One System, Many Models: Designing a Surrogate Model for Sulfur Thermal Energy Storage

Preprint

Kevin Menear,[1,3] Dmitry Duplyakin,[1] Madeleine C. Oliver,[1,4] Munjal Shah,[1,5] Michael James Martin,[1] Janna Martinek,[1] Karthik Nithyanandam,[2] and Zhiwen Ma[1]

1 National Renewable Energy Laboratory
2 Element 16 Technologies
3 Georgia Institute of Technology
4 University of Oklahoma
5 University at Buffalo

Presented at the AAAI 2022 Fall Symposium Series
Arlington, Virginia
November 17-19, 2022

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-2C00-84235
December 2023

One System, Many Models: Designing a Surrogate Model for Sulfur Thermal Energy Storage

Preprint

Kevin Menear,[1,3] Dmitry Duplyakin,[1] Madeleine C. Oliver,[1,4] Munjal Shah,[1,5] Michael James Martin,[1] Janna Martinek,[1] Karthik Nithyanandam,[2] and Zhiwen Ma[1]

1 National Renewable Energy Laboratory
2 Element 16 Technologies
3 Georgia Institute of Technology
4 University of Oklahoma
5 University at Buffalo

Suggested Citation
Menear, Kevin, Dmitry Duplyakin, Madeleine C. Oliver, Munjal Shah, Michael James Martin, Janna Martinek, Karthik Nithyanandam, and Zhiwen Ma. 2023. *One System, Many Models: Designing a Surrogate Model for Sulfur Thermal Energy Storage: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-2C00-84235. https://www.nrel.gov/docs/fy24osti/84235.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-2C00-84235
December 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported with funding from the US DOE HPC4EI program, additional funding from the US DOE AMO office through ORISE, and time on NREL's high performance computing system. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

One System, Many Models: Designing a Surrogate Model for Sulfur Thermal Energy Storage

Kevin Menear[1,3], Dmitry Duplyakin[1], Madeleine C. Oliver[1,4], Munjal Shah[1,5], Michael James Martin[1], Janna Martinek[1], Karthik Nithyanandam[2], Zhiwen Ma[1]

[1]National Renewable Energy Laboratory, Golden, CO, USA [2]Element 16 Technologies, Arcadia, CA, USA
[3]College of Computing, Georgia Institute of Technology, Atlanta, GA, USA
[4]School of Chemical, Biological and Materials Engineering, University of Oklahoma, Norman, OK, USA
[5]Department of Mechanical and Aerospace Engineering, University at Buffalo, Buffalo, NY, USA

Abstract

Industrial process heating (IPH) relies primarily on thermal energy generated by fossil fuel combustion to produce, treat, and alter manufactured goods. Thermal energy storage (TES) helps reduce the carbon footprint of IPH systems by facilitating the utilization of renewable and waste heat sources. A promising new TES technology uses elemental sulfur as the heat-storage medium. The design of sulfur TES systems can be evaluated with the aid of computational fluid dynamics (CFD). However, the computational cost of such CFD efforts is prohibitive to comprehensive optimizations over design parameters. To reduce this computational cost, machine learning (ML) models can be developed to act as surrogates for CFD. In this paper, we describe the process of building and evaluating surrogate ML models for facilitating optimization of sulfur TES systems for IPH. To enforce the thermodynamic relationship between the two modeled quantities, we develop a hybrid model for sulfur temperature using both direct predictions of temperature and calculations of temperature from predictions of the heat transfer coefficient. The hybrid model enforces this constraint at the expense of the slightly reduced accuracy compared to two disjoint models. The overall high accuracy observed in our model evaluation demonstrates the usefulness of such surrogate modeling for studying TES systems. This work contributes to the field of TES surrogate modeling by offering a novel accurate hybrid approach to predicting simultaneously the heat transfer coefficient and temperature of the heat storage medium.

Introduction

IPH constitutes ∼70% of US manufacturing energy use (Kurup and Turchi 2015). TES can reduce IPH costs and carbon footprint by reducing energy lost to waste heat and increasing the use of renewable sources of thermal energy. Molten sulfur can operate in the 100-300 °C temperature range, which encapsulates ∼60% of IPH energy demand (McMillan and Ruth 2019; Schoeneberger et al. 2020), and is an order of magnitude lower cost than molten salt, which is the state-of-the-art heat storage medium for higher-temperature TES systems (Barde et al. 2018, 2020; Jin et al. 2019; Jin and Wirz 2020).

Modeling TES systems using CFD simulations enables rapid iterative design modifications without the cost of constructing prototypes (Al-Abidi et al. 2013). However, CFD simulations can take hours or days to run (Yaïci et al. 2013). Under certain conditions, the thermodynamic behavior of a system can be estimated with dimensionless parameter correlations (Cheng, Yang, and Huang 2009), which mitigates the computational cost of CFD. However, complex nonlinear variations in fluid properties with temperature cause the dimensionless parameter correlations approach to break down.

In the work described here, ML surrogate models were developed to emulate CFD for a sulfur TES system for IPH. Typical dimensionless parameter correlations break down for this system due to the complex variations of sulfur properties (especially viscosity and specific heat capacity) with temperature, necessitating an alternative numerical method and leading to the development of the surrogate ML models. Training data was developed by running CFD simulations of the system for a representative set of operating conditions. The input features for the ML models are the temperature of the heat transfer wall (T_w), the initial temperature of sulfur (T_i), and the flow-time (t); and the targets of the models are heat transfer coefficient ($h(t)$) and sulfur average temperature ($T_s(t)$).

To ensure the surrogate model predicts physically realizable system behavior, efforts were made to enforce the thermodynamic constraint between heat transfer coefficient and sulfur average temperature. Significant error was encountered when enforcing the constraint, which motivated the development of a hybrid model. The hybrid model augments the ML-based predictions with calculations via the thermodynamic formula relating sulfur temperature and heat transfer coefficient for the TES system.

In this paper, we describe:

- Selected ML model architectures
- Calculation of $h(t)$ based on $T_s(t)$ predictions
- Calculation of $T_s(t)$ based on $h(t)$ predictions
- Hyperparameter optimization
- Development of the final hybrid model
- Evaluation of surrogate models on the per-scenario basis

Finally, we discuss the impact of this work on the field of TES surrogate modeling and include a brief summary of related efforts.

