

METHODOLOGY

Open Access



Recursive training based physics-inspired neural network for electric water heater modeling

Surya Venkatesh Pandiyan* and Jayaprakash Rajasekharan

From Energy Informatics.Academy Conference 2022 (EI.A 2022)
Vejle, Denmark. 24-25 August 2022

*Correspondence:
surya.v.pandiyan@ntnu.no

Department of Electric
Power Engineering, NTNU,
7491 Trondheim, Norway

Abstract

Aggregating flexibility from residential electric water heaters (EWHs) is fast gaining commercial interest. Flexibility modeling of an EWH involves highly precise and quick simulation of EWH water temperature using the EWH thermal dynamics model for various flexibility control actions. Since EWH tank water temperature data is usually unavailable or costly to obtain, developing an accurate and computationally inexpensive EWH thermal dynamics model with limited sensor data is essential for devising advanced control strategies for EWH flexibility aggregation. In this paper, we present a novel recursive training-based unsupervised physics-informed neural network (PINN) model for predicting tank water temperature which requires only historical EWH power consumption data to train the model. PINN models enable the integration of domain knowledge from traditional physical processes and methods into neural network (NN) models. Single-zone thermal grey-box differential equation model (DEM) is used as the basis to develop and demonstrate proof-of-concept of the proposed approach. Physics from the single-zone model is encoded into the PINN loss function to incorporate domain knowledge and the PINN architecture is structured to mimic the single-zone DEM. The recursive training approach enables the use of previous-step water temperature as an input to the simulation model. Two separate models for EWH ON- and OFF-states are developed and trained with real-world EWH power consumption data. Water temperature prediction results indicate that the proposed approach has similar performance as the traditional single-zone DEM model, thereby demonstrating the ability of the proposed model to learn the underlying physics behind the single-zone model without water temperature data. The proposed model has high accuracy and performs well outside the control set point temperatures indicating its suitability for simulating load shifting and other DR events. Additionally, EWH simulation results for two different scenarios with different water demand compositions are presented to study the effects of propagation errors on temperature prediction. The proposed approach paves the way for developing advanced EWH flexibility modeling tools for the aggregator to precisely control a large portfolio of EWHs considering user comfort and rebound effects.

Keywords: Physics-informed neural networks, Electric water heater, Temperature prediction, Flexibility modeling, Recursive training

Introduction

To reduce greenhouse gases and achieve carbon emissions reduction targets, many governments are rapidly increasing the share of highly intermittent renewable energy sources in their country's energy mix. These resources are distributed over the grid as opposed to the centralized generation units. Due to their inherent nature and distributed setup, renewable generation will pose considerable challenges to grid stability in the future. This calls for increased power system flexibility for the reliable operation of modern smart energy systems.

Power system flexibility needs are expected to increase by 66% in the EU from 2018 to 2040 under sustainable development scenarios (Iea 2022). The required flexibility will increase significantly once the renewable energy shares increase above 30 percent of annual electricity production in Europe (Huber et al. 2014). Traditional flexibility sources like conventional generators alone will not suffice to meet increasing flexibility requirements. Digitalization of the electricity grid resulting in smart grids presents an opportunity to tap more flexible resources from the demand side. Demand-side flexibility (DSF) is seen as one of the key contributing factors to tackling increasing flexibility needs. IEA estimates a global DSF potential of 4000 TWh/year or 15 percent of global electricity demand (Zhongming et al. 2017). A study estimates 12–23 GW of DSF in the northern European system with a total peak load of 77 GW (Söder et al. 2018). Additionally, DSF can provide various ancillary services such as frequency regulation, reactive power management, etc., to the grid at a lower cost. However, technical challenges involved in harnessing and realizing the full potential of DSF should be addressed first in order to develop appropriate solutions.

In the domestic sector, Electric Water Heaters (EWHs) have a great potential for providing DSF as they are power-intensive devices with thermal storage capabilities. A basic analysis commissioned by NVE shows that EWHs in Europe have an approximate daily flexibility capacity of 20 GW (Norges vassdrags-og energidirektorat 2021) which is equal to the total installed capacity in Czech Republic and more than the generation capacity in Finland. Aggregated flexibility potential of EWHs from 50% of Norwegian households is estimated to be around 1000 MWh/h (Lakshmanan et al. 2021). Aggregated control and operation of distributed EWHs can significantly address the flexibility needs at both distribution and transmission levels. Thus, it is of great importance to develop tools that can aggregate flexibility from EWHs for providing grid solutions.

Estimating near-term flexibility available from EWHs requires the prediction of future water temperature for different control actions considering future hot water demand. It is crucial to maintain the tank water temperature within a range, both for the user's comfort and to avoid legionella bacteria growth. Various control strategies for EWH have been explored in the literature and can be classified into model-based control (Ahmed et al. 2018; Nehrir et al. 2007; Vrettos et al. 2012) and model-free control (Ruelens et al. 2016; Cao et al. 2020; Kazmi et al. 2016). As the name suggests, model-based control requires an EWH model for finding optimum control action as opposed to model-free control. Most approaches in the model-based control use a simplified model with low

accuracy or assume access to data that is not possible/expensive in real world or use a complex model with high computational complexity. Most popular model-free control techniques such as reinforcement learning require a temperature sensor for online training or a simulator model for offline training. In-built temperature sensor from EWHs usually does not provide data and it is expensive to install and collect data from a new sensor. Therefore, an EWH model is required for developing model-free control. Thus, the control strategy for scheduling domestic EWHs for flexibility activation without affecting user comfort and severe rebound effect requires precise knowledge of appliance thermal dynamics and hot water usage profile.

EWH modeling can be classified into white-box, grey-box, and black-box based models. White-box or physics-based modeling of EWHs requires complete individual system information of each EWH which is either very difficult to obtain or unavailable to aggregators (Hossain et al. 2021). These models are developed using detailed knowledge of the underlying physical processes (Farooq et al. 2015) with few or no assumptions and do not require input-output data. However, they require significant effort to develop and tune the parameters such as friction parameter, mixing ratio, thickness of tank wall, etc. Even though white-box based models have good generalization capabilities, they have low accuracy compared to data-driven models (Farooq et al. 2015).

The grey-box differential equation model (DEM) based on energy flow analysis is a less complex approach compared to white-box modeling. Grey-box models use simplified mathematical structure from white-box models and require input-output data to estimate parameters such as thermal capacity and thermal resistance. Most studies use single-zone grey-box model for EWH modeling due to its simplicity and analytical tractability (Paull et al. 2010; Xiang et al. 2019; Ahmed et al. 2018). Single-zone models assume uniform tank water temperature which is quite simplistic. A DEM for single-zone EWH that uses only power consumption for parameter estimation and average water usage estimation is proposed in Shad et al. (2015). Despite its simplicity, accuracy of single-zone model is not suited for precise control of EWHs (Xu et al. 2014). Some grey-box approaches include thermal stratification effects in their models, assuming two or more thermal zones inside the EWH (Farooq et al. 2015; Nel et al. 2016; Alvarez et al. 2019; Zuñiga et al. 2017) and need more than one first-order differential equation for each thermal zone. While multiple-zone grey-box models have higher accuracy than single-zone models, they are computationally expensive.

Alternatively, data-driven black-box models can model non-linear dynamics and provide faster computation during temperature estimation but require expensive and extensive data collection (Han and Jentzen 2018). Moreover, black-box models cannot guarantee learning underlying EWH physical processes (Karniadakis et al. 2021). Therefore, there is a need for computationally efficient models that require limited data but also capture the detailed physics of EWH operation.

Scientific machine learning (SML) is an emerging field that seeks to integrate traditional engineering methods into machine learning-based techniques to utilize prior domain knowledge and physical phenomena in the learning process (Rackauckas et al. 2020). SML-based physics-informed neural networks (PINN) have been explored for modeling building thermal dynamics (Drgoña et al. 2021; Gokhale et al. 2022) and for estimating battery state of charge (Luzi et al. 2019). PINN is gaining popularity across

many fields and their application in power systems are discussed in Huang and Wang (2022).

