

Research Article

Predictive Power of Machine Learning for Optimizing Solar Water Heater Performance: The Potential Application of High-Throughput Screening

Hao Li,^{1,2} Zhijian Liu,³ Kejun Liu,⁴ and Zhien Zhang⁵

¹Department of Chemistry, The University of Texas at Austin, 105 E. 24th Street, Stop A5300, Austin, TX 78712, USA

²Institute for Computational and Engineering Sciences, The University of Texas at Austin, 105 E. 24th Street, Stop A5300, Austin, TX 78712, USA

³Department of Power Engineering, School of Energy, Power and Mechanical Engineering, North China Electric Power University, Baoding 071003, China

⁴Department of Computer Science, Rice University, 6100 Main Street, Houston, TX 77005-1827, USA

⁵School of Chemistry and Chemical Engineering, Chongqing University of Technology, Chongqing 400054, China

Correspondence should be addressed to Hao Li; lihao@utexas.edu

Received 12 July 2017; Revised 6 August 2017; Accepted 5 September 2017; Published 24 September 2017

Academic Editor: Zhonghao Rao

Copyright © 2017 Hao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Predicting the performance of solar water heater (SWH) is challenging due to the complexity of the system. Fortunately, knowledge-based machine learning can provide a fast and precise prediction method for SWH performance. With the predictive power of machine learning models, we can further solve a more challenging question: how to cost-effectively design a high-performance SWH? Here, we summarize our recent studies and propose a general framework of SWH design using a machine learning-based high-throughput screening (HTS) method. Design of water-in-glass evacuated tube solar water heater (WGET-SWH) is selected as a case study to show the potential application of machine learning-based HTS to the design and optimization of solar energy systems.

1. Introduction

How to cost-effectively design a high-performance solar energy conversion system has long been a challenge. Solar water heater (SWH), as a typical solar energy conversion system, has complicated heat transfer and storage properties that are not easy to be measured and predicted by conventional ways. In general, an SWH system uses solar collectors and concentrators to gather, store, and use solar radiation to heat air or water in domestic, commercial, or industrial plants [1]. For the design of high-performance SWH, the knowledge about correlations between the external settings and coefficients of thermal performance (CTP) is required. However, some of the correlations are hard to know for the following reasons: (i) measurements are time-consuming [2]; (ii) control experiments are usually difficult to perform; and (iii) there is no current physical model that can precisely

connect the relationships between external settings and intrinsic properties for SWH. Currently, there are some state-of-the-art methods for the estimation of energy system properties [3–5] and for the optimization of performances [6–11]. However, most of them are not suitable for the solar energy system. These problems, together with the economic concerns, significantly hinder the rational design of high-performance SWH.

Fortunately, machine learning, as a powerful technique for nonlinear fitting, is able to help us precisely acquire the values of CTP with the knowledge of some easy-measured independent variables. With a sufficiently large database, a machine learning technique with appropriate algorithms can “learn” from the numerical correlations hidden in the dataset via a nonlinear fitting process and perform precise predictions. With such a technique, we do not need to exactly find out the physical models for each CTP and can directly

acquire a precise prediction with a well-developed predictive model. During the past decades, Kalogirou et al. have done a large number of machine learning-based numerical predictions of some important CTPs for solar energy systems [12–19]. Their results show that there is a huge potential application of machine learning techniques to energy systems. Based on their successful works, we recently developed a series of machine learning models for the predictions of heat collection rates (daily heat collection per square meter of a solar water system, MJ/m²) and heat loss coefficients (the average heat loss per unit, W/(m³K)) to a water-in-glass evacuated tube solar water heater (WGET-SWH) system [2, 20, 21]. Our results show that with some easy-measured independent variables (e.g., number of tubes and tube length), both heat collection rates and heat loss coefficients can be precisely predicted after some proper trainings from the datasets, with proper algorithms (e.g., artificial neural networks (ANNs) [2, 20], support vector machine (SVM) [2], and extreme learning machine (ELM) [21]). An ANN-based user-friendly software was also developed for quick measurements [20]. These novel machine learning-assisted measurements dramatically shorten the measurement period from weeks to seconds, which has good industrial benefits. However, all the machine learning studies mentioned here are only the predictions and/or measurements. So far, very few industries really put these methods into practical applications. To the best of our knowledge, very few references concern about the optimization of thermal performance of energy systems using such a powerful knowledge-based technique [22]. To address this challenge, we recently used a high-throughput screening (HTS) method combined with a well-trained ANN model to screen 3.5×10^8 possible designs of new WGET-SWH settings, in good agreement with the subsequent experimental validations [23]. This is so far the first trial of HTS to a solar energy system design. The HTS method (roughly defined as the screening of the candidates with the best target properties using advanced high-throughput experimental and/or computational techniques) has already been widely used in biological [24–28] and computational [29–31] areas. With the basic concept that screening thousands or even millions of possible cases to discover the candidates with the best target functions or performances, HTS helps people dramatically reduce the required regular experiments, saving much economic cost and manpower.