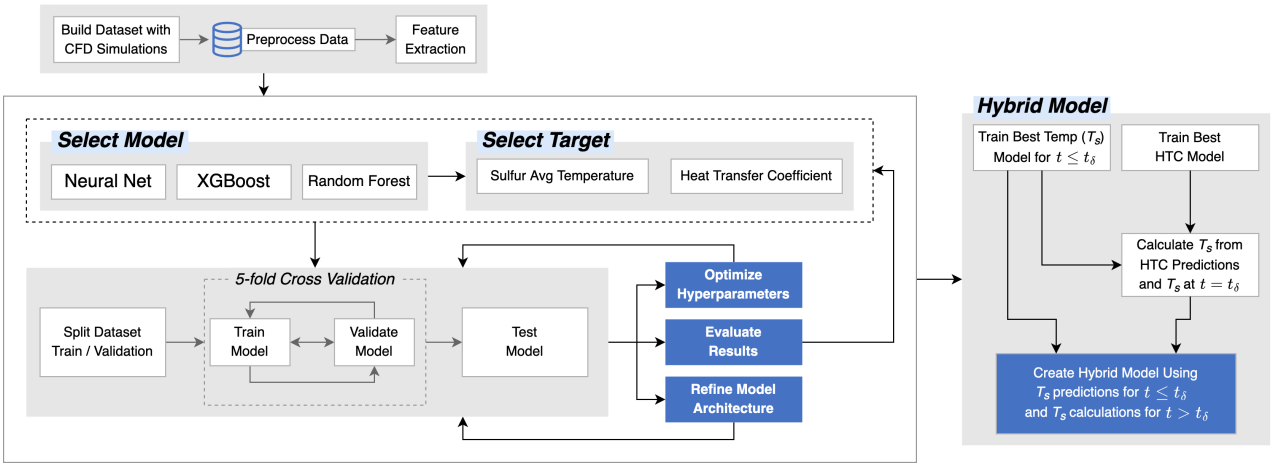

Figure 1: The proposed ML-based computational workflow. The CFD-based dataset characterizing a sulfur TES system is processed, and key features are extracted for model training. 5-fold cross-validation is used. 20 repetitions (with unique dataset splits into training and testing subsets) are applied to each model's validation. Hyperparameters are optimized to obtain the best models, and a hybrid model is developed. t_δ is the transition time between predictions and calculations in the hybrid model; currently, $t_\delta = 360s$. Before and after t_δ, different schemes are used to estimate $T_s(t)$ for best results.

Background

Industrial Process Heating

IPH facilities operate around the clock in the food and beverage, bricks and blocks, textiles, paper, and chemicals industries, among many others (Hasanuzzaman et al. 2012). The dominant power source for these facilities is the combustion of fossil fuels (Sharma et al. 2017). Renewable energy sources are increasingly cost-competitive with fossil fuels. In fact, most new renewable energy technologies are less expensive than new fossil fuel technologies (IRENA 2021). Incorporating renewables as sources of energy for IPH thus provides an opportunity for both economic and environmental benefits. However, the intermittency of renewables presents a significant obstacle to their integration into this industry (Kumar, Hasanuzzaman, and Rahim 2019).

Thermal Energy Storage

TES development has historically been dominated by technologies suitable for deployment with concentrating solar power (CSP) (Kuravi et al. 2013). State-of-the-art thermal storage deployed commercially with power tower CSP plants uses a 60%/40% $NaNO_3$/KNO_3 molten salt and operates between temperatures of approximately 280°C and 570°C using a two-tank configuration (Calvet et al. 2021). However, these nitrate salts are unsuitable for operation outside of this temperature range due to a high freezing point of approximately 220°C, and limits on high-temperature salt stability and corrosion resistance of containment alloys (Intrator and Dunn 2011). Other materials being investigated for TES include those based on: (1) sensible energy storage (various molten salt compositions, inert solid particles, rocks or pebble beds, sulfur, water, concrete, graphite, etc.), (2) latent energy storage in materials that undergo solid-liquid phase change at relevant temperatures (organic materials for low-temperature applications, inorganic salts and/or metals for high-temperature applications), or (3) thermochemical energy storage (hydrides, hydroxides, carbonates, metal oxides, etc.) (Alva, Lin, and Fang 2018; Sarbu and Sebarchievici 2018). The application temperature and challenges pertaining to storage material and/or containment cost, energy density, long-term thermal and cyclic stability, and charge/discharge heat transfer effectiveness drive material selection for a given IPH or electricity generation application (Hasnain 1998; Zhang et al. 2016).

Molten sulfur TES provides a low-cost solution to store and deliver thermal energy due to its low cost, high thermal and chemical stability, and high heat transfer rates due to natural convection dynamics within molten sulfur. Sulfur is a cheap commodity at $80/ton compared to $1100-1300/ton for conventional salts. When using a metric of storage cost per kWh, sulfur costs around 2-3 $/kWh. Sulfur is orders of magnitude cheaper than alternatives such as solar salt (44 $/kWh), HitecXL (22 $/kWh), and thermal oil (42 $/kWh), making this an economically viable thermal storage concept for industrial heat supply (Nithyanandam et al. 2018; Nithyanandam, Barde, and Wirz 2018; Wang et al. 2020).

The complex thermophysical properties variations in TES modeling presents various challenges to optimization and design. The cost of conducting large scale TES experiments is not a viable option (Mirzaei and Haghighat 2012; Nithyanandam, Barde, and Wirz 2018). Specifically, experiments specific to an application cannot be used for optimization or innovative design. Recent advancements in HPC scalable CFD codes have enabled large scale model support with high resolution and accuracy. Flexibility in controlling thermophysical property variations, model size, and scalability have made CFD a desirable tool among researchers in TES modeling (Barde et al. 2018; Nithyanandam et al. 2018). Hence, a transient thermo-fluid characterization CFD framework including all relevant property variability and us-

ing high-performance computing (HPC) resources is critical in optimizing the TES design for both cyclic performance and the lifetime cost. CFD can accelerate the adoption of TES by eliminating technical risks through full product-scale performance model development and simulations, followed by design optimization using the validated model. HPC CFD simulations and optimization of TES through numerical simulations will minimize design iterations, reduce product-to-market time, and accelerate adoption of TES (Kim, Samad, and Benini 2019).

Surrogate ML Modeling

The computational cost associated with full-scale CFD models of TES systems can be mitigated by the development of machine learning surrogate models (Morozova et al. 2022). Using data from the results of selected CFD simulations, ML models can be trained to emulate the utility of CFD (getting the state of a system over time for a set of operating conditions and design parameters) at a greatly reduced computational expense. Predicting output for a new input (i.e. inferencing) with surrogate ML models occurs on the scale of seconds, while full-scale CFD simulations can take hours or days (Duplyakin et al. 2022; Yaïci et al. 2013). This reduction in computational cost comes with limitations on the accuracy of the predictions. The accuracy of the CFD model is an upper limit on the accuracy of the surrogate model, and while the same CFD model can be used to predict behavior for any reasonable operating conditions, accurate inferencing with surrogate ML models is usually limited to scenarios within the scope of the training dataset (Gao et al. 2021). Thus, before developing a surrogate ML model, special care must be taken in building the CFD model and selecting a representative set of operating conditions (McBride and Sundmacher 2019).