SML-based approaches replacing traditional methods are gaining interest due to their speed and comparable performance. A recursive training based unsupervised PINN is proposed in this paper for developing EWH water temperature prediction model where the current water temperature is used as an important variable for accurate prediction of future water temperature. This paper provides proof-of-concept for applying SML-based approaches for developing data-driven temperature prediction models for EWHs when only historical power consumption data is available. The goal is to replace DEM with computationally efficient PINN models that can achieve similar or better performance. In addition to learning the physics of EWH operation, results from the proposed approach provide plenty of scope and insights for further developing advanced high-precision PINN-based EWH models for large-scale flexibility aggregation. The proposed model can also be easily extended to multiple-zone stratified model as the physics of each zone is similar to the single-zone model.

The main contributions of this paper are:

- Developing an unsupervised PINN-based temperature prediction model for EWHs when only historical power consumption data is available.
- Formulating a recursive training approach to model previous temperature as an input parameter in unsupervised learning setting.
- Demonstrating that the proposed model performs comparably with traditional single-zone DEM.

The rest of the paper is organized as follows. “[Background](#)” section provides theoretical background on single-zone DEM and summarizes parameter and water-rate estimation steps. Under the “[Methodology](#)” section, the loss functions, the PINN architecture and the recursive training procedure are discussed. Simulation results are presented, compared and discussed in “[Results and discussion](#)” section and the conclusion and future work are summarized in “[Conclusion and future work](#)” section.

Background

Thermostat temperature settings determine the ON/OFF states of the EWH. When water temperature reaches a lower threshold temperature T_{lo} , the in-built controller turns on the heater and vice-versa when the temperature reaches a higher threshold temperature T_{hi} . The evolution of water temperature profile depends upon EWH characteristics, current water temperature $T(t)$, future hot water demand, ambient temperature T_{amb} , inlet water temperature T_{in} , etc. This section presents the theory behind single-zone EWH model and also summarizes the steps involved in parameter estimation and water rate estimation when only power consumption data is available.

Single-zone DEM

The single-zone DEM assumes that water inside the EWH is perfectly mixed and the water temperature is uniform. It also assumes that the EWH has a single heating

element. The schematic representation of single-zone DEM model is shown in Fig. 1. The energy balance equation representing single-zone DEM is calculated as follows.

$$\dot{T}(t) = \frac{1}{C}[Q_{flow} + Q_{loss} + Q_H]. \tag{1}$$

where $C = \rho c_p V$ is the thermal capacity($Ws/^\circ C$), ρ represent the water density (1000 kg/m^3), c_p is specific heat capacity ($4196 \text{ Ws/Kg}^\circ C$) of water. Q_{flow} represents energy loss due to hot water demand and is a function of the water demand and the temperature difference between $T(t)$ and T_{in} . Q_{loss} represents natural heat loss to the atmosphere due to the temperature difference between $T(t)$ and T_{amb} . Q_H represents the energy gain due to the heating element converting electrical energy into thermal energy. Water temperature can be obtained by solving Eq. (1).

The Ordinary Differential Equation (ODE) describing temperature rate $\dot{T}(t)$ in the single-zone thermal model (Shad et al. 2015) is given in Eq. (2).

$$\dot{T}(t) = \frac{1}{C}[Q(t) - G(T(t) - T_{amb}(t)) - \rho c_p W_d(t)(T(t) - T_{in}(t))], \tag{2}$$

where $G = A/R$ is the thermal conductance($W/^\circ C$) of the tank, V is the volume of the tank, A is the surface area of the tank, and R is the thermal resistance of the tank. $W_d(t)$ denotes the hot water demand.

$Q(t)$ is the rate of energy input and is a function of water temperature in a thermostatically controlled water heater as described in Eq. (3). Q_0 denotes the nominal power rating of the EWH and $Q(t - \delta t)$ denotes the previous value of consumed power.

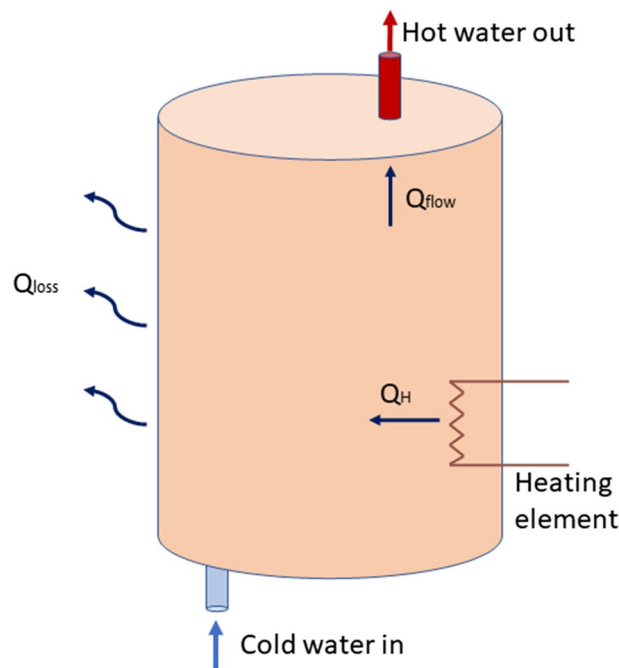


Fig. 1 Schematic representation of the energy flow in single-zone model

$$Q(t) = \begin{cases} Q_0 & T(t) \leq T_{lo} \\ 0 & T(t) \geq T_{hi} \\ Q(t - \delta t) & \text{otherwise.} \end{cases} \quad (3)$$

Analytical solution to Eq. (2). is given as:

$$T(t) = T(t_0)e^{-(t-t_0)/\tau} + K(1 - e^{-(t-t_0)/\tau}), \quad (4)$$

$$\tau = \frac{C}{G + \rho c_p W_d}, \quad K = \frac{GT_{out} + \rho c_p W_d T_{in} + Q}{G + \rho c_p W_d},$$

where t and t_0 represents current and initial time step and $T(t_0)$ represents temperature at initial time-step. Eq. (4) can be used estimate water temperature at any time-step.

Parameter and water rate estimation

There are multiple parameters in the single-zone DEM model that needs to be estimated. T_{amb} and T_{in} are assumed to be constant for simplicity. Thermostat set-points T_{hi} and T_{lo} are assumed based on literature review (Lakshmanan et al. 2021). These values are easily available for a particular geographical area and from the EWH’s technical specifications sheet.

Physical parameters C and G need to be estimated solely from power consumption data. Using the method proposed in Shad et al. (2015), C and G in Eq. (2) are estimated and the steps are summarized as follows.

- It can be fairly assumed that maximum OFF duration Max_{off} and minimum ON duration Min_{on} happen when hot water demand is zero.
- Max_{off} and Min_{on} can be obtained from the power consumption data and water demand is assumed zero.
- Replacing $(t - t_0)$ in Eq. (4) using Max_{off} and Min_{on} and solving the pair of the obtained equation provides an estimated value for C and G .

Hot water demand $W_d(t)$ is stochastic to some extent and can not be assumed constant. The duration for which EWH stays on or off depends upon the hot water demand during that duration. The average water rate during a particular event is also estimated using the method proposed in Shad et al. (2015) and is summarized as follows.

- From the power consumption data, the duration of each ON event and OFF event can be obtained. ON/OFF event represents the operation/non-operation of EWH between two control commands (turn-on and turn-off).
- Calculated duration along with estimated parameters C and G can be used in Eq. (4) to obtain the average water demand during a particular event.

Methodology

Prior knowledge from EWH thermal dynamics can be employed in PINN loss functions and in designing PINN architecture. Two separate temperature prediction models with one-minute prediction intervals for ON- and OFF-events of an EWH are developed.