In this paper, we aim to propose an HTS framework for optimizing a solar energy system. Picking SWH as a case study, we show how this optimization strategy can be applied to a novel solar energy system design. Different from the study by Liu et al. [23], this paper shows the predictive power of machine learning and the development of a general HTS framework. Instead of listing tedious mathematical works, in this paper, we provide vital details about the general modeling and HTS process. Since tube solar collectors have a substantially lower heat loss coefficient than other types of collectors [12, 32], WGET-SWHs gradually become popular during the past decades [33–35], with the advantages of excellent thermal performance and easy transportability [36, 37]. With this reason, we chose the WGET-SWH system

as a typical SWH, to show how a well-developed ANN model can be used to cost-effectively optimize the thermal performance of an SWH system, using an HTS method.

2. Machine Learning Methods

2.1. Principles of an ANN. There are various machine learning algorithms that have been effectively applied to the prediction of properties for energy systems, such as ANN [12, 13, 17, 18, 20, 38], SVM [20, 39, 40], and ELM [21, 41]. Because the ANN method is the most popular algorithm for numerical predictions [42], we only introduce the basic principle of ANN here. A general schematic ANN structure is shown in Figure 1, with the input, hidden, and output layers constructed by certain numbers of “neurons.” Each neuron (also called a “node”) in the input layer, respectively, represents a specific independent variable. The neuron in the output layer represents the dependent variable that is needed to be predicted. Usually, the independent variables should be the easy-measured variables that have a potential relationship with the dependent variable. The dependent variable is usually the variable that is hard to be detected from experiments and is expected to be precisely predicted. The layer between the input and output layers shown in Figure 1 is the hidden layer. The optimal number of neurons in the hidden layer depends on the study object and the scale of the dataset. Each neuron connects to all the neurons in the adjacent layer, with the connection called the weight (usually represented as w), which directly decides the predictive performance of the ANN, using the activation functions. For the training of an ANN, the initial weights will be first selected randomly, and then following iterations would help find out the optimal weight values that fulfill the prediction criterions. All the data move only in the same direction (from left to right, as shown in Figure 1). A well-trained ANN should consist of the optimal numbers of hidden layer neurons, hidden layer(s), and weight values, which sufficiently avoid the risk of either under- or overfitting. In practical applications, there is a large number of neural networks with modified algorithms, such as ELM [43–45], back-propagation neural network (BPNN) [46–48], and general regression neural network (GRNN) [49–51]. Though there are various network models, the basic principles for model training are similar.

2.2. Training of an ANN. To train a robust ANN, several factors should be considered: (i) percentages of the training and testing sets; (ii) number of hidden neurons; (iii) number of hidden layers; and (iv) required time for training. When training a practical ANN for real applications, a large training set is recommended. For predicting the heat collection rates of WGET-SWHs, we found that with a relatively large dataset (>900 data groups), the training set higher than 85% could help acquire a model with good predictive performance in the testing set [2]. Another reason to use a large training set is that if the training set percentage is small, it would be a waste of data for practical applications. The reason is simple: more data groups for training would usually lead to a better predictive performance. For the selection of the number of

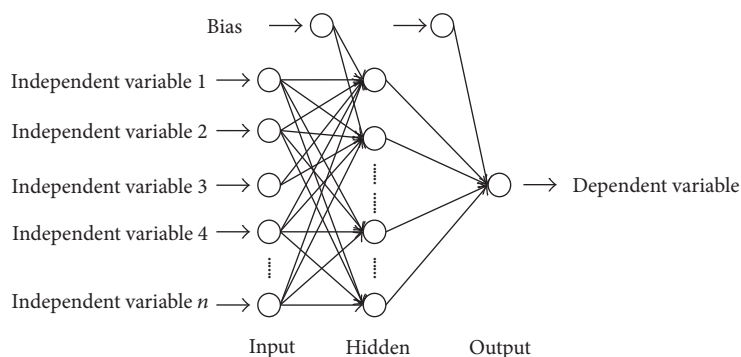


FIGURE 1: Schematic structure of a typical ANN. Circles represent the neurons in the algorithm.