After a training dataset has been developed for a surrogate model, an appropriate algorithm must be chosen (Sarker 2021). Though many such algorithms exist (Ray 2019), and many candidates were implemented and analyzed as a part of this work, only the most successful algorithms are presented here. These include the Random Forest, XGBoost, and Neural Network algorithms.

The Random Forest algorithm builds a set of randomly generated decision trees. Using a "wisdom of the crowd" process, the predictions of the trees are averaged and the result is output as the prediction of the model (Biau and Scornet 2016). The XGBoost algorithm is similar to a random forest, except the trees are generated in sequential boost rounds, and the results of the previous boost rounds are used to inform the creation of the trees in the next boost round (Chen and Guestrin 2016).

A Neural Network consists of multiple layers of nodes that are networked together between an input and an output layer (Jain, Mao, and Mohiuddin 1996). The data in the input layer is fed into the first layer of nodes, which transforms the data according to the weights of the connections between the nodes in the two layers. This data is then passed through weighted connections to the nodes in the next layer, and so on, until it is passed to the output layer, which gives a predicted value for the given input. The weights of the connections are iteratively optimized according to the accuracy of the prediction, with the objective of minimizing the difference between predictions and expectations.

Thermodynamic Constraints

The behavior of a thermodynamic system is constrained by the relationships between its physical properties. Surrogate ML models that predict multiple physical properties must explicitly enforce these constraints to ensure the predictions are physically realizable (Henze 2003). Relationships can be enforced during model training with soft or hard constraints (Chen and Zhang 2022), or after training by directly calculating properties from model predictions. Soft constraints encourage but do not require adherence to the relationship, typically by penalizing constraint violations with an additional item in the loss function. Hard constraints explicitly enforce the relationships within the model architecture (Xu and Darve 2022; Gao, Sun, and Wang 2021). Calculating a property from model predictions is only possible when the governing equation can be solved for the property of interest. Any deviation from the thermodynamic relationship in the ML training data must be rectified for these methods to be viable.

Design Optimization

Due to their short runtime, surrogate ML models can be used for iterative design optimization processes which would be computationally intractable with CFD models. These processes seek optimal design parameters to achieve desired system behavior by minimizing some objective function subject to a set of constraints. Many optimization algorithms are present in the literature, with gradient descent (Ruder 2016) and genetic (Whitley 1994) optimization algorithms among the most widely studied and used.

In the case of TES systems, among the desired behaviors is heat transfer effectiveness. Heat transfer effectiveness is dependent upon the physical state of the system. Developing surrogate ML models to predict physical states of the system, such as the temperature of the storage medium and heat transfer coefficient, can thus provide valuable insight to the process of developing models for direct use in design optimization.

Methods

Dataset creation

A set of CFD simulations were run for a simplified system, and the output of these simulations served as training data for the surrogate models. The geometry of the system was simplified to a pair of two-dimensional concentric pipes (see Figure 2). The heat transfer fluid (HTF) is in the inner pipe and the molten sulfur is in between the outer and inner pipes. Heat is transferred between HTF and sulfur at the wall of the inner pipe. During charging, the temperature of HTF is higher than the initial temperature of sulfur (T_i) and heat is transferred from HTF to sulfur. During discharging, the opposite is true and heat is transferred from sulfur to the HTF.

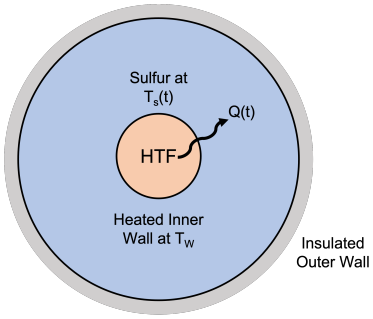

Figure 2: The sulfur TES system with a simplified geometry of two concentric pipes. Heat is transferred between HTF and sulfur at the wall of the inner pipe. The temperature at the wall of the inner pipe is T_w. The inner and outer pipes have a diameter of 4 cm and 20 cm, respectively.

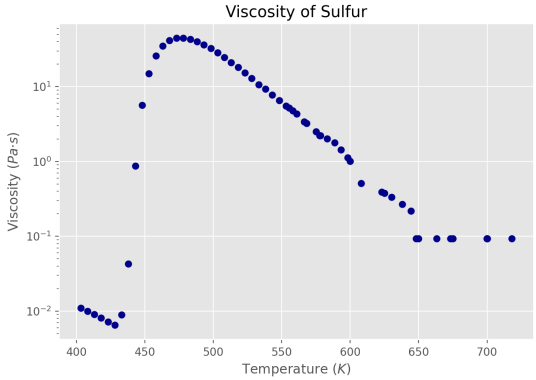

Figure 3: Variation of sulfur viscosity with temperature. A concentration of impurities of 35 ppmw was assumed.

While both charging and discharging are key to predicting overall system performance, the current study focuses on charging. Because the goal is to model over a wide range of operating conditions, both T_w and T_i were varied between 400 and 660 K while creating the data set used for training and testing of the system. Splitting this range into a grid by incrementing T_w and T_i by 20 K produced a representative set of 91 unique scenarios.

A major consideration in creating this data set is the variation in fluid properties. As shown in Figure 3, liquid sulfur has extreme variations in viscosity (Bacon and Fanelli 1943). The viscosity increases by approximately 4 orders of magnitude between 420 K and 480 K as rings of sulfur break into long chains, and then decreases by 3 orders of magnitude as the chains break apart. The data set developed in this work captures temperature ranges that include the entire range of behaviors this polymerization can create.

Developing initial models

Surrogate ML models were developed to predict the behavior of the sulfur TES system. Time-dependent tube-average heat transfer rates and bulk average sulfur temperature were extracted from the CFD simulation results. The heat transfer rates and sulfur average temperature were used to calculate heat transfer coefficient. The inputs to the model include T_i, T_w, and the time elapsed from the initial state (t). The output of the model is the average temperature of the sulfur or the heat transfer coefficient at time t. Of the many possible algorithms available to train the models under such formulations, Random Forest (RF), XGBoost (XGB), and Neural Network (NN) models were implemented. Long short-term memory (LSTM) neural network (Hochreiter and Schmidhuber 1997), physics-informed neural network (PINN) (Raissi, Perdikaris, and Karniadakis 2019), and recursive multi-step forecasting (Cheng et al. 2006) models were also implemented, and governing equations were predicted with a method based on sparse identification of nonlinear dynamic systems (de Silva et al. 2020), but none of these initial efforts were sufficiently fruitful to warrant their continued exploration.

