}}(s,t) \tag{2}$$

where s is the control strategy, t is the time, P°_{ewh} is the chosen authorized heating power, and δ_{sat} is the hysteresis function describing the saturation state of the tank, equal to 1 if saturation is occurring and 0 otherwise:

$$\delta_{\text{sat}}(t) = \begin{cases} 0 & \text{if } E(t) < E_{\text{clear}} \\ 1 & \text{if } E(t) \geq E_{\text{sat}} \\ \delta_{\text{sat}}(t^-) & \text{if } E(t) \in [E_{\text{clear}}, E_{\text{sat}}[\end{cases} \tag{3}$$

with t^- denoting the time directly preceding t.

The variable δ_{sat} in Eq. (2) describes the behavior of the EWH internal thermostat, allowing heating to remain activated if the water temperature has not yet reached the saturation threshold or to resume if it has dropped below the clearance temperature.

The function P°_{ewh} is defined as follows. The EWH under consideration has a power consumption of either 0 or $\bar{P}_{\text{ewh}}$, its constant power rating, triggered by an On/Off authorization mechanism. In this work, to simplify the heating control strategy, this authorization is chosen to be On only during a unique and continuous period $s = [t_{\text{ewh}}, \tau_{\text{ewh}}]$ determined by the EMS, which corresponds to choosing a strategy s. In detail,

$$P^{\circ}_{\text{ewh}}(s,t) = \bar{P}_{\text{ewh}}\,\mathbb{1}_{[t_{\text{ewh}},\tau_{\text{ewh}}]}(t) \tag{4}$$

in which $\mathbb{1}_{\text{I}}$ is the indicator function of the time interval I. Inside this authorized period, the EWH internal thermostat might start and stop effective power activation according to Eq. (3).

2.3. Optimization problem formulation

2.3.1. Objective function and decision variables

This study focuses on optimizing the self-consumption (SC) because it is a key evaluation criterion for such installations. It is usually considered as a proxy for reducing electricity bills, and high SC levels can be critical in locations where the state of the distribution grid cannot allow large volumes of PV surplus to be injected.

SC is defined as the part of local photovoltaic production that is locally consumed to meet electric consumption (see Fig. 4). Maximizing it is equivalent (see [30] for details) to maximizing the positive part of the remaining PV production, which is obtained after subtracting the uncontrollable electric consumption (heating and cooling, lights, white goods). Observing from Eqs. (2)–(4) that $P_{\text{ewh}}(s,t)$ is null when $t \notin [t_{\text{ewh}}, \tau_{\text{ewh}}]$, the mathematical definition of the SC becomes the integral over a period $[t_{\text{ewh}}, \tau_{\text{ewh}}]$ of the minimum between the positive part of the surplus of local production P_{PV}, and the controllable load (see Fig. 2):

$$SC(s, P_{\text{PV}}) = \int_{t_{\text{ewh}}}^{\tau_{\text{ewh}}} \min(P_{\text{ewh}}(s,t), P_{\text{PV}}(t))dt \tag{5}$$

where $s = (t_{\text{ewh}}, \tau_{\text{ewh}}) \in \mathbb{R}^2$ is the decision variable and $P_{\text{PV}}(t) \geq 0, \forall t$.

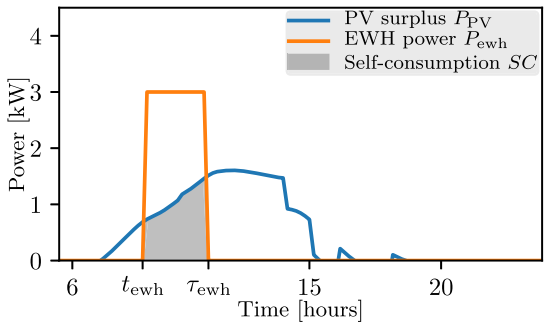

Fig. 2. Optimizing SC, which is the controllable part of the self-consumption, corresponds to maximizing the size of the gray area.

2.3.2. Constraints

The optimal control problem considers the following boundary and control constraints:

$$E(t_0) = E_0 \tag{6}$$

$$E(s, t_f) \in [E_{clear}, E_{sat}] \tag{7}$$

where t_0 is a generic starting time at which the energy level is known.

Eq. (7) is a comfort constraint: the control shall guarantee that the EWH reaches an energy interval at a final time t_f. Here, this energy interval is chosen to be centered around the energy level corresponding to a hot-water tank and the final time is the beginning of the evening. Enforcing this constraint ensures that the tank has stored enough energy to cover the hot-water consumption of the end of the day.

2.3.3. Reducing the number of variables

Once the energy reaches the interval $[E_{clear}, E_{sat}]$, it remains in it as long as heating is allowed by Eq. (4). Hence, the constraint (7) naturally leads to the choice $\tau_{ewh} = t_f$. τ_{ewh} is no longer a decision variable, and this makes it possible to replace s with only the starting time t_{ewh}.

Furthermore, taking Eq. (6) into account, one obtains the existence of an upper bound t_{lim} for t_{ewh}. According to (1)–(4), it is defined as the unique solution of the equation

$$E_{clear} = e^{-k(t_f - t_0)} E_0 + \int_{t_{lim}}^{t_f} e^{-k(t_f - t)} [\bar{P}_{ewh} - Q(t)] dt \tag{8}$$

that is, the latest acceptable heating starting time to reach E_{clear} at $\tau_{ewh} = t_f$.

Combining these two transformations makes it possible to replace the constraints (6)-(7) with the simple inequality constraint $t_{ewh} \leq t_{lim}$.

As a result, after the EWH reaches the interval $[E_{clear}, E_{sat}]$, the internal energy will keep on oscillating inside it. Fig. 3 gives an example of such an evolution for one day and a specific starting time.

This behavior defines (i) saturation times, when the EWH internal thermostat detects a temperature reaching the allowed upper limit and halts the heating signal, and (ii) clearance times, when thermal losses or hot-water consumption have led the internal temperature to drop below the lower hysteresis value, thus making heating resume. These times are respectively denoted by $t_{sat(n+1)}$ and $t_{clear(n)}$, $n \in \mathbb{N}$ and are defined mathematically as the unique solutions to

$$E_{sat} = e^{-k(t_{sat(n+1)} - t_{clear(n)})} E_{clear} + \int_{t_{clear(n)}}^{t_{sat(n+1)}} e^{-k(t_{sat(n+1)} - t)} [\bar{P}_{ewh} - Q(t)] dt \tag{9}$$

$$E_{clear} = e^{-k(t_{clear(n)} - t_{sat(n)})} E_{sat} - \int_{t_{sat(n)}}^{t_{clear(n)}} e^{-k(t_{clear(n)} - t)} Q(t) dt \tag{10}$$

with the initialization $t_{clear(0)} = t_{ewh}$. Note that N is the index of the last clearance time $t_{clear(n)}$ before t_f, that is, such that $t_{clear(N)} \leq t_f < t_{clear(N+1)}$.