This section describes data preparation, setting up physics-based loss functions, PINN architecture, and recursive training steps for the proposed approach.

Data preparation

The training data set used in this work contains minute-wise continuous power consumption data. The number of ON/OFF events and their duration can be inferred from power consumption data. From the data set, ON and OFF events are separated for training ON- and OFF-dynamics models respectively. Average water rates are estimated for randomly selected events and used for training.

Loss functions

Custom loss functions developed to train the neural network (NN) are explained here. Since temperature data is unavailable, it is necessary that developed loss functions well represent the physics of EWH in order for the NN to learn. Multiple loss functions that represent the physics are developed and listed here.

Temperature range

In a thermostat-controlled EWH, under normal operation, the water temperature will not exceed T_{hi} and will not drop below T_{lo} . This knowledge is modeled as a loss function L_1 and is shown in Eq. (5).

$$L_1 = \begin{cases} (T_{prd}(t) - \frac{(T_{hi}+T_{lo})}{2}) & T_{prd}(t) > T_{hi} \\ & \text{or } T_{prd}(t) < T_{lo} \\ 0 & T_{lo} < T_{prd}(t) < T_{hi}. \end{cases} \tag{5}$$

Loss function L_1 penalizes the network during training when the predicted temperature is not between T_{hi} and T_{lo} .

Temperature change

When the EWH is OFF, water temperature continues to decrease until T_{lo} before the EWH turns on. So in the OFF-dynamics model, the difference between two consecutive predictions for two consecutive time steps of an event should be positive. For the OFF-dynamics model, this information is modeled as a loss function L_2^{off} as shown in (Eq. 6).

$$L_2^{off} = \begin{cases} (T_{prd}(t) - T_{pred}(t - 1)) & T_{prd}(t) \geq T_{pred}(t - 1) \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

When the EWH is ON, water temperature will increase when water rate is sufficiently low or will decrease/stay the same when water rate is sufficiently high. Since the estimated average water rate is used in this model, water temperature will not decrease at any instant when EWH is ON according to DEM. For the ON-dynamics model, this information is modeled as loss function L_2^{on} as shown in Eq. (7).

$$L_2^{on} = \begin{cases} (T_{prd}(t - 1) - T_{pred}(t)) & T_{prd}(t - 1) \geq T_{pred}(t) \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Total temperature change per event

The water temperature at start (T_{start}) and end of an event (T_{end}) is known for a particular event. T_{start} and T_{end} are equal to thermostat control set-points T_{hi} and T_{lo} for OFF events. T_{start} and T_{end} are equal to T_{lo} and T_{hi} for ON events. Time t_{end} can be inferred from the power consumption data. This information is modeled as loss function L_3 as shown in Eq. (10).

$$T_{spd} = | (T_{hi} - T_{lo}) | . \tag{8}$$

$$T_{spd}^{prd} = | (T(t_{start}) - T_{pred}(t_{end})) | \tag{9}$$

$$t_{start} = 0 \quad T^{on}(t_{start}) = T_{lo} \quad T^{off}(t_{start}) = T_{hi}.$$

$$L_3 = \begin{cases} | (T_{spd} - T_{spd}^{prd}) | & T_{spd}^{prd} \neq T_{spd} \\ 0 & otherwise. \end{cases} \tag{10}$$

T_{spd} and T_{spd}^{prd} represent actual and predicted set points difference. L_3 loss function ensures that the predicted temperature t_{end} is equal to the set-point temperature for a particular event.

Relevant-ODE

L_1 - L_3 loss functions do not reflect the influence of water temperature $T(t)$ on the rate of change in water temperature $\dot{T}(t)$. Temperature change is continuous in the real world. Hence, $\dot{T}(t)$ also changes continuously. The predictions can have constant $\dot{T}(t)$ and still satisfy L_1 - L_3 losses. For a constant water rate, $\dot{T}(t)$ will decrease when $T(t)$ moves closer to T_{hi} when EWH is ON and $\dot{T}(t)$ will decrease when $T(t)$ moves closer to T_{lo} when EWH is OFF.

In a supervised PINN model, generally relevant ODE is used to model an additional loss function to incorporate prior scientific knowledge. ODE used in this work is given in Eq. (2). The goal is to minimize the difference between the gradient of the network’s output with respect to its inputs and the gradient calculated using DEM. This ensures that PINN solutions are consistent with the known physics. The modeled loss function L_4 is shown in Eq. (11).

$$L_4 = \left(\frac{dT_{DEM}(x_i)}{dt} - \frac{dNN(x_i)}{dt} \right), \tag{11}$$

$\frac{dT(x_i)}{dt}$ is calculated using Eq. (2). L_4 loss function ensures that for a constant water rate, $\dot{T}(t)$ will decrease when $T(t)$ moves closer to T_{hi} when EWH is on and $\dot{T}(t)$ will decrease when $T(t)$ moves closer to T_{lo} when EWH is off.

All loss values are scaled and combined together for training the PINN.

Network architecture

A simple feed-forward artificial neural network is used to develop the proposed architecture. A priori knowledge from single-zone model is used to define the network

architecture. More precisely, the initial layers are structured to mimic the differential equation defined in Eq. (2). This ensures that physics from single-zone model is embedded in the NN architecture and inputs are processed according to it before feeding it to neural layers. The schematic representation of the proposed architecture for the ON- and OFF-dynamics model is shown in Fig. 2. The number of layers and neurons in the figure does not reflect the actual model. NN parameters such as the number of layers,