The RF and XGB models were developed with the *sklearn* API (Pedregosa et al. 2011). Initially, the optimal number of trees built for each model was discovered with a simple grid search, and default values were used for all other hyperparameters. The NN model was developed with the Tensorflow framework (Abadi et al. 2015). A base model architecture was implemented with three dense layers of 50 hidden units and rectified-linear unit (ReLU) activation functions. The model was trained with either normalized or unnormalized data and the best method was chosen. Normalization consisted of scaling the data for each feature to have a mean of zero and a standard deviation of one.

The models were trained and validated using a five-fold cross-validation process. In this process, the dataset is split into five subsets. Training and validation is performed in five rounds. In each round, a different subset is selected as the validation set, and the other four sets are used for training. After training is complete, the model is validated using the validation set. The validation results are averaged across the five rounds to determine the average model performance.

Early stopping was implemented for the XGB and NN models to determine the optimal number of training boost rounds and epochs, respectively. With early stopping, the model is trained and validated during each training round. If the validation loss ceases to improve, the training process is stopped to prevent overfitting. The metric used to calculate the training and validation loss was root mean squared error (RMSE) between predictions and expectations. Early stopping is not possible with the RF model, as all decision trees are built simultaneously.

Calculating physical properties from predictions

The XGB, RF, and NN models were trained to predict either $T_s(t)$ or $h(t)$. With the predictions for one property, the other property could be calculated using the thermodynamic relationship between the two. This relationship arises from the process of equating the following formulas for heat transfer rate ($Q(t)$):

$$Q(t) = mc_p(T_s)\frac{dT_s}{dt} \tag{1}$$

$$Q(t) = h(t)A_s(T_w - T_s(t)) \tag{2}$$

Symbol	Description	Unit
i	Current timestep	None
A_s	Heat transfer surface area	m^2
m	Mass of sulfur	kg
$c_p(T_s)$	Specific heat of sulfur	$\frac{J}{kg \cdot K}$
Δt	Timestep length	s
t_δ	Hybrid model transition time	s

Table 1: Additional symbols used in this paper. Note the dependency of specific heat on sulfur average temperature.

Descriptions of the new symbols used in these and the following three equations are given in Table 1. Equating eq. 1 and 2 and solving for $T_s(t)$ leads to the following discrete formula for calculating $T_s(t)$ from predictions of $h(t)$:

$$T_{s,i+1} = T_{s,i} + \frac{h_i A_s (T_w - T_{s,i})}{m c_p(T_s)} \Delta t \tag{3}$$

In this equation, $T_{s,0}$ is provided by the initial temperature of sulfur (T_i).

Similarly, equating eq. 1 and 2 and solving for $h(t)$ leads to the following discrete formula for calculating $h(t)$ from predictions of $T_s(t)$:

$$h_i = \frac{m c_p(T_s)}{A_s (T_w - T_{s,i})} \frac{dT_s}{dt} \tag{4}$$

Using the centered-difference approximation for the time-derivative of $T_s(t)$ results in the following:

$$h_i = \frac{m c_p(T_s)}{A_s (T_w - T_{s,i})} \frac{T_{s,i+1} - T_{s,-1}}{2\Delta t} \tag{5}$$

By calculating one property from predictions for the other, the thermodynamic constraint between the two properties is satisfied.

Hybrid Model for T

A hybrid model was built to increase the accuracy of predictions for $T_s(t)$ while adhering to the thermodynamic constraint between $T_s(t)$ and $h(t)$. The hybrid model, represented in the diagram in Figure 4, consists of an NN model trained to predict $T_s(t)$ for $t \leq t_\delta$ and calculations of $T_s(t)$ for $t > t_\delta$ from an XGBoost model predictions of h (using eq. 3). The value of $T_s(t)$ predicted by the NN model for $t = t_\delta$ is provided as the initial point for calculating $T_s(t)$ from $h(t)$ predictions. The transition time t_δ was manually chosen to exceed the initial period of discrepancy between $T_s(t)$ and $h(t)$ discussed later.

Hyperparameter Optimization

Model hyperparameters were optimized with Optuna. Optuna uses a Tree-structured Parzen Estimator algorithm to optimize hyperparameters to minimize an objective function (Akiba et al. 2019). The metric used for the objective function was the RMSE of the model predictions.

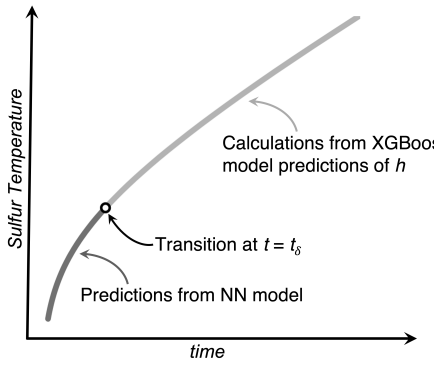

Figure 4: The hybrid model uses both predictions and calculations for $T_s(t)$. This figure illustrates the hybrid model approach in general rather than any specific timeseries.

	XGB	RF	NN
$T_s\ (K)$	11.103	11.994	**1.87**
$h\ (\frac{W}{m^2 K})$	**4.21**	6.88	11.103

Table 2: RMSE results for the preliminary XGBoost, Random Forest, and Neural Network models.

A single trial of the optimization process, called a study, consists of the algorithm selecting values for the hyperparameters, training the model with those values, and validating the model with five-fold cross-validation. As each trial finishes, Optuna uses the results of the previous trials to guide its choice of hyperparameter values for the next trial.

Evaluating the final models

The dataset for this model consisted of 91 scenarios. For the five-fold cross-validation (CV) process, the data was split into 80% training data and 20% validation data, which rounds to 73 training and 18 validation scenarios. For the final models, to improve confidence in the results, this process was repeated 20 times with the data randomly shuffled between each five-fold CV set. This results in a total of 100 different train-validation splits. Because the model is tested on each scenario once for each five-fold CV set, and this process was repeated twenty times, twenty unique results were available to determine the average performance of the model for each scenario.

Results

Initial model results

As shown in Table 2, the initial Neural Network model performed better than both the XGBoost and Random Forest models in predicting $T_s(t)$, while the XGBoost model outperformed the others in predicting $h(t)$.