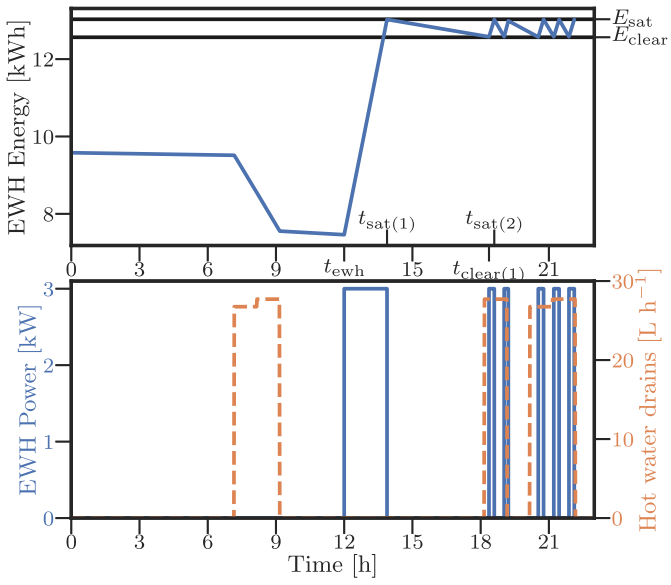

Fig. 3. Hot-water drains (lower figure, dashed curve) decrease the EWH energy (top figure). EWH heating intervals $[t_{clear(n)}, t_{sat(n+1)}]$ (lower figure, plain curve), calculated from a chosen t_{ewh}, increase the EWH energy and keep it oscillating between the deadband $[E_{clear}, E_{sat}]$.

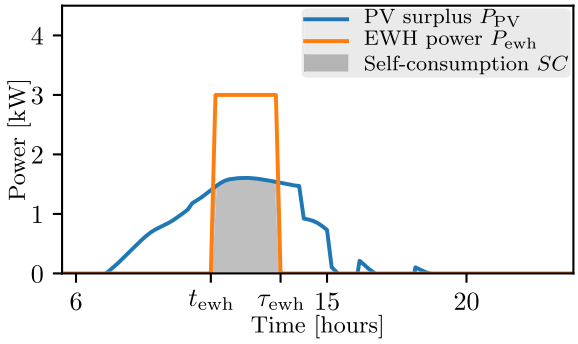

Fig. 4. The optimized placement of the EWH heating period maximizes the gray area, i.e. the objective SC.

Therefore, the EWH consumption curve can be represented by a sum of non-overlapping boxcar functions corresponding to the intervals where $\delta_{sat}(t) = 0$: the EWH is On between $[t_{ewh}, t_{sat(1)}]$ and between each $[t_{clear(n)}, t_{sat(n+1)}]$ pair, with the last interval potentially shortened to $[t_{clear(N)}, t_f]$ if $t_f < t_{sat(N+1)}$.

Note that the hot-water drains depicted by orange dashed lines in the lower part of Fig. 3 have a visible impact on the internal EWH energy (upper part).

2.3.4. Problem statement

From previous considerations and as illustrated in [30], the problem of optimizing self-consumption can be summarized as:

Problem 1. Given a PV production curve P_{PV} and a final time t_f, solve for each day

$$\max_{t_{ewh} \leq t_{lim}} \int_{t_{ewh}}^{t_f} \min(P_{ewh}(t_{ewh}, t), P_{PV}(t)) dt \tag{11}$$

where t_{lim} is defined in (8) and P_{ewh} is defined as the sum of boxcar functions

$$P_{ewh}(t_{ewh}, t) = \bar{P}_{ewh} \sum_{n=0}^{N} \mathbb{1}_{[t_{clear(n)}, \min(t_{sat(n+1)}, t_f)]}(t)$$

with $t_{\text{clear}(0)} = t_{\text{ewh}}$ and $t_{\text{clear}(n)}$ and $t_{\text{sat}(n+1)}$ depending on t_{ewh} and defined in (9)–(10) along with N.

2.4. Optimization algorithm

Problem 1 is non-concave. To solve it, this study proposed to identify the critical points of its objective function.

2.4.1. Smoothness analysis of the objective function

Following steps similar to those used in [30], one can prove that SC is continuously differentiable with respect to t_{ewh} (the proof is omitted for brevity). Furthermore, with $t_{\text{clear}(0)} = t_{\text{ewh}}$, its derivative is given as

$$\frac{dSC(t_{\text{ewh}}, P_{\text{PV}})}{dt_{\text{ewh}}} = \sum_{n=0}^{N} \left[\frac{dt_{\text{sat}(n+1)}}{dt_{\text{ewh}}} \min(P_{\text{ewh}}(t_{\text{ewh}}, t_{\text{sat}(n+1)}), P_{\text{PV}}(t_{\text{sat}(n+1)})) - \frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}} \min(P_{\text{ewh}}(t_{\text{ewh}}, t_{\text{clear}(n)}), P_{\text{PV}}(t_{\text{clear}(n)})) \right] \quad (12)$$

if $t_{\text{sat}(N+1)} \leq t_{\text{f}}$, and a similar expression otherwise, omitted for brevity.

2.4.2. Algorithm description

Because the problem is non-concave, it is necessary to assess the nature of every stationary point. The stationary condition can be written as

$$\frac{dSC(t_{\text{ewh}}, P_{\text{PV}})}{dt_{\text{ewh}}} = 0 \quad (13)$$

Corresponding SC values are then evaluated and compared to solve the optimization problem. Typically, for the examples treated in the paper, fewer than ten extrema need to be considered.

A key advantage of this approach is that evaluating this stationary condition through Eq. (12) is much less computationally costly than evaluating the integral in Eq. (11).

2.4.3. Implementation

The specific times $t_{\text{sat}(n+1)}$ and $t_{\text{clear}(n)}$, $n \in \mathbb{N}$, as well as their derivatives, must be computed because they are involved in the expression of the derivative of the objective function in Eq. (12), but cannot be explicitly determined from their definitions (9)–(10). Besides, the derivatives of these times, $\frac{dt_{\text{sat}(n+1)}}{dt_{\text{ewh}}}$ and $\frac{dt_{\text{clear}(n)}}{dt_{\text{ewh}}}$, also appear in (12) and involve $t_{\text{sat}(n+1)}$ and $t_{\text{clear}(n)}$. The following discussion will clarify how these terms are evaluated numerically.

Simplified description of the calculation of the t_{sat} times For the EWH saturation times $t_{\text{sat}(n+1)}$, a fixed-point algorithm is used. For $n = 0$, one cannot deduce an explicit expression for $t_{\text{sat}(1)}$ from (9) due to the presence of the integral term corresponding to the hot-water drains. However, one can rewrite (9) for $n = 0$ as $t_{\text{sat}(1)} = g_1(t_{\text{ewh}}, t_{\text{sat}(1)})$ with

$$g_1(t_{\text{ewh}}, t_{\text{sat}(1)}) = \frac{1}{k} \left[\log(\bar{P}_{\text{ewh}} e^{kt_{\text{ewh}}} - kE_0 e^{kt_0}) - \log\left(\bar{P}_{\text{ewh}} - kE_{\text{sat}} - k \int_{t_0}^{t_{\text{sat}(1)}} e^{-k(t_{\text{sat}(1)}-t)} Q(t) dt \right) \right] \quad (14)$$

which can be shown to be a contractive mapping with respect to $t_{\text{sat}(1)}$ for the parameter values under consideration in this study. Hence, fixed-point iterations were used to estimate $t_{\text{sat}(1)}$ numerically, with successive estimates of $t_{\text{sat}(1)}$ computed through $g_1(t_{\text{ewh}}, t_{\text{sat}(1)})$ and progressively reaching this fixed point. Usually, fewer than one dozen iterations are needed to reach an accurate value.