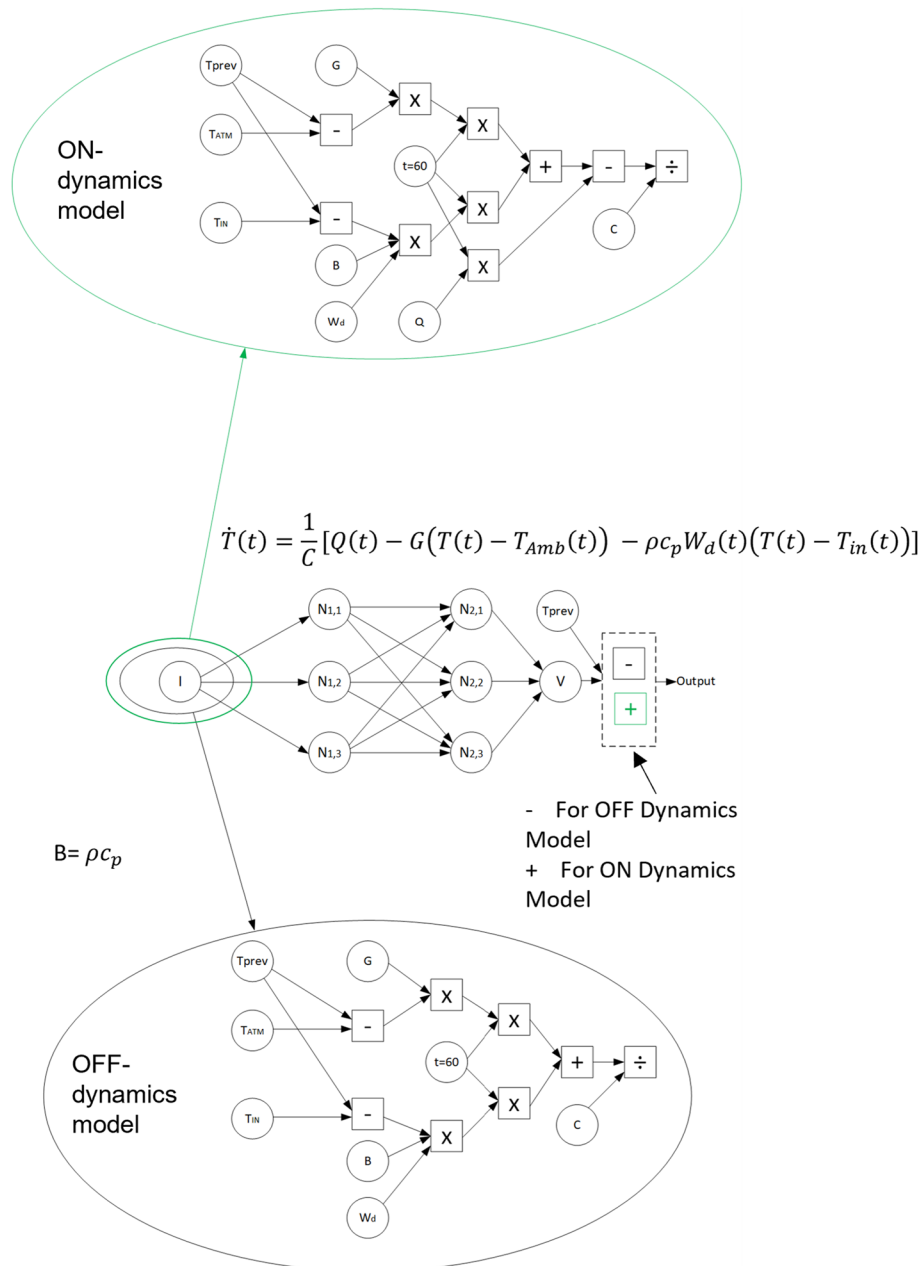


Fig. 2 Schematic representation of the proposed architecture for PINN for learning EWH ON- and OFF-dynamics. The difference in network architecture between ON- and OFF-dynamics model is shown using the zoomed-in view of the initial layers of PINN using different colors. The green view represents initial layers of ON-dynamics model and the black view represents initial layers of OFF-dynamics model

neurons, activation functions, etc., were selected heuristically. The inputs to the models are previous time step temperature, average water rate, and other relevant inputs to Eq. (2) which are assumed constant. The output is current water temperature.

There is also another constant input representing the time step (60s) of the prediction. This input is constant for all predictions and yet relevant to training the network. Without the time input, there is no time reference in the network. The time input is also required to calculate the L_4 loss function as it requires a gradient of NN with respect to the input time-step.

The proposed architecture mimicking the DEM requires previous step temperature as an input. However, the goal is to learn the underlying EWH thermal dynamics in the absence of temperature data. A recursive training approach is selected to train the network.

Recursive training

Algorithm 1: Pseudocode for recursive training

```

Initialize random weights;
for epoch=1 do
  for event=1 do
    set  $T_{prev}$  = initial set-point temperature;
    while  $event - step \neq event - duration$  do
      Calculate output;
      Update  $T_{prev}$ =output
    end
  end
  calculate loss value;
  Update network weights
end

```

In time series data, the current value is highly correlated with the immediate past values. Hence, previous step temperature is highly relevant for current temperature prediction model. Recursive training is an iterative training approach in which the next input is created based on the output from the previous training input. During training, new input is created recursively after processing the previous input. In this case, the previous step temperature input is created based on the PINN output to the previous input. The pseudo-code for recursive training approach is given as follows.

First, the network weights are initialized randomly. For each event, the initial point for previous step temperature is required and is available. It is known that the initial temperature for an event is equal to T_{hi} or T_{low} depending upon whether it is an OFF or ON event. T_{hi} or T_{low} will act as an initial point for previous step temperature. The next input is created based on the output predicted by PINN. This process repeats until the end of an event. Since the event duration is known, once an event ends, the input is reset to T_{hi} or T_{low} for the next event and the steps are repeated. Once all events are completed, the

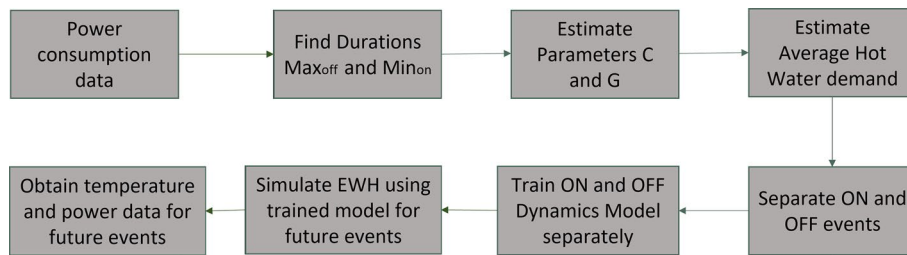


Fig. 3 The flow chart summarizing the entire model development procedure to model EWH dynamics without water temperature data

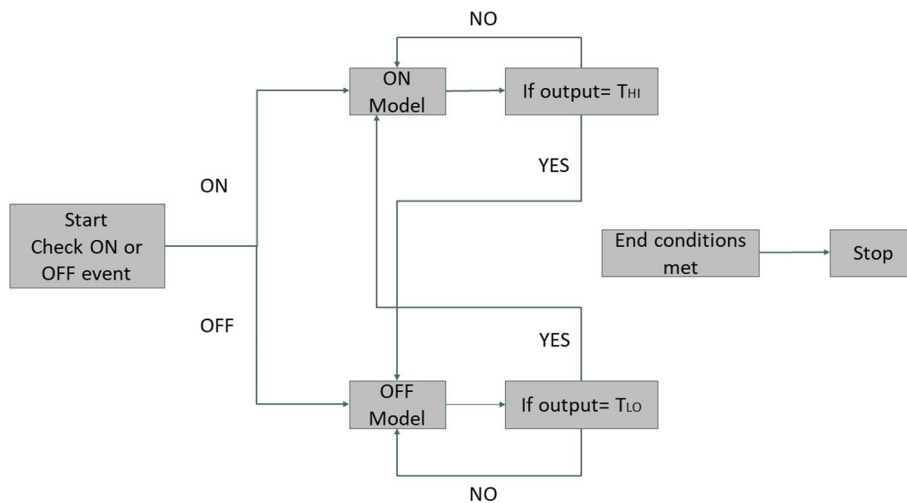


Fig. 4 The schematic representation of the procedure to use the trained model to simulate EWH over a certain period

loss value is calculated for the epoch and the weights are updated. The same procedure is repeated for the next remaining epochs.

The flow chart summarizing the entire model development is shown in Fig. 3. Once the NN is trained to reduce total custom loss function, ON- and OFF-dynamics models can be used recursively to predict the water temperature and power consumption can be inferred from the $T_{prd}(t)$ for the future events. The schematic representation of the procedure to use the trained model is shown in Fig. 4. The model requires an initial point for starting the prediction. If the initial point is the beginning of the ON-event, it will continue to use the ON-dynamics model until $T_{prd}(t)$ reaches T_{hi} , after which it switches to the OFF-dynamics model. This continues till the end of the simulation period.

Results and discussion

The proposed approach was tested using real-world power consumption data for an EWH available from the open-source data set Pecan Street Dataport (Dataport 2022). The frequency of power measurements is 1 min. An EWH was chosen randomly and was assumed to have a single heating element and uniform water tank temperature. The predicted temperature profiles were compared with temperature profiles estimated by single-zone DEM as true temperature data is unavailable in the data set. The

presented discussions are still relevant as the goal is to demonstrate that PINN can replace DEM methods and can have similar performance.

Max_{off} and Min_{on} were obtained from power consumption data and the parameters were estimated. The estimated parameters were used in the loss function L_4 . The obtained values are presented in Table 1.

50 random events each for ON- and OFF-dynamics model were selected. The duration of events were obtained from power consumption data and the average water demand for selected events were estimated using the duration and estimated parameters. The duration and water demand of selected events are presented in Figs. 5 and 6. The selected events well reflect the actual distribution of the water demand and the duration of EWH operation. It can be observed from Fig. 6 that water demand is higher during ON events than during OFF events. It is fair to assume that this is due to major activities such as taking showers, washing clothes, etc. This also indicates that temperature will fall quickly during demand response and supports the goal of this work.