These models were trained and tested on a limited dataset, as the CFD dataset creation process was not yet complete. The hyperparameters had not yet been optimized, with the exception of using a simple grid search to optimize the number of decision trees for the XGB and RF models. With the

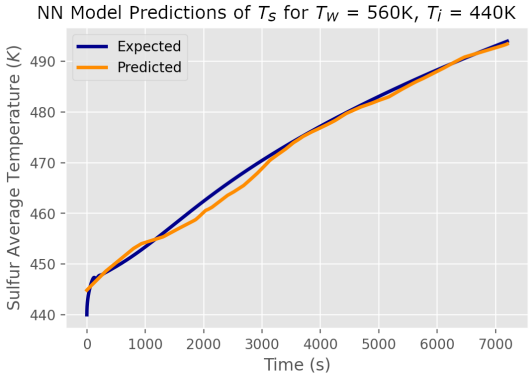

Figure 5: A sample result of NN model predictions for $T_s(t)$.

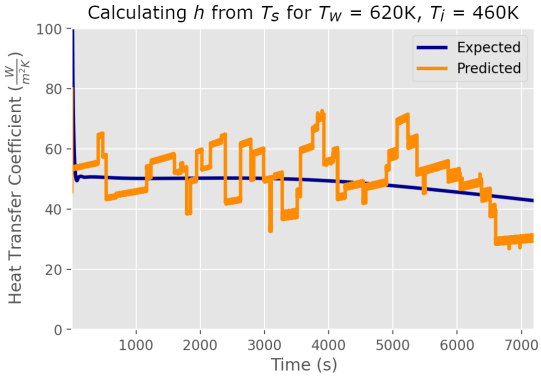

Figure 6: A sample result of calculating $h(t)$ from NN model predictions for $T_s(t)$.

	XGB_α	XGB_β	Neural Network	
learning rate	0.3	0.274	*learning rate*	.0005
subsample	0.1	0.494	*layers*	2
max depth	4	4	*hidden units*	72
gamma	0	0	*batch size*	13

Table 3: The final values of the hyperparameters for each model. The values reported for NN are for the short-run model. XGB_α is trained for $t \leq 360$, XGB_β—for $t > 360$.

NN model, normalizing the data produced better results for predicting $T_s(t)$, while unnormalized data produced better results for predicting $h(t)$. Based on these preliminary results, the NN model was chosen to predict $T_s(t)$ (see Figure 5) and the XGB model was chosen for $h(t)$.

Hyperparameter optimization

Hyperparameters were optimized for the XGB and NN models. The results of this process are displayed in Table 3. For the NN model, the hyperparameters were optimized to predict $T_s(t)$ for both a full run (up to 7200 seconds) and a short run (up to 360 seconds). The short-run NN model was developed for the hybrid model. For the XGBoost model, the hyperparameters were optimized to predict $h(t)$ for two separate cases: (1) for $t \leq 360s$ and (2) for $t > 360s$. The results of the two separate XGBoost models were combined to generate predictions of $h(t)$. Developing two separate models split at $t = 360s$ provided better results than did training a single XGBoost model for the entire simulation time and increased the accuracy of the hybrid model.

Calculating physical properties from predictions

The calculation of $h(t)$ from $T_s(t)$ suffers from errors due to the variance in the time derivative of $T_s(t)$ (see Figure 6). This variance is caused by erroneous fluctuations in the NN model predictions of $T_s(t)$. As the calculation of $h(t)$ depends on the time derivative of $T_s(t)$, small variance in the $T_s(t)$ prediction leads to large variance in the results of calculating $h(t)$. The calculations of $T_s(t)$ are free from this issue as $T_s(t)$ is dependent not on the derivative but on the integral of $h(t)$.

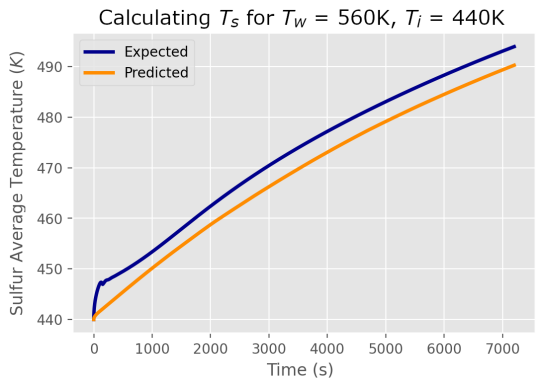

Figure 7: A sample result of calculating $T_s(t)$ from XGBoost model predictions for $h(t)$.

As can be seen in Figure 7, there is error in the calculation of $T_s(t)$ from $h(t)$ predictions during the initial phase of heat transfer. The CFD simulation data shows a sudden jump in $T_s(t)$ during the first few hundred seconds. This jump is not seen in the calculation of $T_s(t)$ from $h(t)$ predictions. This error is present even when calculating $T_s(t)$ directly from the $h(t)$ output of the CFD simulations.

The deviation between expected and calculated $T_s(t)$ during the initial portion of the simulation can likely be attributed to a combination of factors. The temperature distribution within the sulfur in the CFD model is non-uniform, particularly during the initial phase of the simulation when the sulfur temperature near the tube wall increases much more quickly than the bulk average sulfur temperature ($T_s(t)$). In addition to the sulfur viscosity variability illustrated in Figure 3, the sulfur heat capacity exhibits a sharp discontinuity at approximately 160 °C. Correspondingly, the sulfur heat capacity at the bulk sulfur average temperature may differ from the domain-averaged sulfur heat capacity, which contributes to errors in the predicted bulk sulfur temperature from equation 3 even if the heat transfer coefficient ($h(t)$) is exactly captured by the ML model.

As the value of $T_s(t)$ at a given timestep is dependent

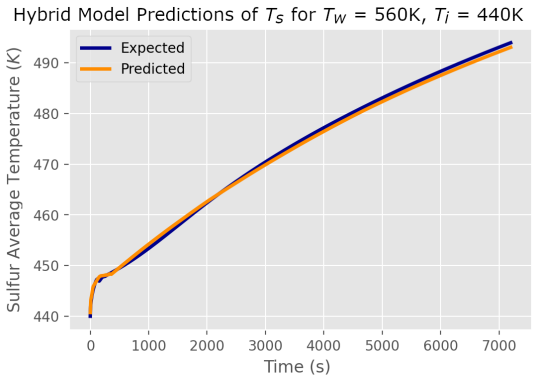

Figure 8: A sample result of hybrid model predictions of $T_s(t)$.

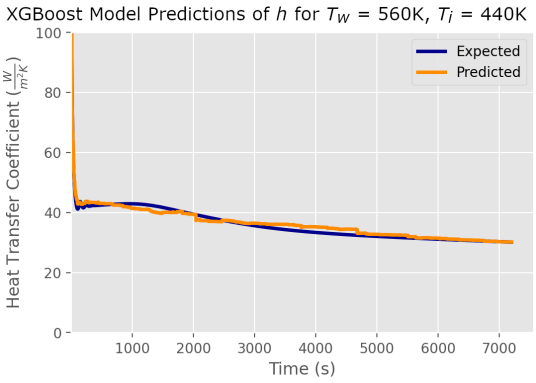

Figure 9: A sample result of the final XGBoost model predictions of $h(t)$. This is the model used to generate the predictions of $h(t)$ used to calculate $T_s(t)$ for the hybrid model.

on its value at the previous timestep, the error in the initial calculations is passed on to the calculations of $T_s(t)$ for the entire scenario. The pursuit of rectifying this error led to the development of the hybrid model.