Similarly, $t_{\text{sat}(n+1)}$, for $n > 0$, satisfies a relation of the type $t_{\text{sat}(n+1)} = g_2(t_{\text{clear}(n)}, t_{\text{sat}(n+1)})$ with g_2 a contractive mapping, and this can be calculated with fixed-point iterations as well. Note that g_2 requires a prior calculation of $t_{\text{clear}(n)}$, which will now be described in detail.

Simplified description of the calculation of the t_{clear} times The $t_{clear(n)}$ times satisfy a similar fixed-point equation; however, the resulting mapping is not contractive. Hence, a dichotomy procedure was followed instead to evaluate $t_{clear(n)}$. Indeed, it can be proven that the energy obtained from Eq. (10) between a saturation time $t_{\text{sat}(n)}$ and a clearance time $t_{\text{clear}(n)}$, $n \in \mathbb{N}^*$ is a decreasing function of $t_{\text{clear}(n)}$, ensuring the convergence of the method.

Hence, in a nutshell, the proposed procedure computes alternatively $t_{\text{sat}(n+1)}$ through a fixed-point procedure and $t_{\text{clear}(n)}$ with a dichotomy algorithm and repeats this process iteratively until t_{f} is reached.

Algorithm The solutions to the stationary condition were exhaustively determined, and the corresponding values of the objective function were then compared to solve Problem 1. This procedure is summarized in Algorithm 1. Note that this algorithm is of course intended to be used over a finite time grid. Specifically steps 1 and 8 consider a finite number of elements, and a solution of Eq. (13) is identified in step 3 when the objective function derivative (12) goes from a positive to a non-positive value between two consecutive timesteps. This algorithm can now be applied in a high-fidelity environment.

Algorithm 1 Calculate $t^{\star}_{\text{ewh}}$ solution of Problem 1.

Requires: $P_{\text{PV}}, \bar{P}_{\text{ewh}}, t_0, E_0, t_{\text{f}}, E_{\text{f}}, E_{\text{sat}}, E_{\text{clear}}, Q, I = [0, t_{\text{lim}}]$
1: **for all** $t \in I$ **do**
2: Compute times = [every $t_{\text{sat}(n)}$ and $t_{\text{clear}(n)}$ up to t_{f}].
3: **if** times are solution of Eq. (13), **then**
4: Compute $SC(t, P_{\text{PV}})$
5: Update Buffer = [Buffer, $SC(t, P_{\text{PV}})$]
6: **end if**
7: **end for**
8: Identify $t^{\star}_{\text{ewh}} = \text{argmax}_t$ Buffer

Fig. 4 displays an example of the result of Algorithm 1, with the EWH heating period chosen so that the SC will be maximized.

3. Generation of stochastic PV production scenarios

Now that an optimization strategy for self-consumption has been proposed, it is desirable to evaluate the impact of a discrepancy between the unique deterministic production forecast used in the Energy Management System and the actual realized production. To the authors' knowledge, only a few studies in the literature have numerically evaluated this kind of impact on their EMS, among them Thomas et al. [31] and Rabiee et al. [32]. However, in both cases, only a few scenarios were considered, which cannot account for potential intra-day variability.

The following discussion describes in detail the methodology designed to generate a representative set of PV production scenarios.

3.1. Context

Starting from a unique deterministic forecast named here *most likely*, a set of PV production scenarios is needed to illustrate the uncertainties linked to the initial deterministic forecast. The set as a whole shall represent with proportionality the range of possible production outcomes, and each individual scenario shall present a realistic intra-day variability.

Associated with weather-based power models, ensemble forecasts provided by meteorological services could serve this purpose. Indeed, they give an indication of the range of possible future states of the atmosphere. However, only a limited number of possible outcomes are produced. In addition, their standard timestep is several hours, hence providing low resolution. Therefore, these data were discarded as a data source for our study.

Instead, the uncertainties of deterministic PV load factor forecasts can be described by quantile forecasts. In Thorey et al. [28], quantile regressions are applied to deterministic PV production forecasts to build

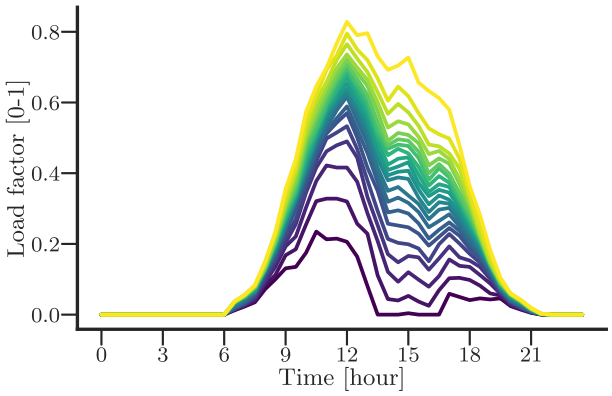

Fig. 5. Quantile values at each timestep associated with the PV production forecast for one day in June with a cloudy afternoon (timezone: UTC + 2). Data generated using the method of [28].

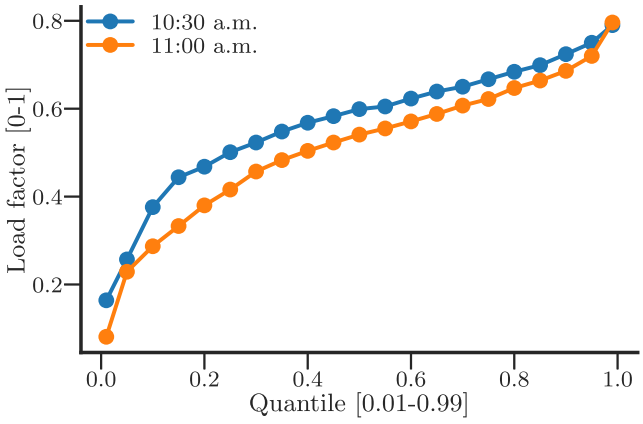

Fig. 6. The cumulative distribution function for two consecutive timesteps for the same day in June do not follow any standard random process *a priori*. The differences between the consecutive curves demonstrate the changing weather conditions.

19 quantile forecasts (of order 5 to 95). Each deterministic PV production forecast itself is the result of a statistical model fed with the following variables from a deterministic weather forecast: solar irradiance, total cloud cover and 2-m temperature. The PV load factor quantile forecasts are the input data to the presented method.