The duration and water demand were used to create inputs to the NN. The models were built in Python using the Keras library and Adam optimizer was used for

Table 1 The estimated parameters for single-zone DEM and the maximum and minimum duration of OFF and ON events found in the dataset

Parameter	C	G	Max_{off}	Min_{on}
Value	1,167,000 (Ws/°C)	2.05 (W/°C)	959 min	21 min

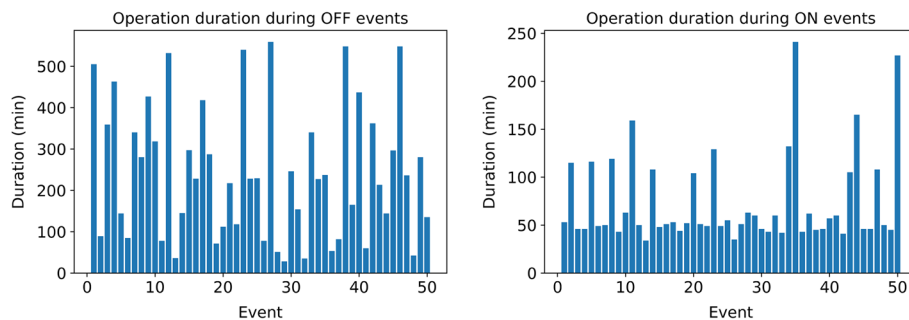


Fig. 5 EWH operation duration during ON and OFF events found in the training dataset

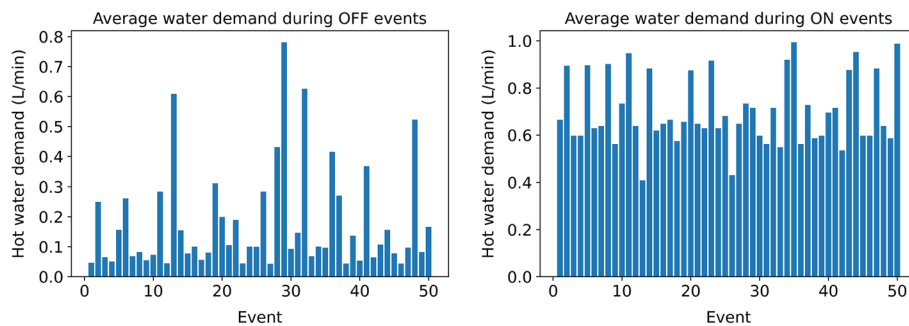


Fig. 6 Estimated average water demand for OFF and ON events found in the training data set

training. The proposed approach is generic and applies to any EWH with power consumption data.

The models were trained and their performance for the training dataset was analyzed first. Different error metrics were calculated and presented in Table 2 and Table 3 for both models. For the OFF-dynamics model, the mean absolute error (MAE) for training dataset shows that PINN predictions differ on average only by 0.15 °C from DEM predictions. OFF-dynamics model predictions for three different events with different water demands are shown in Fig. 7. It shows that PINN predictions are similar to DEM predictions. However, event-1 predictions seem to be inaccurate compared to events 2 and 3. This can be explained by the fact that water demand is very high for event 1 and high water demand is not well represented in the OFF-training dataset as can be seen in Fig. 6. This shows the importance of having an evenly distributed dataset for training.

Table 2 Prediction performance of trained OFF-dynamics PINN model for the training dataset compared to DEM

Metric	MSE	RMSE	MAE
Value (°C)	0.13	0.36	0.15

Table 3 Prediction performance of trained ON-dynamics PINN model for the training dataset compared to DEM

Metric	MSE	RMSE	MAE
Value (°C)	0.37	0.61	0.25

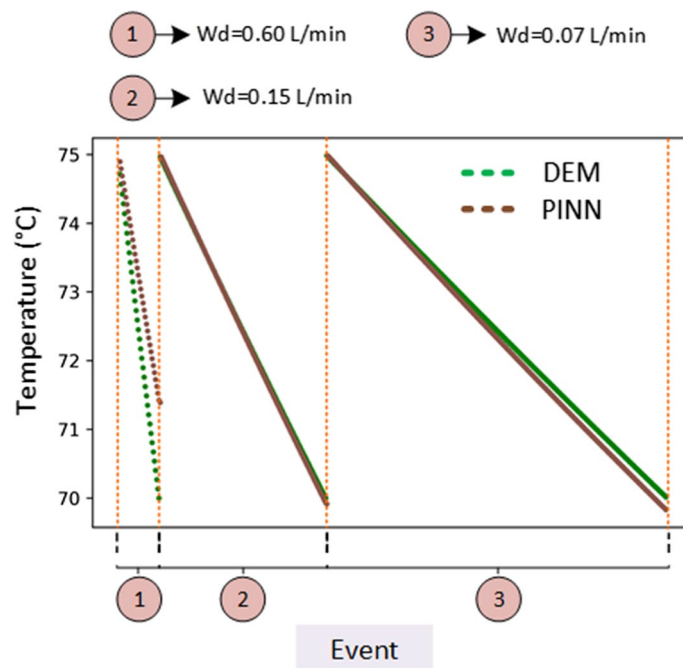


Fig. 7 OFF-dynamics model predictions for three different OFF-events with different water demands compared with DEM predictions

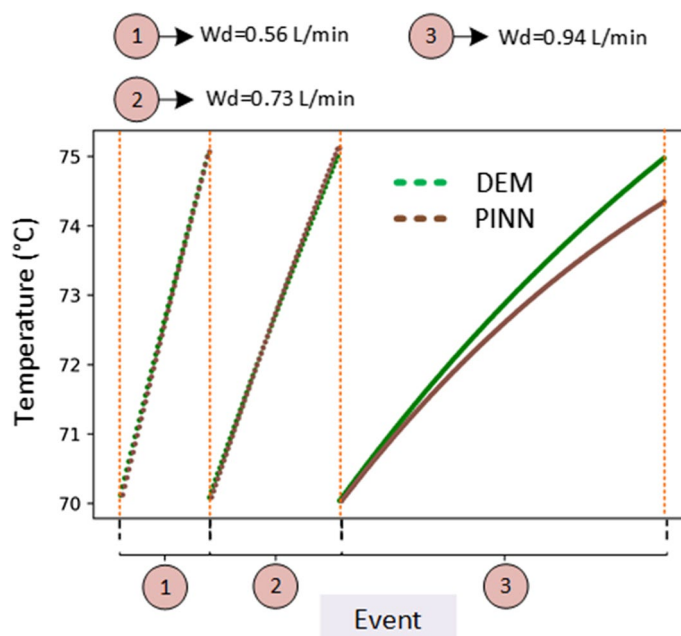


Fig. 8 ON-dynamics model predictions for three different ON-events with different water demands compared with DEM predictions

Similarly, predictions for three different ON events with different water demands by the ON-dynamics model are shown in Fig. 8. It can also be seen in Table 3 that the error is higher ($MAE = 0.25\text{ }^{\circ}\text{C}$) for ON-dynamics model compared to the OFF-dynamics model. However, in Fig. 8, the predictions for events 1 and 2 appear almost accurate. The inaccurate prediction for events with very high water demand like event-3 is the reason for high error in ON-dynamics model predictions. This can again be explained by the non-uniform distribution of water demands in the ON-training dataset and can be seen in Fig. 6.

In general, performance analysis for individual events from the training dataset indicates that the proposed model is capable of modeling the thermal dynamics of EWH between control set-points fairly accurately. However, during Demand response (DR), EWH operates outside control set-point temperatures. Thus it is important to assess the model’s performance outside the thermostat control range. It was assumed that there was no thermostat-based control and the evolution of the temperature profile predicted by PINN was examined for the case when EWH was not turned ON/OFF when it reaches T_{lo}/T_{hi} . The results were compared with the temperature profile provided by DEM and are shown in Figs. 9 and 10. It can be seen that PINN is able to capture the non-linear temperature evolution for both ON and OFF events without thermostat control. This shows that the proposed model can be used for flexibility modeling of EWH for load shifting during a DR event.