hidden neurons, it is quite important to try the neuron numbers from low to high. If the number of hidden neurons is not enough, there would be a risk of underfitting; if it is too many, there would be a risk of overfitting and time-consuming. Therefore, finding the best number of neurons by comparison is particularly important. It should be noted that in some special neural network methods (e.g., GRNN), the number of hidden neurons can be a fixed value once the dataset is defined in some software packages. Under this circumstance, it is no longer necessary to worry about the hidden neuron settings. In addition to the hidden neuron numbers, same tests should be done on the number of layers, in order to avoid either under- or overfitting. The last factor we need to consider is the training time. According to the basic principle of an ANN (Figure 1), the interconnection among neurons would become more complicated with higher numbers of neuron. Therefore, with larger database and larger numbers of independent variables and hidden neurons, the training time would be longer. This means that sometimes an ordinary personal computer (PC) cannot sustain a tedious cross-validation test. From our previous studies with an ANN training [2, 51], we found that if the database was sufficiently large, repeated training and/or cross-validation training would lead to insignificant fluctuation. In other words, for practical applications, the ANN training and testing results would be robust if the database is large, and so a cross-validation process can be rationally skipped after a simple sensitivity test, in order to save computational cost.

2.3. Testing of an ANN. Using a testing result with an ANN for the prediction of heat collection rate as an example (Figure 2), we can see that a well-trained ANN can precisely predict the heat collection rates of the data in the testing set, with relatively low absolute residual values. Though there are still deviations exist in some predicted points, the overall accuracy is still relatively high and acceptable to practical applications. It should be noted that for a solar energy system, the independent variables for modeling should always include some environmental variables, such as solar radiation intensity and ambient temperature [2]. These variables are highly dependent to the external temperature, location, and season. That is to say, the external conditions of the predicted data should be in the similar environmental

conditions as the data used for the model training. Otherwise, the ANN may not perform good predictive performance. In all of our recent studies, all the data measurements were performed in very similar season, temperature, and location, which can sufficiently ensure precise predicted results in both the testing set and subsequent experimental validation.

3. High-Throughput Screening (HTS)

The basic idea of computational HTS is simple: the calculations of all possible systems in a certain time period (using fast algorithms) and the screening of the candidates with target performances. Previously, Greeley et al. used density functional theory (DFT) calculations to screen and design high-performance metallic catalysts for hydrogen evolution reaction via an HTS method, in good agreement with experimental validations [29]. Hautier et al. combined DFT calculations, machine learning, and HTS methods to predict the missing ternary oxide compounds in nature and develop a completed ternary oxide database [31], which shows that a machine learning-assisted HTS process can be precisely used for new material prediction and discovery. However, though the HTS method has been widely used in many areas, its conceptual applications to energy system optimization is not reported during the past decade.

Very recently, our studies show that the machine learning-assisted HTS process can be effectively performed on the optimization of solar energy system [23]. Choosing WGET-SWH as a case study, our results show that an HTS process with a well-trained ANN model can be used for the optimization of heat collection rate of SWH. The first step was to generate an extremely large number of independent variable combinations (3.5×10^8 possible design combinations) as the input of a well-trained ANN model. The heat collection rates of all these combinations were then, respectively, predicted by the ANN. After that, the new designs with high predicted heat collection rates were recorded as the candidate database. For validation, we installed two screened cases and performed rigorous measurements. The experimental results showed that the two selected cases had higher average heat collection rates than all the existing cases in our previous measurement database. Being similar to a previous chemical HTS concept proposed by Pyzer-Knapp et al. [52], we reconstruct the process of this