At each 30-minute timestep,[1] for each day, an estimation of the uncertainties is available in the form of quantile values. Fig. 5 displays these quantile curves, with one color for each quantile value. The darkest and lowest curve is the 0.01 quantile curve: for each timestep, there is a 1% chance that the PV load factor will be below the value of this curve. The lightest and highest curve is the 0.99 quantile curve: there is a 99% chance that the PV load factor will be below the value of this curve. The 19 curves in-between correspond to all the quantile values between 0.05 and 0.95, with a 0.05 increment. This method is *a priori* not climate-dependent, but was tested exclusively for an oceanic temperate climate (covering the vast majority of the French continental territory).

Thanks to these quantile values, the cumulative distribution functions of the PV load factor are known for all the consecutive times in a discretized day. Note that each cumulative distribution function obtained by quantile regression as inspired by Thorey et al. [28] is constrained so that the scenario values cannot be greater than the maximum load factor at the given timestamp. This upper limit is obtained assuming a completely clear sky at a given location. Without a means to ascertain the lower limit of the load factor (through a completely cloud-covered sky, for instance), it is simply set to 0 without loss of generality. With these cumulative distribution functions, the next step is to generate numerous PV load factor scenarios.

3.2. Methodology for generating correlated scenarios

Creating the daily scenarios by successively picking one of the values drawn for each timestep according to the cumulative distribution is not a suitable option. Indeed, the generated PV load factor scenarios might not be consistent, with important changes of values from one timestep to another.

Hence, the aim is to introduce some correlation of the PV load factor between two consecutive timesteps.

Unfortunately, the cumulative distribution function obtained by the quantile regression inspired by [28] has no analytical form and does not correspond to any standard random process *a priori* (see Fig. 6). Hence, it is necessary to design an *ad hoc* procedure to introduce correlation between two random variables $\mathbf{X_1}$ and $\mathbf{X_2}$ of fixed and known respective cumulative distributions F_1 and F_2.

The solution proposed here is based on the Probability Integral Transform. This result states that if $\mathbf{X}$ is a continuous random variable with cumulative distribution function F_X, then the random variable $\mathbf{Y} = F_X(\mathbf{X})$ has a uniform distribution on $[0, 1]$.

Let there be two random variables $\mathbf{X_1}$ and $\mathbf{X_2}$, of respective cumulative distribution functions F_1 and F_2. Consider $\mathbf{U_1} = F_1(\mathbf{X_1})$ and $\mathbf{U_2}$, two uniformly distributed random variables on [0,1]. Another random variable is built with the following relation:

$$\mathbf{W} = (1-\alpha)\mathbf{U_1} + \alpha\mathbf{U_2} \tag{15}$$

with $\alpha \in]0, 1[$.

The probability density function f of $\mathbf{W}$ can be calculated as

$$f(x) = \begin{cases} \frac{x}{ab} & \text{if } 0 \le x \le a \\ \frac{1}{b} & \text{if } a \le x \le b \\ \frac{1-x}{ab} & \text{if } b \le x \le 1 \end{cases} \tag{16}$$

with $a = \min\{\alpha, 1-\alpha\}$ and $b = \max\{\alpha, 1-\alpha\}$. Integrating f yields the cumulative distribution function F of $\mathbf{W}$ as follows

$$F(x) = \begin{cases} \frac{x^2}{2ab} & \text{if } 0 \le x \le a \\ \frac{a}{2b} + \frac{x-a}{b} & \text{if } a \le x \le b \\ \frac{a}{2b} + \frac{b-a}{b} + \frac{x-b-\frac{x^2-b^2}{2}}{ab} & \text{if } b \le x \le 1 \end{cases} \tag{17}$$

$F(\mathbf{W})$ is then uniformly distributed according to the Probability Integral Transform. Besides, when the same result is applied again, $F_2^{-1}(F(\mathbf{W}))$ has the same probability density function as $\mathbf{X_2}$. Interestingly, it is correlated with $\mathbf{X_1}$ through Eq. (15), which is the desired result.

In a nutshell, denoting by F_k the cumulative distribution function of the random variable $\mathbf{X_k}$, the stochastic process used to generate the scenarios is

$$\begin{cases} x_{k+1} = F_{k+1}^{-1}(F((1-\alpha)F_k(x_k) + \alpha u_{k+1})), \text{ for } k > 0 \\ x_0 = 0 \end{cases} \tag{18}$$

with $\alpha \in]0, 1[$, F defined in Eq. (17) and u_{k+1} uniformly distributed on [0,1].

The degree of correlation, given by $\frac{Cov(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}}$ (where σ_{x_i} is the standard deviation of x_i), is customizable by varying the value of α. In the simulations in this study, the value of the parameter α was chosen so that the impact of a past random selection fades after a given number of

[1] Note that we consider PV production forecasts with a 30-minute timestep to be possibly exact, thus knowingly disregarding faster variations in solar irradiance, for instance due to passing clouds.

timesteps. The choice of $\alpha = 0.25$ leads to a steady decrease in the correlation between two values x_i and x_j separated by an increasing duration, with a median correlation of 0.45 between two values separated by 5 hours.

With 500 values drawn for each timestep, the original estimated uncertainty quantile values and the dispersion of the scenarios generated by this method match almost perfectly. This limited number of scenarios does not generate any computational burden and thus represents a reasonable compromise in terms of accuracy.

4. Numerical experiments

From the forecast and production data of a PV installation, a set of PV production scenarios was generated according to the method presented in the previous section. Because the data used were not available for two months from the middle of August to the middle of October, the experiments described were conducted over 309 dates only, instead of 365.

Using these scenarios, the impact of PV production uncertainties on the self-consumption performances of two EMSs was investigated in an individual house case study.

4.1. Test scenario

Consider two crystalline silicon PV arrays of 1.5kW$_p$ each, both inclined at 35° relative to the horizontal and facing south. An EWH of volume V =200L and power rating $\bar{P}_{\text{ewh}}$ =3kW is assumed, to cover the consumption of two inhabitants (Table 2). The scenario of outdoor temperature, input cold water temperature, and solar irradiance predictions correspond to a house located at the research and development facility of EDF near Fontainebleau, in the French region of Ile-de-France. It is stressed that though the results of this study stand only for similar weather conditions (the oceanic temperate climate found in most of north-western Europe), the method is not climate-dependent.

The hot-water drain scenario is perfectly known, and therefore the EMS can predict accurately when and how much energy will be consumed from the EWH. This consumption scenario is repeated each day with the same pattern reported in Fig. 7, with drains in the morning (from 7 a.m. to 9 a.m.) and in the evening (from 6 p.m. to 7 p.m. and from 8 p.m. to 10 p.m), ranging between 26.8L h^{-1} and 27.7L h^{-1}, for a total consumption of 956L per week at 40°. Only slight quantity variations occur seasonally, with the same pattern at lesser amplitude in summer.

The final time t_f was set to 6 p.m. to satisfy the main hot-water drains of the day.

Table 2
Case study specifications.