Hybrid model results

The hybrid model was developed to overcome the error in calculating $T_s(t)$ from $h(t)$ while maintaining adherence to the thermodynamic constraint between them. The model demonstrates significant improvement in predicting $T_s(t)$ when compared to using calculations alone. Using the NN model to predict $T_s(t)$ for the first 360 seconds successfully eliminated the error apparent in Figure 7. Furthermore, the shape of the curve is much better captured by the hybrid model than by the NN model, which fluctuates irregularly and does not predict the characteristic sudden jump in $T_s(t)$ at the beginning of the modelled process. However, the RMSE of the hybrid model is higher than the NN model. This is the trade-off of the hybrid model: the hybrid predictions adhere to the thermodynamic constraint on the system at the cost of slightly lower accuracy in predicting $T_s(t)$.

The XGBoost model for $h(t)$ achieved an R^2 (coefficient of determination) of 0.949. The R^2-score quantifies how well the differences in expected values are matched by differences in predicted values, and the maximum value is 1. Though the XGBoost R^2-score is lower than the equivalent score for the $T_s(t)$ models, it is a high level of accuracy. As can be seen in Figure 9, the predictions of $h(t)$ closely match the expected values.

	XGB	**NN$_{short}$**	**NN$_{full}$**	**Hybrid**
RMSE	3.128 $\frac{W}{m^2K}$	0.345 K	1.202 K	2.210 K
R^2	0.949	0.99996	0.9995	0.9985

Table 4: Results for the final models. The NN$_{short}$ model was trained to predict $T_s(t)$ for $t \leq 360$ s. All other models were trained for the full simulation time of 7200 seconds.

Results of the per-scenario model evaluation are displayed in Figure 10. From these results, it is clear the XGBoost and Neural Network models performance declines at the periphery of the training dataset. This is not the case for the performance of the hybrid model, which presents no clear pattern.

Code and data

The final version of the paper will include the link to the repository with the publicly released data and code artifacts.

Discussion

The hybrid model was the result of a methodical process of developing and testing various ML architectures and techniques to provide accurate predictions for the 2-dimensional sulfur TES model. The initial models gave promising results but were only able to predict one property, and there was no enforcement of the thermodynamic constraint between $T_s(t)$ and $h(t)$. The need for this enforcement led this research down multiple paths, including directly calculating one property from predictions of the other.

Calculating $h(t)$ from $T_s(t)$ was especially prone to error, which is largely attributable to challenges inherent in using the time derivative of $T_s(t)$ to calculate $h(t)$. Small errors in $T_s(t)$ predictions are magnified into significant errors in its time derivative. The error in $h(t)$ is directly proportional to the error in the time derivative of $T_s(t)$. This error is likely to be repeated in any calculation involving a derivative of a predicted property. Calculating $T_s(t)$ from $h(t)$ was prone to error caused by a discrepancy inherent to the CFD calculation of these properties. Because the CFD model is the only source of training data for the surrogate model, this discrepancy is passed on in the form of error in the calculation of $T_s(t)$ from predictions of $h(t)$.

The process of developing the hybrid model, and the reasoning behind its development, is the main contribution of this work. Multiple properties are often of interest when developing CFD-emulating surrogate ML models. Enforcing constraints between these properties is the subject of much research, most recently including the development of physics-informed neural networks. The hybrid model enforces a hard constraint between $T_s(t)$ and $h(t)$ for $t >$ 360

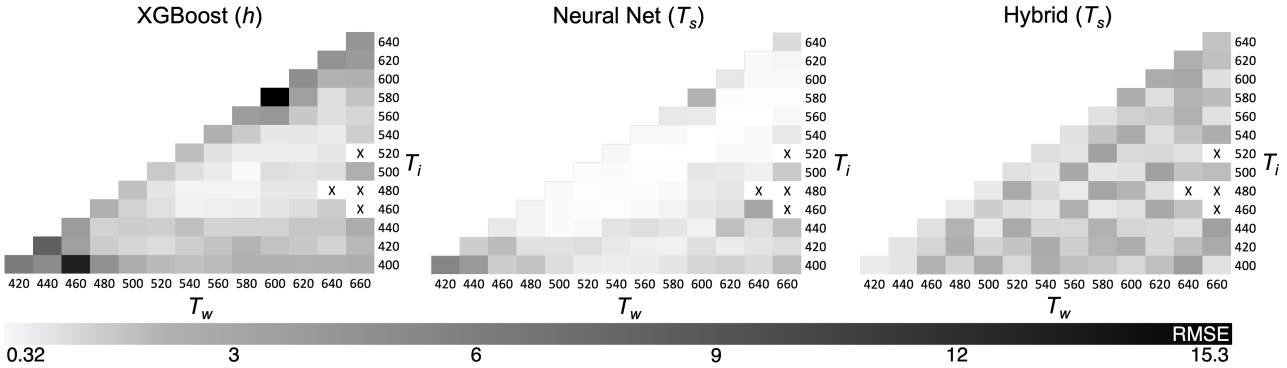

Figure 10: The results of the per-scenario evaluation of the final models. This evaluation process helps identify regions of high or low accuracy, such as the two outlier scenarios (in black) observed at (460,400) and (600,580) for the XGBoost model for $h(t)$. In the figure, *X* represents a failed CFD experiment and thus no data for the scenario.

s. For $t \leq 360$ s, the main source of discrepancy between $T_s(t)$ and $h(t)$ is a direct result of the manner of their calculation in the CFD model. The NN model trained for $t \leq 360$ s is highly accurate (see Table 4), which means any discrepancy caused by errors in this model is negligible.

Generally, surrogate ML models are evaluated by providing a chosen metric (usually RMSE or R^2) for the performance on the entire testing set or the average result for five- or ten-fold cross-validation. However, these methods do not provide details of the model performance for specific scenarios. The evaluation method visualised in Figure 10 gives actionable insight, especially for the XGBoost and NN models. For these models, there is less confidence in predictions for scenarios on the periphery of the representative dataset. If predictions for these operating conditions are critical, the range of the training dataset should be expanded.