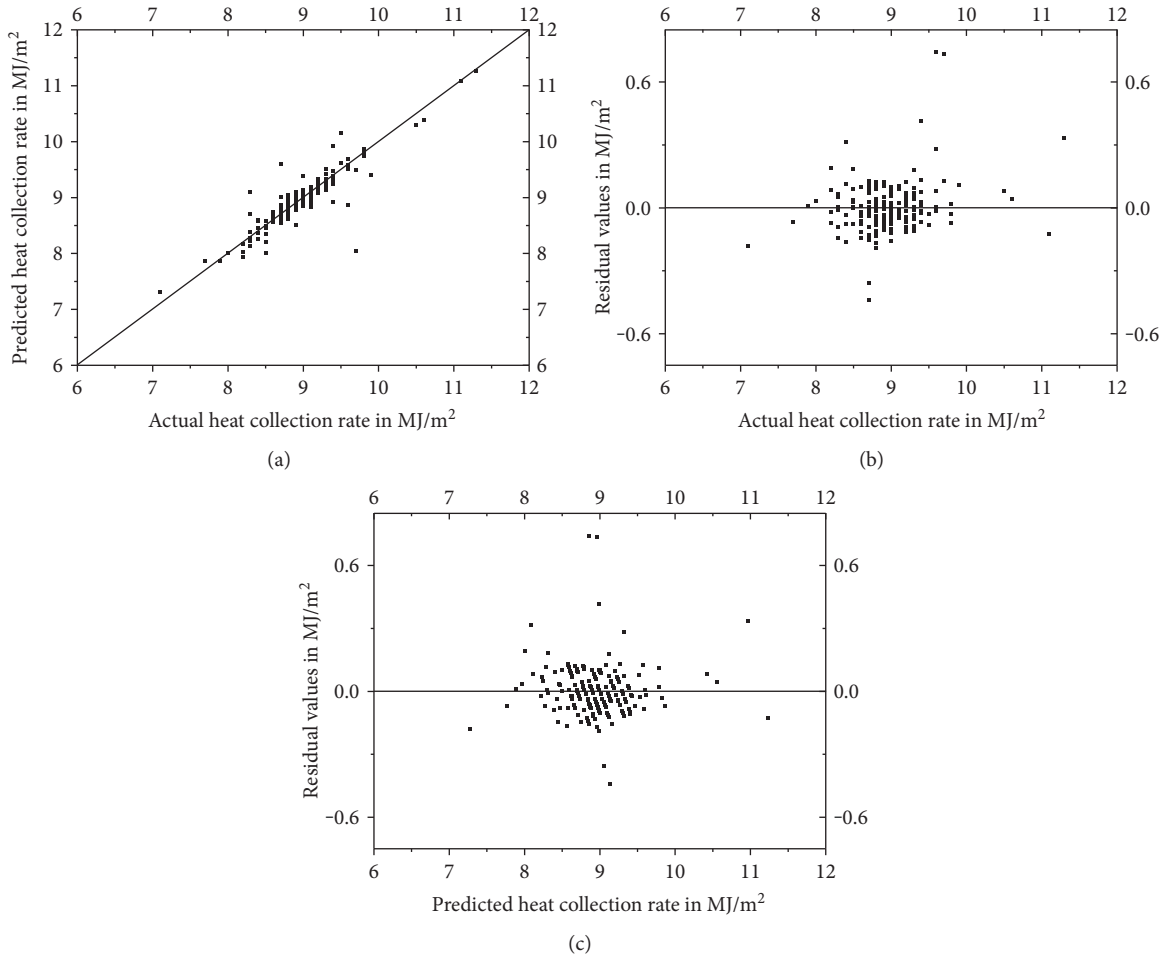


FIGURE 2: Testing results using an ANN model for the prediction of heat collection rate for WGET-SWHs. (a) Predicted values versus actual values; (b) residual values versus actual values; and (c) residual values versus predicted values. Reproduced with permission from Liu et al. [2].

optimization method, as shown in Figure 3. More modeling and experimental details can be found in [23].

4. HTS-Based Optimization Framework

Based on the recent trials on the HTS-based optimization method to the SWH system, here, we propose a framework for the design and optimization of solar energy systems. Though the machine learning-based HTS method is a quick design strategy, the preconditions should be fulfilled rigorously. That means, two vital conditions should be fulfilled, including (i) a well-trained machine learning model and (ii) a rational generation of possible inputs.