PV capacity	3kW$_p$
EWH Power rating	3kW
EWH volume	200L
Weekly hot-water consumption	956L

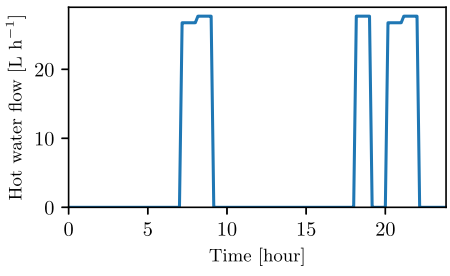

Fig. 7. Hot-water drains for a winter day are distributed between morning and late afternoon consumptions.

4.2. High fidelity model and hardware specifications

The numerical experiments required the controller to communicate with a simulation model. The controller corresponds to the optimization algorithm presented in Section 2.4 based on the simplified EWH model. It is implemented using Python 3.7.3. The simulation model is implemented using the Dymola 2018 software (based on the Modelica 3.2.2 object-oriented modeling language, see [33]). It includes a model for the EWH, and readings of simulated datasets for the house uncontrolled electric loads and for the PV power production. The EWH simulation model is more complex but more accurate than the model described in Section 2.2 for the control algorithm proposed here. It models the stratification phenomenon, with three layers of water constituting the whole volume. The thermal energy dissipated by the resistive element, as well as the incoming cold water, are situated in the lower layer of the tank, whereas the hot-water outlet is situated in the upper layer. As described in [29], at rest, the layers are mixed only by heat diffusion. The effects of this phenomenon are relatively slow compared to the forced convection and mixing induced by draining. Heat losses at the walls of the tank are more precisely accounted for in this simulation model than in the control model.

All computations are run on a Core i3 2.4 GHz processor, with 8 Go RAM. Running the entire calculations took approximately 70 hours for 309 days.

4.3. Performance evaluation procedure

The performances of two EWH control algorithms was compared under two settings regarding PV production uncertainties, resulting in four configurations.

Strategies Two control algorithms can be considered to compute the heating strategy:

- an industrial reference heuristic control;
- the optimization algorithm presented in Section 2.4.

Both are deterministic, meaning that they consider a unique and supposedly exact PV power prediction to decide on the corresponding optimal heating strategy.

The industrial reference control was developed for commercial use in residential housing equipped with PV arrays with the double aim of maximizing consumption of local PV production and ensuring a reduced electricity bill. It follows priority rules to choose each day whether to start the heating authorization according to the following principles:

1. determine when to start the EWH according to a PV surplus threshold set up in advance depending on various criteria;
2. figure out whether heating the tank without PV production is necessary according to whether a daily heating duration target has been met.

Note that the reference control does not require the hot-water drain sequence to perform an estimation of the next heating duration target, whereas the method proposed here requires exact knowledge of it.

Subsequently, s_R denotes a strategy computed with the reference heuristic, whereas s_P denotes a strategy computed with the proposed method.

Scenarios The most likely PV production scenario on day j is denoted by S^j_{ML}. The range of outcomes for the PV production on this day j is represented by the set of K distinct PV production scenarios being generated. The k^{th} realization scenario belonging to this set, with $k \in [0, K]$, is denoted by S^j_k.

We either consider to know only the most likely PV production scenario, or to know the exact PV power predictions. In the first case, $s_R(S^j_{\text{ML}})$ denotes a strategy computed with the reference heuristic, and $s_P(S^j_{\text{ML}})$ denotes a strategy computed with the proposed algorithm, both optimized over S^j_{ML}, the most likely scenario on day j. In the second

Table 3
Scores computed for the proposed algorithm in the conducted numerical simulations. (Crosses indicate scores that were not evaluated.)

Strategies	PV production scenarios for day j					
	Most Likely, S^j_{ML}	Realization 1, S^j_1	...	Realization k, S^j_k	...	Realization K, S^j_K
Most Likely, $s_{\mathrm{P}}(S^j_{\mathrm{ML}})$	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_{\mathrm{ML}}\right)$	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_1\right)$	...	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_k\right)$	...	$SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_K\right)$
Realization 1, $s_{\mathrm{P}}(S^j_1)$	×	$SC\left(s_{\mathrm{P}}(S^j_1), S^j_1\right)$	×	×	×	×
...	×	×	⋱	×	×	×
Realization k, $s_{\mathrm{P}}(S^j_k)$	×	×	×	$SC\left(s_{\mathrm{P}}(S^j_k), S^j_k\right)$	×	×
...	×	×	×	×	⋱	×
Realization K, $s_{\mathrm{P}}(S^j_K)$	×	×	×	×	×	$SC\left(s_{\mathrm{P}}(S^j_K), S^j_K\right)$

case, $s_{\mathrm{R}}(S^j_k)$ denotes a strategy computed with the reference heuristic, and $s_{\mathrm{P}}(S^j_k)$ denotes a strategy computed with the proposed algorithm, both optimized over S^j_k, the realization scenario k of day j. This case corresponds to the availability of what is named a "perfect" PV production forecast.

For a given date j, the performance of the reference control relying on a most likely forecast is the mean SC achieved by this heating strategy evaluated with respect to all the realization scenarios of the generated set. The corresponding mathematical formulation is given in Eq. (19). This configuration is hereafter called Reference - Most Likely. This mean can be considered as an approximation of the expected value of SC for all possible PV production.

Eq. (22) shows the mathematical definition of the mean SC score achievable by the proposed algorithm for a given date j, in the case where a "perfect" forecast is available for every realization scenario of the set. This configuration is hereafter named Proposed - Exact.

Reference - Exact is the mean SC score attainable by the reference heuristic with knowledge of each production realization scenario (Eq. (20)); Proposed - Most Likely is the mean SC score achieved by the heating strategy designed by the proposed algorithm relying only on the most likely forecast (Eq. (21)).

$$\text{Reference - Most Likely: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{19}$$

$$\text{Reference - Exact: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_k), S^j_k\right) \tag{20}$$

$$\text{Proposed - Most Likely: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{P}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{21}$$

$$\text{Proposed - Exact: } \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{P}}(S^j_k), S^j_k\right) \tag{22}$$

Logically, it should hold that (19) < (20), and (21) < (22). However, (20) < (21) is not guaranteed, and quantifying the relation between these values is one of the purpose of this study.

Table 3 details the simulations conducted for the proposed algorithm in this plan of action. Equivalent calculations were conducted for the reference control. Small crosses designate combinations that are *not* associated with calculations.