Related Work

The state-of-the-art of CFD-based surrogate modeling is advancing on multiple fronts. Kaya et al. propose a new approach to CFD-based surrogate modeling to interpolate the volume-averaged hydrodynamic properties of a bioreactor (Kaya et al. 2022). Du, Zhu, and Wang leverage an ensemble of CFD simulations to train a surrogate model to make patient-specific local hemodynamic information predictions (Du, Zhu, and Wang 2022). Du, He, and Martins employed multilayer perceptron, recurrent neural networks, and a mixture of experts surrogate modeling approach based on CFD analysis for rapid optimization of airfoils (Du, He, and Martins 2021). Cho et al. developed a neural network surrogate model of CFD output to determine optimal operating conditions for explosive waste disposal (Cho et al. 2021).

Advances in surrogate modeling techniques include applications to thermal energy storage. Li et al. train a support vector machine surrogate model to emulate CFD analysis of stratified thermal energy storage tanks, and use the surrogate model in a genetic algorithm design optimization process (Li et al. 2022). Jin et al. develop a neural network-based surrogate model paired with a genetic algorithm optimization process to determine the optimal scenarios and locations for high-temperature reservoir thermal energy storage sites (Jin et al. 2022). Lee et al. develop a surrogate-model based model predictive control strategy to reduce operating costs of a thermal energy storage building energy system (Lee et al. 2021).

Conclusion

Sulfur TES is a promising new technology with the potential for facilitating the incorporation of renewable energy and waste heat into industrial process heating facilities. Machine learning models can be developed to emulate the utility of computational fluid dynamics methods in simulating the behavior of sulfur TES systems. These models, known as surrogate models, can then be incorporated into design optimization processes at a greatly reduced computational expense. Our research explored various avenues for developing surrogate models for sulfur TES systems. Random Forest, XGBoost, and Neural Network models were created to predict either the sulfur average temperature or the heat transfer coefficient of the system. Our experimental evaluation showed that the Neural Network model is the best approach for the former, while the Random Forest and XGBoost models are better choices for the latter. To enforce the thermodynamic constraint between the two properties, this research explored calculating one property from predictions of the other. We observed the substantial errors resulting from calculation of sulfur temperature from the heat transfer coefficient and vice versa. To overcome this issue, we built a hybrid model by combining direct predictions of sulfur temperature with calculations of sulfur temperature from predictions for heat transfer coefficient. This hybrid model is able to predict both quantities with high accuracy while maintaining adherence to the thermodynamic constraint.

Acknowledgments

This work was supported with funding from the US DOE HPC4EI program, additional funding from the US DOE AMO office through ORISE, and time on NREL's high performance computing system. The authors also want to thank Mohammad Rahimi for contributing to the CFD modeling.

References

Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M.; Jia, Y.; Jozefowicz, R.; Kaiser, L.; Kudlur, M.; Levenberg, J.; Mané, D.; Monga, R.; Moore, S.; Murray, D.; Olah, C.; Schuster, M.; Shlens, J.; Steiner, B.; Sutskever, I.; Talwar, K.; Tucker, P.; Vanhoucke, V.; Vasudevan, V.; Viégas, F.; Vinyals, O.; Warden, P.; Wattenberg, M.; Wicke, M.; Yu, Y.; and Zheng, X. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org.

Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; and Koyama, M. 2019. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2623–2631.

Al-Abidi, A. A.; Mat, S. B.; Sopian, K.; Sulaiman, M.; and Mohammed, A. T. 2013. CFD applications for latent heat thermal energy storage: a review. *Renewable and sustainable energy reviews*, 20: 353–363.

Alva, G.; Lin, Y.; and Fang, G. 2018. An overview of thermal energy storage systems. *Energy*, 144: 341–378.

Bacon, R. F.; and Fanelli, R. 1943. The viscosity of sulfur. *Journal of the American Chemical Society*, 65(4): 639–648.

Barde, A.; Jin, K.; Shinn, M.; Nithyanandam, K.; and Wirz, R. E. 2018. Demonstration of a low cost, high temperature elemental sulfur thermal battery. *Applied Thermal Engineering*, 137: 259–267.

Barde, A.; Nithyanandam, K.; Shinn, M.; and Wirz, R. E. 2020. Sulfur heat transfer behavior for uniform and non-uniform thermal charging of horizontally-oriented isochoric thermal energy storage systems. *International Journal of Heat and Mass Transfer*, 153: 119556.

Biau, G.; and Scornet, E. 2016. A random forest guided tour. *Test*, 25(2): 197–227.

Calvet, N.; Slocum, A. H.; Gil, A.; Grange, B.; Lahlou, R.; Hamer, T. T.; Diago, M.; Tetreault-Friend, M.; Codd, D. S.; Trumper, D. L.; et al. 2021. Dispatchable solar power using molten salt directly irradiated from above. *Solar Energy*, 220: 217–229.

Chen, T.; and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Chen, Y.; and Zhang, D. 2022. Integration of knowledge and data in machine learning. *arXiv preprint arXiv:2202.10337*.

Cheng, H.; Tan, P.-N.; Gao, J.; and Scripps, J. 2006. Multistep-ahead time series prediction. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 765–774. Springer.

Cheng, X.; Yang, Y.; and Huang, S. 2009. A simplified method for heat transfer prediction of supercritical fluids in circular tubes. *Annals of Nuclear Energy*, 36(8): 1120–1128.

Cho, S.; Kim, M.; Lyu, B.; and Moon, I. 2021. Optimization of an explosive waste incinerator via an artificial neural network surrogate model. *Chemical Engineering Journal*, 407: 126659.

de Silva, B. M.; Champion, K.; Quade, M.; Loiseau, J.-C.; Kutz, J. N.; and Brunton, S. L. 2020. Pysindy: a python package for the sparse identification of nonlinear dynamics from data. *arXiv preprint arXiv:2004.08424*.

Du, P.; Zhu, X.; and Wang, J.-X. 2022. Deep learning-based surrogate model for 3-dimensional patient-specific computational fluid dynamics. *Physics of Fluids*.

Du, X.; He, P.; and Martins, J. R. 2021. Rapid airfoil design optimization via neural networks-based parameterization and surrogate modeling. *Aerospace Science and Technology*, 113: 106701.

Duplyakin, D.; Beckers, K. F.; Siler, D. L.; Martin, M. J.; and Johnston, H. E. 2022. Modeling subsurface performance of a geothermal reservoir using machine learning. *Energies*, 15(3): 967.

Gao, H.; Sun, L.; and Wang, J.-X. 2021. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain. *Journal of Computational Physics*, 428: 110079.

Gao, K.; Yang, Y.; Zhang, T.; Li, A.; and Qu, X. 2021. Extrapolation-enhanced model for travel decision making: an ensemble machine learning approach considering behavioral theory. *Knowledge-Based Systems*, 218: 106882.