The combined performance of OFF-dynamics and ON-dynamics models in simulating EWH for future events is evaluated through two constructed scenarios. Scenario 1 consists of three ON and OFF events each. Scenario 2 consists of five ON and OFF events each. Water demands during events for scenarios 1 and 2 are presented in Tables 4 and 5. Water demands were carefully selected to be different from the data

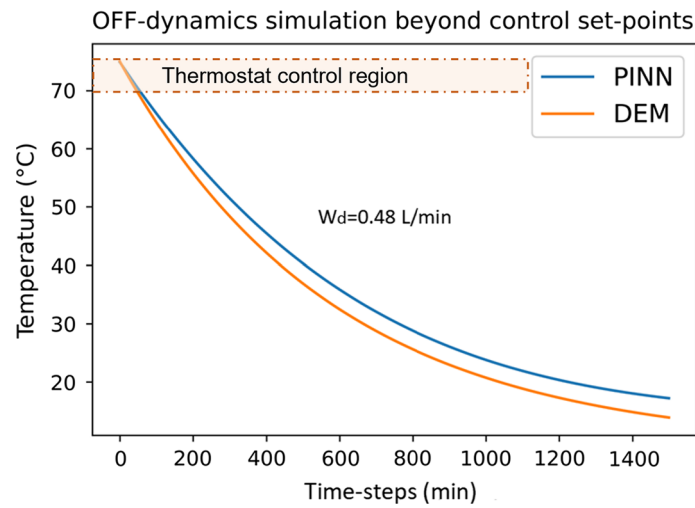


Fig. 9 OFF dynamics model performance outside thermostat control region

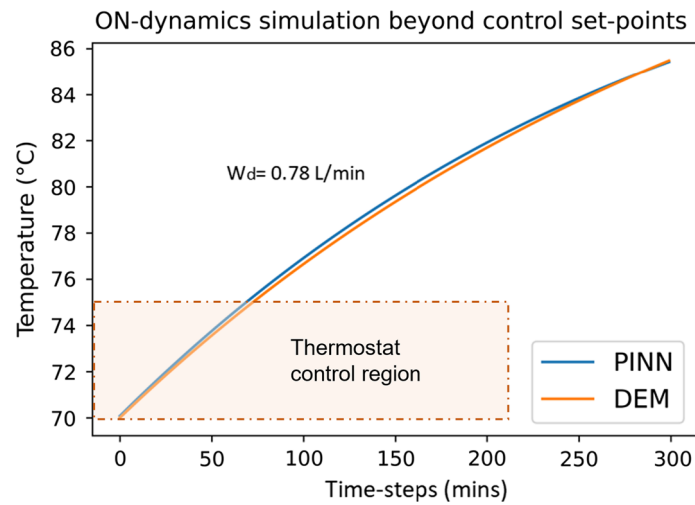


Fig. 10 ON dynamics model performance outside thermostat control region

Table 4 Scenario-1

Event	1	2	3	4	5	6
Operation	ON	OFF	ON	OFF	ON	OFF
Water demand (L/min)	0.60	0.13	0.33	0.3	0.76	0.42

present in the training dataset. Since PINN is based on a single-zone DEM model, the same is used for comparison purposes. The temperature predicted by PINN and DEM models for both scenarios are presented in Figs. 11 and 12. The plots show that for both scenarios, temperature predicted by PINN is fairly similar to the temperature predicted by DEM at least in the first few events. A lag in PINN predictions compared to DEM predictions can be observed in the later stages of the scenarios. This is due

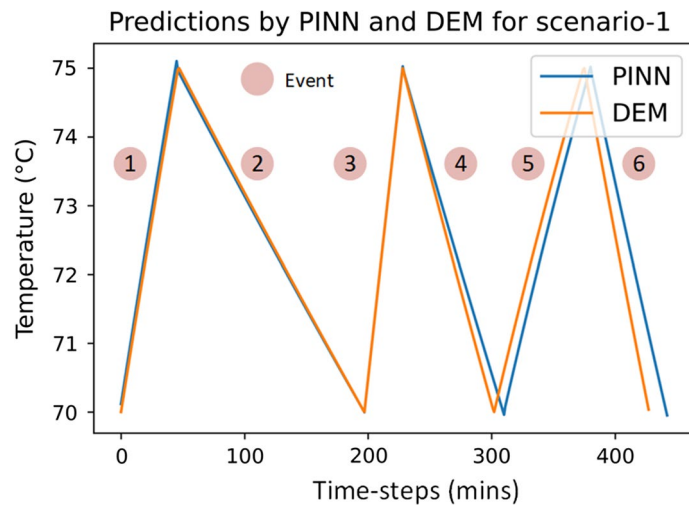


Fig. 11 The temperatures predicted by the PINN for the scenario-1 compared to DEM predictions

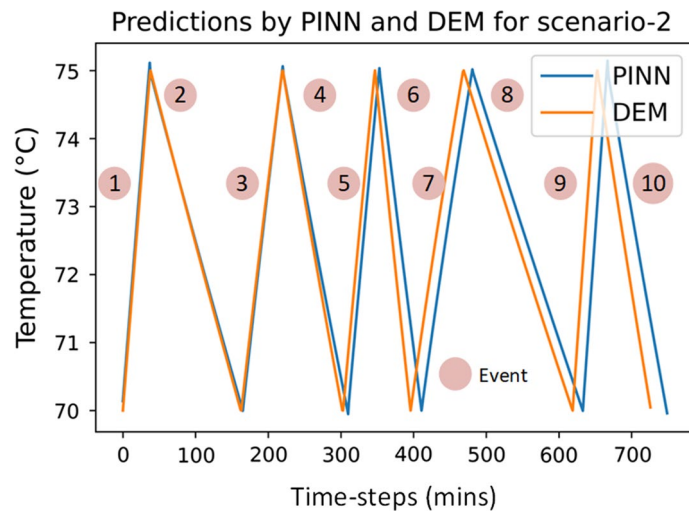


Fig. 12 The temperatures predicted by the PINN for the scenario-2 compared to DEM predictions

Table 5 Scenario-2

Event	1	2	3	4	5	6	7	8	9	10
Operation	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF
Water demand (L/min)	0.48	0.18	0.66	0.27	0.58	0.45	0.76	0.13	0.39	0.3

to the effect of error propagation from initial stages occurring in recursive prediction strategy and is a very well-known phenomenon.

Different error metrics were calculated for both scenarios to compare the performance of PINN with DEM and are shown in Table 6. The MAE for scenario-1 with the simulated duration of almost 7.5 h shows that PINN predictions can differ on average only by 0.22 °C from DEM predictions. For scenario-2 with the simulated duration of almost 12 h, MAE is 0.47 °C. This clearly shows the impact of error propagation from initial stages.