4.4. Results

The impact of the exact PV load factor predictions and of the chosen algorithm can be seen in Fig. 8. It displays the sum, from January 1st to the end of each month, of the expected daily SC over the set of drawn scenarios $S^j_k, k \in [0, K]$. The control can use either the reference heuristic (orange color) or the proposed tailored algorithm (blue color). The strategy has been optimized either following a most likely PV production forecast (dashed lines) or assuming that the exact PV production forecast is available for each realization (solid lines). The

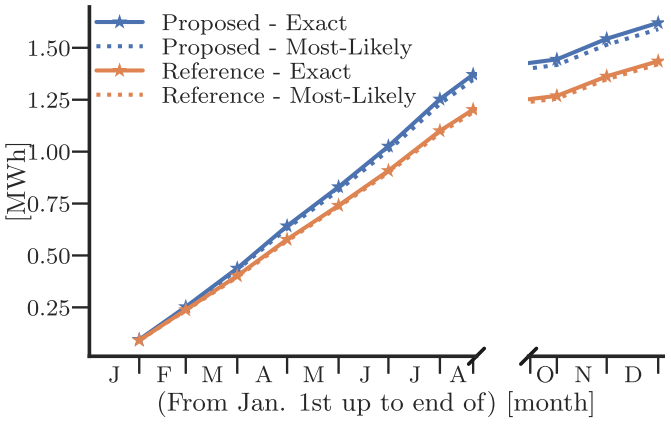

Fig. 8. The cumulated expected SC is significantly more impacted by the algorithm choice than the forecast precision.

following equation gives the corresponding mathematical formulation with $J \in [\text{Jan.}31^{\text{st}}, \ldots, \text{Dec.}31^{\text{st}}]$, the date of each end of month:

$$\sum_{j=\text{Jan.}1^{\text{st}}}^{J}\left[\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S^j_k\right)\right] \tag{23}$$

with X taking the following values in each case:

$$\text{Reference - Most Likely: } X = s_{\mathrm{R}}(S^j_{\mathrm{ML}}) \tag{24}$$

$$\text{Reference - Exact: } X = s_{\mathrm{R}}(S^j_k) \tag{25}$$

$$\text{Proposed - Most Likely: } X = s_{\mathrm{P}}(S^j_{\mathrm{ML}}) \tag{26}$$

$$\text{Proposed - Exact: } X = s_{\mathrm{P}}(S^j_k) \tag{27}$$

The cumulated SC values reached at the end of the year and the corresponding relative increase compared with the Reference - Most Likely case are listed in Table 4.

It is clear that, for the studied setup, the impact of a "perfect" PV production forecast is negligible compared with the impact of the choice of control algorithm. The hypothetical benefit of switching from a most likely PV power prediction to an "exact" one (at a 30-minute timestep) would be slightly greater with the proposed algorithm (+1.97%) than with the Reference heuristic (+0.85%), but remains limited compared with the gains coming from the choice of control algorithm. Note that the SC values consider both the components due to EWH consumption and those due to uncontrolled demand. It is recalled that the most likely forecast already conveys some information about the PV production curves that will finally occur. Shall a "perfect" PV production forecast be compared with a very low level production forecast (e.g. based on clear-sky solar irradiance only), the impact of the forecast precision would be much greater.

Table 4
The cumulated self-consumed energy for 309 days can be increased by 11% with a smarter optimization algorithm.

Strategy	Cumulated Expected SC	Relative increase w.r.t. Reference - Most Likely
Reference - Most Likely (24)	1.42MWh	-
Reference - Exact (25)	1.43MWh	+0.85%
Proposed - Most Likely (26)	1.58MWh	+11%
Proposed - Exact (27)	1.61MWh	+13%

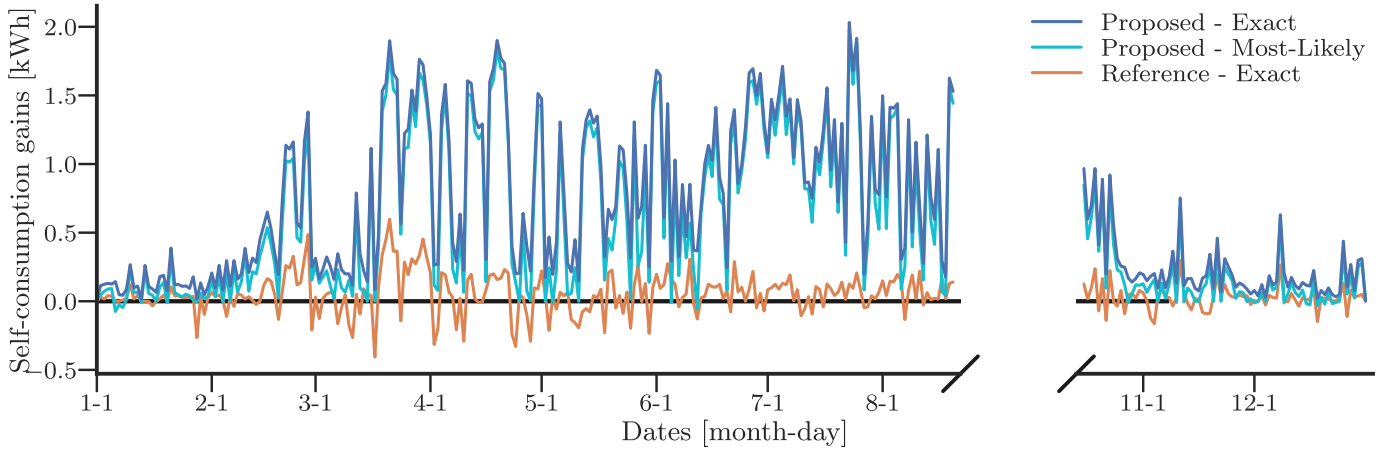

Fig. 9. The daily SC gains for each configuration with respect to the Reference - Most Likely configuration are higher during summer days.

Remark The figures display a gap from the middle of August to the middle of October, because the PV production forecasts were not available for these dates. The "annual" values were hence calculated for 309 days only.

Fig. 9 shows the daily expected self-consumption gains for each configuration compared with the Reference heuristic - Most Likely forecast configuration. For greater clarity, the daily mean Reference - Most Likely SC score on day j is denoted as

$$\mathrm{RML}(j) = \frac{1}{K}\sum_{k=0}^{K} SC\left(s_{\mathrm{R}}(S^j_{\mathrm{ML}}), S^j_k\right) \tag{28}$$

The curves correspond to the following equation, with $j \in$ [Jan.1$^{\mathrm{st}}$, Dec.31$^{\mathrm{st}}$]:

$$\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S^j_k\right) - \mathrm{RML}(j) \tag{29}$$

The equation is evaluated for these three values of X: $s_{\mathrm{R}}(S^j_k)$, $s_{\mathrm{P}}(S^j_{\mathrm{ML}})$, $s_{\mathrm{P}}(S^j_k)$.

The gains are almost always greater with the Proposed algorithm than with the Reference heuristic and Exact PV forecasts. Neither the Reference heuristic nor the Proposed algorithm succeeds in benefiting sensibly from the exact PV forecasts. The ratio of these SC gains over the daily score of the Reference - Most Likely configuration is not depicted because it follows almost the same pattern, with y-values ranging from -10% to +30%.

Fig. 9 suggests that the performance difference depends on the weather conditions of the day (e.g., a sunny summer day associated with a high load factor or an intermittently cloudy autumn day associated with a low one). To further analyze these differences, the 309 days of the simulation were partitioned according to their distribution profile. The clustering was made automatically through a K-means method fed with the quantile values of the distribution of the PV load factor values at noon for the 500 scenarios. Because it seems that the partitioning depends mainly on the median value of this distribution, the partitioning is represented in Fig. 10 with only the median noon value for each day.