Hasanuzzaman, M.; Rahim, N.; Hosenuzzaman, M.; Saidur, R.; Mahbubul, I.; and Rashid, M. 2012. Energy savings in the combustion based process heating in industrial sector. *Renewable and Sustainable Energy Reviews*, 16(7): 4527–4536.

Hasnain, S. 1998. Review on sustainable thermal energy storage technologies, Part I: heat storage materials and techniques. *Energy conversion and management*, 39(11): 1127–1138.

Henze, G. P. 2003. An overview of optimal control for central cooling plants with ice thermal energy storage. *J. Sol. Energy Eng.*, 125(3): 302–309.

Hochreiter, S.; and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, 9(8): 1735–1780.

Intrator, J.; and Dunn, B. S. 2011. *2020 Strategic Analysis of Energy Storage in California: Final Project Report: Public Interest Energy Research (PIER) Program*. California Energy Commission.

IRENA. 2021. IRENA renewable power generation costs in 2020. *International Renewable Energy Agency: Abu Dhabi, UAE*.

Jain, A. K.; Mao, J.; and Mohiuddin, K. M. 1996. Artificial neural networks: A tutorial. *Computer*, 29(3): 31–44.

Jin, K.; Barde, A.; Nithyanandam, K.; and Wirz, R. 2019. Sulfur heat transfer behavior in vertically-oriented isochoric thermal energy storage systems. *Applied Energy*, 240: 870–881.

Jin, K.; and Wirz, R. E. 2020. Sulfur heat transfer behavior in vertically-oriented and nonuniformly-heated isochoric thermal energy storage systems. *Applied Energy*, 260: 114287.

Jin, W.; Atkinson, T.; Doughty, C.; Neupane, G.; Spycher, N.; McLing, T.; Dobson, P.; Smith, R.; and Podgorney, R. 2022. Machine-learning-assisted high-temperature reservoir thermal energy storage optimization. *Renewable Energy*.

Kaya, U.; Gopireddy, S.; Urbanetz, N.; Nopens, I.; and Verwaeren, J. 2022. Predicting the hydrodynamic properties of a bioreactor: Conditional density estimation as a surrogate model for CFD simulations. *Chemical Engineering Research and Design*, 182: 342–359.

Kim, K.-Y.; Samad, A.; and Benini, E. 2019. *Design Optimization of Fluid Machinery: Applying Computational Fluid Dynamics and Numerical Optimization*. John Wiley & Sons.

Kumar, L.; Hasanuzzaman, M.; and Rahim, N. 2019. Global advancement of solar thermal energy technologies for industrial process heat and its future prospects: A review. *Energy Conversion and Management*, 195: 885–908.

Kuravi, S.; Trahan, J.; Goswami, D. Y.; Rahman, M. M.; and Stefanakos, E. K. 2013. Thermal energy storage technologies and systems for concentrating solar power plants. *Progress in Energy and Combustion Science*, 39(4): 285–319.

Kurup, P.; and Turchi, C. 2015. Initial investigation into the potential of CSP industrial process heat for the Southwest United States. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Lee, D.; Ooka, R.; Matsuda, Y.; Ikeda, S.; and Choi, W. 2021. Experimental Investigation of Model Predictive Control for Thermal Energy Storage System Using Artificial Intelligence. In *2021 29th Mediterranean Conference on Control and Automation (MED)*, 961–966. IEEE.

Li, Y.; Deng, Y.; Niu, M.; Sun, D.; and Yu, B. 2022. Surrogate modeling and optimization for the unequal diameter radial diffuser of stratified thermal energy storage tanks. *Energy Science & Engineering*.

McBride, K.; and Sundmacher, K. 2019. Overview of surrogate modeling in chemical process engineering. *Chemie Ingenieur Technik*, 91(3): 228–239.

McMillan, C. A.; and Ruth, M. 2019. Using facility-level emissions data to estimate the technical potential of alternative thermal sources to meet industrial heat demand. *Applied Energy*, 239: 1077–1090.

Mirzaei, P. A.; and Haghighat, F. 2012. Modeling of phase change materials for applications in whole building simulation. *Renewable and Sustainable Energy Reviews*, 16(7): 5355–5362.

Morozova, N.; Trias, F. X.; Capdevila, R.; Schillaci, E.; and Oliva, A. 2022. A CFD-based surrogate model for predicting flow parameters in a ventilated room using sensor readings. *Energy and Buildings*, 266: 112146.

Nithyanandam, K.; Barde, A.; Lakeh, R. B.; and Wirz, R. 2018. Charge and discharge behavior of elemental sulfur in isochoric high temperature thermal energy storage systems. *Applied Energy*, 214: 166–177.

Nithyanandam, K.; Barde, A.; and Wirz, R. 2018. Heat transfer behavior of elemental sulfur for low temperature thermal energy storage applications. *International Journal of Heat and Mass Transfer*, 127: 936–948.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12: 2825–2830.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378: 686–707.

Ray, S. 2019. A quick review of machine learning algorithms. In *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 35–39. IEEE.

Ruder, S. 2016. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

Sarbu, I.; and Sebarchievici, C. 2018. A comprehensive review of thermal energy storage. *Sustainability*, 10(1): 191.

Sarker, I. H. 2021. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3): 1–21.

Schoeneberger, C. A.; McMillan, C. A.; Kurup, P.; Akar, S.; Margolis, R.; and Masanet, E. 2020. Solar for industrial process heat: A review of technologies, analysis approaches, and potential applications in the United States. *Energy*, 206: 118083.

Sharma, A. K.; Sharma, C.; Mullick, S. C.; and Kandpal, T. C. 2017. Solar industrial process heating: A review. *Renewable and Sustainable Energy Reviews*, 78: 124–137.

Wang, Y.; Barde, A.; Jin, K.; and Wirz, R. 2020. System performance analyses of sulfur-based thermal energy storage. *Energy*, 195: 116996.

Whitley, D. 1994. A genetic algorithm tutorial. *Statistics and computing*, 4(2): 65–85.

Xu, K.; and Darve, E. 2022. Physics constrained learning for data-driven inverse modeling from sparse observations. *Journal of Computational Physics*, 453: 110938.

Yaïci, W.; Ghorab, M.; Entchev, E.; and Hayden, S. 2013. Three-dimensional unsteady CFD simulations of a thermal storage tank performance for optimum design. *Applied Thermal Engineering*, 60(1-2): 152–163.

Zhang, H.; Baeyens, J.; Caceres, G.; Degreve, J.; and Lv, Y. 2016. Thermal energy storage: Recent developments and practical aspects. *Progress in Energy and Combustion Science*, 53: 1–40.