Table 6 Simulation performance of PINN for different scenarios compared to DEM

-	Metric	MSE	RMSE	MAE
Scenario-1	Value (°C)	0.11	0.33	0.22
Scenario-2	Value (°C)	0.44	0.66	0.47

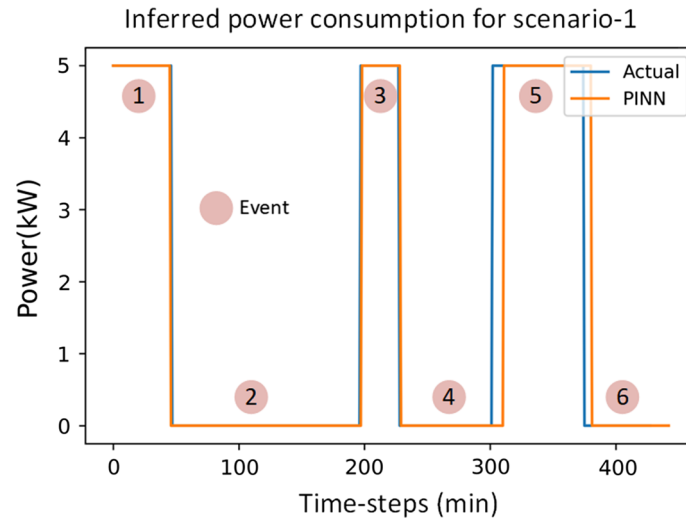


Fig. 13 The power inferred from the predicted temperatures from PINN for the scenario-1 compared to actual power consumption

Also, mean error metrics can be misleading as they can not explain the impact of the propagated error. Though the error values are small, the plots show that the error for a single point will be large when the time step is further from zero. This would cause the entire event to lag by large time-steps if not corrected. This is not desirable for DR programs as the aggregator would have to pay a huge penalty for not fulfilling contractual obligations in terms of flexibility delivered or user comfort. However, the more accurate predictions for the first few events in initial stages is promising since aggregator participating in DR programs usually requires shorter simulation periods.

Power consumption for scenarios 1 and 2 can be inferred from predicted temperatures and is presented in Figs. 13 and 14. The plots show that inferred power consumption using PINN model closely matches the actual power consumption. The lag in output observed for PINN model is due to an error in the time step at which the model switches between the ON- and OFF-dynamics model. This in turn is due to the propagating error from initial stages as explained earlier. Precision, Recall, and F1 scores were calculated based on the correct prediction of EWH ON and OFF states using inferred power consumption data and are presented in Table 7. The results for scenario-1 show that PINN can infer 92% of actual ON states correctly and 95% of total inferred ON states are actual ON states. The performance degrades for scenario-2 as the values are 82% and 85%, respectively. This is expected as scenario-2 is simulated for a longer period.

This analysis emphasizes the impact of propagating error from initial stages on flexibility modeling if not corrected. For example, consider two cases where EWH power

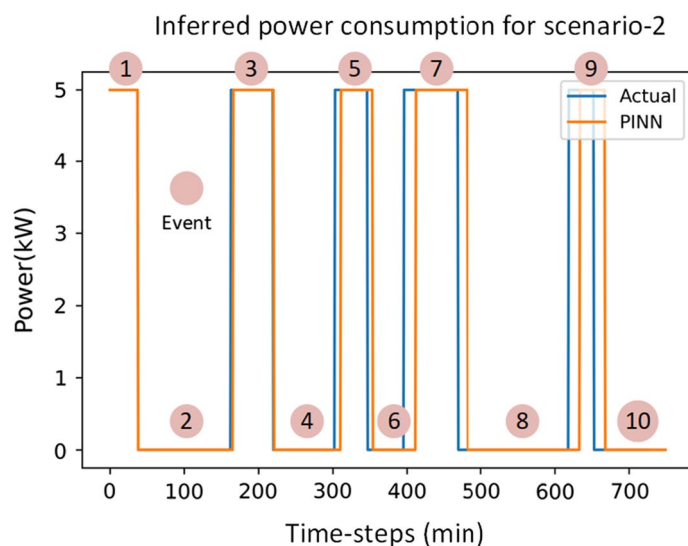


Fig. 14 The power inferred from the predicted temperatures from PINN for the scenario-2 compared to actual power consumption

Table 7 Performance of PINN in predicting EWH operational states compared with actual states

Metric	Precision	Recall	F1-score
Scenario-1	0.95	0.92	0.93
Scenario-2	0.85	0.82	0.83

consumption data is available after simulation in one case and not available in another case. EWH power consumption data can be used to correct the lag in predicted events simulation since the true power consumption of a simulated event is available after sometime. Non-intrusive load monitoring (NILM) can be used as a correcting strategy for the case where EWH power consumption data is not available after simulation. NILM is a technique to disaggregate individual appliance power consumption from total power consumption from household smart meter measurements (Garcia et al. 2020; Zufferey et al. 2020). As aggregators can have access to total consumption through smart meter data, it is possible to infer EWH consumption to a good extent from total power consumption using NILM techniques. An aggregator can compare the NILM output and PINN simulation. The PINN simulation model can be readjusted for the accounted error and EWH can be simulated again. For example, consider an aggregator simulated EWH for the next 24 h. After 12 h, the aggregator can use NILM to infer EWH consumption for the past 12 h from total consumption data. PINN results and NILM results can be compared for these 12hrs and any observed deviation can be noted. The aggregator can then reset PINN simulation for the next 12 h based on NILM results as it can provide the actual time-step of recent event start/end. It should be noted that such approaches are suitable only for short-term decision-making problems.

To summarize, individual event analysis and simulation beyond control set-points indicate that the proposed model learns the physics described in Eq. (2) and the trained model has high prediction accuracy for individual events. Average water demand in the

training set should be uniformly distributed. The propagating error is of concern for simulating EWH for longer periods using the proposed approach.

The advantages of the proposed approach are:

- The proposed approach is generic and can be used for any water heater.
- Neural network based EWH temperature prediction model can be developed without having actual temperature data.
- Recursive approach allows previous step temperature as input and thus NN can mimic the differential equation.
- Can easily be integrated into decision support tools for DR programs. This is mainly because the previous temperature input allows starting point of the simulation from any point in the temperature range of interest. This makes it convenient to simulate DR event without additional steps required.

The disadvantages of the proposed approach are:

- Temperature stratification effects in the water tank are not captured.
- The average water demand doesn't reflect real-world scenarios making the proposed approach slightly less accurate for flexibility modeling. However, the proposed approach can be easily modified to adapt to stochastic water demand input.
- It is not possible to predict the time-step of interest directly. Recursively moving towards the time-step of interest might not be desirable at sometimes.

Conclusion and future work

This paper presents a framework for developing PINN based EWH simulation model with only historical power consumption data. The proposed approach is generic and can be applied to any EWH. Training dataset preparation involving thermal parameter estimation and water demand estimation is described. The PINN architecture is structured to mimic DEM and physics-based custom loss functions are designed to facilitate unsupervised learning. A recursive training procedure supporting previous time-step temperature to be modeled as a PINN input is introduced for learning EWH thermal dynamics. Two different models representing EWH ON and OFF states are trained with real-world EWH power consumption data from open-source data set available in Pecan Street Dataport.

The temperature prediction results for individual events show that the proposed approach has comparable performance to traditional single-zone-DEM. The results promise the possibility of employing unsupervised learning for advanced thermal modeling of EWH using PINN. The trained model also showed considerably good results when the EWH was simulated outside the thermostat control range. The proposed model was also used to simulate EWH for two scenarios with different event compositions. The usability of the proposed approach for EWH simulation during a DR event is demonstrated through these two analyses. However, the recursive prediction strategy introduces propagation errors into the prediction results necessitating measures to reduce or correct propagating errors.

This work demonstrates the application of PINN for developing an EWH simulation model based on single-zone DEM. The proposed approach can serve as a stepping stone for building advanced PINN-based thermal dynamics models that could potentially result in less computationally intensive, high accuracy model that can help the aggregator to develop effective control strategies for precisely controlling EWH portfolio without causing user discomfort and rebound effects.

Future work will focus on (1) incorporating EWH stratification effects into PINN and comparing the results with actual temperature measurements, (2) improving loss function and PINN architecture for better performance, (3) evaluating the developed model for EWH flexibility modeling and aggregation and (4) methods to reduce propagation errors.