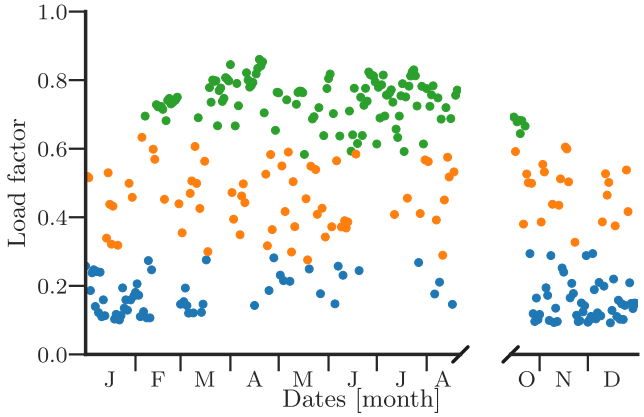

Fig. 10. PV load factor days clustering.

The "Low" cohort represented in blue contains 104 days, the "Medium" one, corresponding to the orange dots, contains 86 days, and there are 119 "High" PV scenario days represented by green dots. Unsurprisingly, it can be observed that winter days are mainly grouped in the "Low" cohort, and summer days are in the "High" one.

The clustering makes it possible to break down the relative and absolute expected SC gains (w.r.t. the Reference - Most Likely configuration) according to the type of day. The lower window of Fig. 11 represents the variability of the 309 absolute SC gains (same data as Fig. 9), according to the PV scenario types classification. The upper window displays the variability of the 309 relative SC gains according to the classification, as detailed in the following equations

$$\frac{1}{\mathrm{RML}(j)}\left[\frac{1}{K}\sum_{k=0}^{K} SC\left(X, S^j_k\right) - \mathrm{RML}(j)\right] \tag{30}$$

with $\mathrm{RML}(j)$ defined in Eq. (28). The equation is valued for the three same values of X as in Eq. (29).

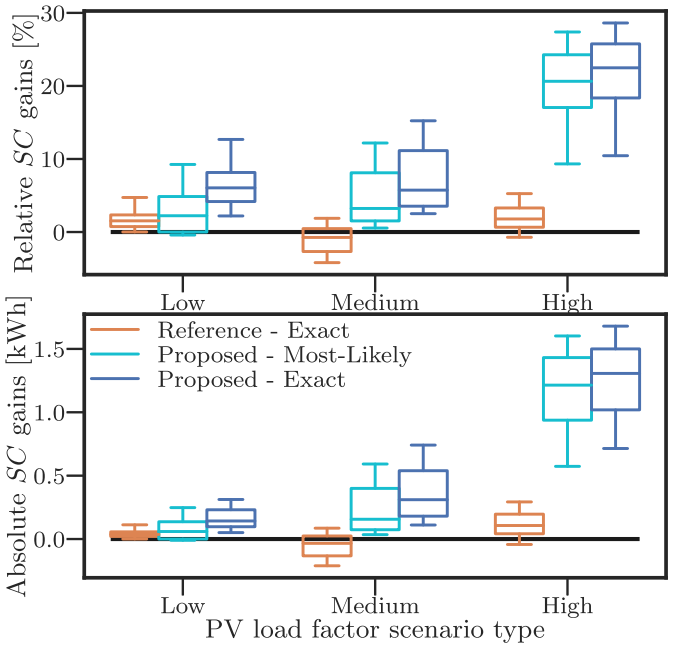

Fig. 11. The relative and absolute expected SC gains w.r.t. Reference - Most Likely tend to be higher with the proposed algorithm, and especially during "High" PV production days.

It is clear that the gains over the Reference - Most Likely case, both absolute and relative, are observed during high PV power production days. Because the reference control is activated as soon as the PV production reaches a threshold, it cannot tap into the higher production values later in the day, thus leading to lower self-consumption scores on clear sunny days. Besides, sunny summer days usually display low variability. Hence a most likely forecast is good enough for the proposed algorithm to reach high SC scores. Moreover, it is evident that a too-basic heuristic control can poorly benefit from the valuable information carried by an "exact" PV production forecast. This underlines the fact that further research efforts might be better spent on developing better control algorithms than on finding better forecasting methods for the considered time resolution.

5. Conclusions

This paper has studied the impact of PV production forecasts uncertainties on the performance of an EMS optimizing the self consumption of a residential PV installation. To this end, an optimal EWH scheduling algorithm has been proposed. A novel methodology to generate a set of realistic PV production scenarios was also presented. This method ensures that the ensemble of PV power scenarios is representative of the variability associated with most likely weather forecasts. When combined, the efficient optimization algorithm and these sets of PV production scenarios make it possible to assess the performance of such an energy management system according to various production scenarios, either assuming the availability of "perfect" PV production predictions or of only the most likely forecasts.

This study has shown that self-consumption performance benefits more from an efficient optimal management system like the one proposed in this paper than from "perfect" PV production predictions at a 30-minute timestep. Indeed, considering a simple rule-based commercial heuristic accessing a unique deterministic most likely forecast as a baseline and upgrading to the proposed optimization algorithm yields an 11% increase in annual cumulated expected self-consumed energy, whereas combining this proposed optimization algorithm and a "perfect" PV production forecast at a 30-minute timestep brings only an additional two-percentage-point increase (from +11% to +13%). A detailed analysis shows that these gains are concentrated in days of high PV power production.

This study suggests that for these installations:

- the current level of accuracy of PV production forecasts is good enough;
- improving the control algorithms can be valuable.

This work should be extended to study as well the impact of inhabitants' behavior uncertainties, both for hot-water usage and for uncontrolled electrical demand.

Assuming that PV production scenarios with a 30-minute timestep can be exact is also an oversimplification, so future research is needed to assess the performance loss of this kind on system due to faster production variations.

A shift towards jointly minimizing the total electricity bill and controlling multiple appliances would be a useful extension resulting in a more valuable EMS. Additional optimization criteria could also be considered, such as user discomfort when appliances are not run at desired times. Finally, this work could lead to a transition towards optimizing the behavior of a microgrid of diverse houses with multiple controlled appliances and shared energy resources, with guarantees of individual profits and data privacy.

Declaration of Competing Interest

The authors have no competing interests to declare.

References

[1] SolarPower Europe. Global Market Outlook for Solar Power 2018–2022. Tech. Rep.; 2018.
[2] IEA. World Energy Outlook 2018. Paris; 2018. doi: 10.1787/weo-2018-en
[3] Nousdilis AI, Christoforidis GC, Papagiannis GK. Active power management in low voltage networks with high photovoltaics penetration based on prosumers self-consumption. Appl Energy 2018;229:614–24. doi:10.1016/j.apenergy.2018.08.032.
[4] EU Directive. Directive 2018/2001 of the european parliament and of the council of 11 december 2018 on the promotion of the use of energy from renewable sources. Off J Eur Union 2018;L328:82–209.
[5] Quoilin S, Kavvadias K, Mercier A, Pappone I, Zucker A. Quantifying self-consumption linked to solar home battery systems: statistical analysis and economic assessment. Appl Energy 2016;182:58–67. doi:10.1016/j.apenergy.2016.08.077.
[6] Goebel C, Cheng V, Jacobsen H-A. Profitability of residential battery energy storage combined with solar photovoltaics. Energies 2017;10(7). doi:10.3390/en10070976.
[7] Roberts MB, Bruce A, MacGill I. Impact of shared battery energy storage systems on photovoltaic self-consumption and electricity bills in apartment buildings. Appl Energy 2019;245:78–95. doi:10.1016/j.apenergy.2019.04.001.
[8] Barbato A, Capone A. Optimization models and methods for demand-side management of residential users: a survey. Energies 2014;7(9):5787–824. doi:10.3390/en7095787.
[9] Ha DL, Ploix S, Zama E, Jacomino M. Control of energy in home automation by ressource constraint scheduling. In: Proceedings of the 15th international conference on control system and computer science; 2005.
[10] Malisani P. Pilotage dynamique de l'nergie du btiment par commande optimale sous contraintes utilisant la pnalisation intrieure. Ecole Nationale Suprieure des Mines de Paris; 2012.
[11] Oldewurtel F, Parisio A, Jones CN, Gyalistras D, Gwerder M, Stauch V, et al. Use of model predictive control and weather forecasts for energy efficient building climate control. Energy Build 2012;45:15–27. doi:10.1016/j.enbuild.2011.09.022.
[12] Sou KC, Weimer J, Sandberg H, Johansson KH. Scheduling smart home appliances using mixed integer linear programming. In: Proceedings of the IEEE conference on decision and control and European control conference. Orlando, FL, USA: IEEE; 2011. p. 5144–9. doi:10.1109/CDC.2011.6161081. ISBN 978-1-61284-801-3 978-1-61284-800-6 978-1-4673-0457-3 978-1-61284-799-3
[13] Wytock M, Moehle N, Boyd S. Dynamic energy management with scenario-based robust MPC. In: Proceedings of the American control conference (ACC). Seattle, WA, USA: IEEE; 2017. p. 2042–7. ISBN 978-1-5090-5992-8. doi:10.23919/ACC.2017.7963253.
[14] Eurostat. Disaggregated final energy consumption in households - dataset. 2018. Code : NRG_D_HHQ; URL https://ec.europa.eu/eurostat/databrowser/bookmark/fe627fab-b02c-4913-83fe-8cfa16ef44a0?lang=en.
[15] VHK. Review study of ecodesign and energy labelling for water heaters and tanks. Tech. Rep.; 2019.
[16] Cao S, Hasan A, Sirn K. Analysis and solution for renewable energy load matching for a single-family house. Energy Build 2013;65:398–411. doi:10.1016/j.enbuild.2013.06.013.

[17] Lefort A, Bourdais R, Ansanay-Alex G, Guguen H. Hierarchical control method applied to energy management of a residential house. Energy Build 2013;64:53–61. doi:10.1016/j.enbuild.2013.04.010.
[18] Sossan F, Kosek AM, Martinenas S, Marinelli M, Bindner H. Scheduling of domestic water heater power demand for maximizing PV self-consumption using model predictive control. In: Proceedings of the IEEE PES ISGT Europe. Lyngby, Denmark: IEEE; 2013 http://ieeexplore.ieee.org/document/6695317/. ISBN 978-1-4799-2984-9.
[19] Heleno M, Rua D, Gouveia C, Madureira A, Matos MA, Lopes JP, et al. Optimizing PV self-consumption through electric water heater modeling and scheduling. In: Proceedings of the IEEE Eindhoven PowerTech. Eindhoven, Netherlands: IEEE; 2015 http://ieeexplore.ieee.org/document/7232636/. ISBN 978-1-4799-7693-5.
[20] Beeker N, Malisani P, Petit N. Discrete-time optimal control of electric hot water tank. IFAC-PapersOnLine 2016;49(7):882–8. doi:10.1016/j.ifacol.2016.07.301.
[21] Pacaud F. Decentralized optimization methods for efficient energy management under stochasticity. Paris-Est; 2018.
[22] Inman RH, Pedro HT, Coimbra CF. Solar forecasting methods for renewable energy integration. Prog Energy Combust Sci 2013;39(6):535–76. doi:10.1016/j.pecs.2013.06.002.
[23] Paridari K, Parisio A, Sandberg H, Johansson KH. Robust scheduling of smart appliances in active apartments with user behavior uncertainty. IEEE Trans Autom Sci Eng 2016;13(1):247–59. doi:10.1109/TASE.2015.2497300.
[24] Pflaum P, Alamir M, Lamoudi MY. Comparison of a primal and a dual decomposition for distributed MPC in smart districts. In: Proceedings of the IEEE international conference on smart grid communications (SmartGridComm). Venice, Italy: IEEE; 2014. p. 55–60. ISBN 978-1-4799-4934-2. doi:10.1109/SmartGridComm.2014.7007622.
[25] Parisio A, Wiezorek C, Kyntj T, Elo J, Johansson KH. An MPC-based Energy Management System for multiple residential microgrids. In: Proceedings of the IEEE international conference on automation science and engineering (CASE). Gothenburg: IEEE; 2015. p. 7–14. ISBN 978-1-4673-8183-3. doi:10.1109/CoASE.2015.7294033.
[26] Parisio A, Wiezorek C, Kyntj T, Elo J, Strunz K, Johansson KH. Cooperative MPC-based energy management for networked microgrids. IEEE Trans Smart Grid 2017;8(6):3066–74. doi:10.1109/TSG.2017.2726941. Conference Name: IEEE Transactions on Smart Grid
[27] Oldewurtel F. Stochastic model predictive control for energy efficient building climate control. ETH Zurich; 2011.
[28] Thorey J, Chaussin C, Mallet V. Ensemble forecast of photovoltaic power with online CRPS learning. Int J Forecast 2018;34(4):762–73. doi:10.1016/j.ijforecast.2018.05.007.
[29] Beeker-Adda N. Modeling and control of electric hot water tanks: from the single unit to the group. PSL Research University; 2016.
[30] Amabile L, Bresch-Pietri D, El Hajje G, Labbe S, Petit N. An optimization methodology for self-consumption of residential photovoltaic energy. In: Proceedings of the IFAC world congress; 2019 https://hal.archives-ouvertes.fr/hal-02939016.
[31] Thomas D, Deblecker O, Ioakimidis CS. Optimal operation of an energy management system for a grid-connected smart building considering photovoltaics uncertainty and stochastic electric vehicles driving schedule. Appl Energy 2018;210:1188–206. doi:10.1016/j.apenergy.2017.07.035.
[32] Rabiee A, Sadeghi M, Aghaeic J, Heidari A. Optimal operation of microgrids through simultaneous scheduling of electrical vehicles and responsive loads considering wind and PV units uncertainties. Renew Sustain Energy Rev 2016;57:721–39. doi:10.1016/j.rser.2015.12.041.
[33] Wistrm U. Dymola - dynamic modeling laboratory; 2013.