Abbreviations

EWH	Electric water heater
PINN	Physics inspired neural network
DSF	Demand side flexibility
DR	Demand response
NN	Neural network
DEM	Differential equation model
SML	Scientific machine learning
ODE	Ordinary differential equation
MAE	Mean absolute error
RMSE	Root mean squared error
MSE	Mean squared error
NILM	Non-intrusive load monitoring

Nomenclature

t	Time step of an event [min]
$T(t)$	Water temperature at time t [$^{\circ}\text{C}$]
T_{amb}	Ambient temperature [$^{\circ}\text{C}$]
T_{in}	Inlet water temperature [$^{\circ}\text{C}$]
T_{lo}	Thermostat lower set-point temperature [$^{\circ}\text{C}$]
T_{hi}	Thermostat higher set-point temperature [$^{\circ}\text{C}$]
$Q(t)$	Power consumption [W]
$\dot{T}(t)$	Rate of change of temperature [$^{\circ}\text{C}$]
$W_d(t)$	Hot water demand [L/min]
$T_{prd}(t)$	Predicted water temperature [$^{\circ}\text{C}$]
T_{st}	Water temperature at the start of an event [$^{\circ}\text{C}$]
T_{end}	Water temperature at the end of an event [$^{\circ}\text{C}$]
T_{spd}	Temperature difference between control set-points [$^{\circ}\text{C}$]

Acknowledgements

Not applicable.

Author contributions

SVP was involved in conceiving the idea for applying PINN to EWH modeling, developing the model, generating results and creating the manuscript. JR was involved in conceptualizing the idea, reviewing the approach, results and manuscript. All authors read and approved the final manuscript.

About this supplement

This article has been published as part of Energy Informatics Volume 5 Supplement 4, 2022: Proceedings of the Energy Informatics Academy Conference 2022 (EIA 2022). The full contents of the supplement are available online at <https://energyinformatics.springeropen.com/articles/supplements/volume-5-supplement-4>.

Funding

This paper is funded by the authors affiliations.

Availability of data and materials

The data used in this work is from open source dataset available in Pecan Street Dataport. The data can be obtained from <https://www.pecanstreet.org/dataport/>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Published: 21 December 2022

References

- Ahmed MT, Faria P, Abrishambaf O, Vale Z (2018) Electric water heater modelling for direct load control demand response. In: 2018 IEEE 16th International Conference on Industrial Informatics (INDIN). IEEE; p. 490–495
- Alvarez MAZ, Agbossou K, Cardenas A, Kelouwani S, Boulon L (2019) Demand response strategy applied to residential electric water heaters using dynamic programming and K-means clustering. *IEEE Trans Sustain Energy*. 11(1):524–533
- Cao J, Dong L, Xue L (2020) Load Scheduling for an Electric Water Heater With Forecasted Price Using Deep Reinforcement Learning. In: 2020 Chinese Automation Congress (CAC). IEEE; p. 2500–2505
- Dataport (2022). Available from: <https://www.pecanstreet.org/dataport/about-dataport/>. Accessed 10 Mar 2022
- Drgoña J, Tuor AR, Chandan V, Vrabie DL (2021) Physics-constrained deep learning of multi-zone building thermal dynamics. *Energy Build* 243:110992
- Farooq AA, Afram A, Schulz N, Janabi-Sharifi F (2015) Grey-box modeling of a low pressure electric boiler for domestic hot water system. *Appl Therm Eng* 84:257–267
- Garcia FD, Souza WA, Diniz IS, Marafão FP (2020) NILM-based approach for energy efficiency assessment of household appliances. *Energy Informat* 3(1):1–21
- Gokhale G, Claessens B, Devellder C (2022) Physics informed neural networks for control oriented thermal modeling of buildings. *Appl Energy* 314:118852
- Han J, Jentzen AEW (2018) Solving high-dimensional partial differential equations using deep learning. *Proc National Acad Sci* 115(34):8505–8510
- Hossain MM, Zhang T, Ardakanian O (2021) Identifying grey-box thermal models with Bayesian neural networks. *Energy Build* 238:110836
- Huang B, Wang J (2022) Applications of physics-informed neural networks in power systems—a review. *IEEE Trans Power Syst*
- Huber M, Dimkova D, Hamacher T (2014) Integration of wind and solar power in Europe: assessment of flexibility requirements. *Energy* 69:236–246
- iea (2022) World energy outlook 2019 analysis; 2019. Available from: <https://www.iea.org/reports/world-energy-outlook-2019>. Accessed 4 Apr 2022
- Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. *Nat Rev Phys* 3(6):422–440
- Kazmi H, D'Oca S, Delmastro C, Lodeweyckx S, Corgnati SP (2016) Generalizable occupant-driven optimization model for domestic hot water production in NZEB. *Appl Energy* 175:1–15
- Lakshmanan V, Sæle H, Degefa MZ (2021) Electric water heater flexibility potential and activation impact in system operator perspective-Norwegian scenario case study. *Energy* 236:121490
- Luzi M, Mascioli FMF, Paschero M, Rizzi A (2019) A white-box equivalent neural network circuit model for SoC estimation of electrochemical cells. *IEEE Trans Neural Netw Learn Syst* 31(2):371–382
- Nehrir MH, Jia R, Pierre DA, Hammerstrom DJ (2007) Power management of aggregate electric water heater loads by voltage control. In: 2007 IEEE Power Engineering Society General Meeting. IEEE :1–6
- Nel P, Booysen MJ, van der Merwe B (2016) A computationally inexpensive energy model for horizontal electric water heaters with scheduling. *IEEE Trans Smart Grid* 9(1):48–56
- Paull L, Li H, Chang L (2010) A novel domestic electric water heater model for a multi-objective demand side management program. *Electric Power Syst Res* 80(12):1446–1451
- Rackauckas C, Ma Y, Martensen J, Warner C, Zubov K, Supekar R et al. (2020) Universal differential equations for scientific machine learning. arXiv preprint [arXiv:2001.04385](https://arxiv.org/abs/2001.04385)
- Ruelens F, Claessens BJ, Quaiyum S, De Schutter B, Babuška R, Belmans R (2016) Reinforcement learning applied to an electric water heater: from theory to practice. *IEEE Trans Smart Grid* 9(4):3792–3800
- Shad M, Momeni A, Errouissi R, Diduch CP, Kaye ME, Chang L (2015) Identification and estimation for electric water heaters in direct load control programs. *IEEE Trans Smart Grid* 8(2):947–955
- Söder L, Lund PD, Koduvere H, Bolkesjø TF, Rossebø GH, Rosenlund-Soysal E et al (2018) A review of demand side flexibility potential in Northern Europe. *Renewable Sustain Energy Rev* 91:654–664
- Value of flexibility from electrical storage water heaters (2022) Norges vassdrags- og energidirektorat; 2021. Available from: https://publikasjoner.nve.no/eksternrapport/2021/eksternrapport2021_05.pdf. Accessed 15 Apr 2022
- Vrettos E, Koch S, Andersson G (2012) Load frequency control by aggregations of thermally stratified electric water heaters. In: 2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe). IEEE:1–8
- Xiang S, Chang L, Cao B, He Y, Zhang C (2019) A novel domestic electric water heater control method. *IEEE Trans Smart Grid* 11(4):3246–3256

- Xu Z, Diao R, Lu S, Lian J, Zhang Y (2014) Modeling of electric water heaters for demand response: a baseline PDE model. *IEEE Trans Smart Grid* 5(5):2203–2210
- Zhongming Z, Linong L, Xiaona Y, Wangqiang Z, Wei L et al (2017) A world in transformation: World Energy Outlook 2017
- Zufferey T, Valverde G, Hug G (2020) Unsupervised disaggregation of water heater load from smart meter data processing. In: The 12th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2020). vol. 2020; p. 283–288
- Zuñiga M, Agbossou K, Cardenas A, Boulon L (2017) Parameter estimation of electric water heater models using extended kalman filter. In: IECON 2017-43rd Annual Conference of the IEEE Industrial Electronics Society. IEEE; p. 386–391

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