4.1. A Well-Trained Machine Learning Model. To acquire a well-trained machine learning model, in addition to the regular training and testing processes as shown in Sections 2.2 and 2.3, another key step is to define the independent variables for training. Since the dependent variable is usually the quantified performance of the energy system, the selection of an independent variable which has potential relationships with the dependent variable would directly decide the predictive precision of the model. In our previous case [23],

we chose seven independent variables as the inputs, including tube length, number of tubes, tube center distance, tank volume, collector area, final temperature, and tilting angle (the angle between tubes and the ground). A 3-D schematic design of a WGET-SWH system is shown in Figure 4 [23], which shows that only with these independent variables can we reconstruct a WGET-SWH system quickly with some other minor empirical settings. Unlike a physical model (which requires rigorous mathematical deduction and hypothesis), machine learning does not require the user to know exactly about the potential relationships between the independent and dependent variables. This feature also leads to the fact that machine learning prediction method is more flexible than conventional methods. From these seven inputs, we can see that except for the final temperature, all the other six variables are the important parameters of a WGET-SWH. In terms of the final temperature, we found that this is extremely important to ensure a precise model for heat collection rate prediction. The reason is simple: the heat collection performance of a WGET-SWH is not only decided by the mechanical settings of the system but also depends on the environmental conditions such as solar radiation intensity, ambient temperature, and the final temperature. Since

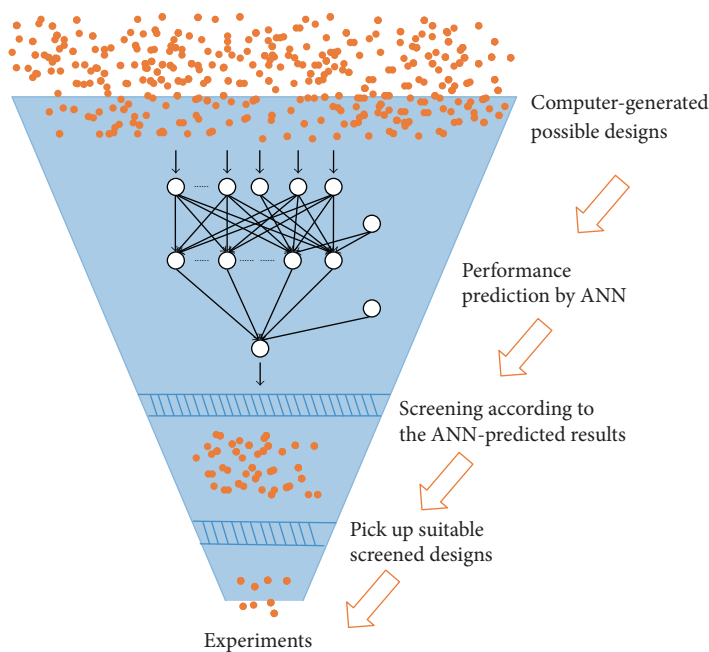


FIGURE 3: An HTS process for solar energy system optimization. Each orange circle represents a possible design.

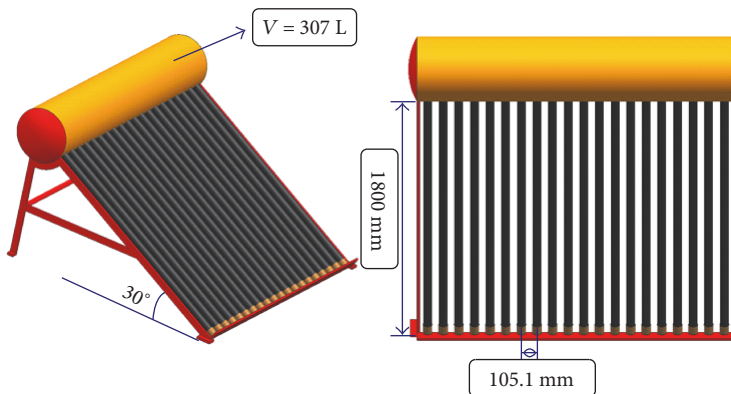


FIGURE 4: A 3-D schematic design for WGET-SWH installation. Reproduced with permission from Liu et al. [23].

the solar radiation intensity correlates well with the final temperature in a nonphotovoltaic heat transfer system, and it is not easy to be measured, we did not consider this as a variable for model training. Also, because the ambient temperatures are very similar during the measurements of all the SWHs in our database (we performed all the measurements in the similar months and locations), we also removed it from the variable list. It should be noted that for the measurements gathered from various seasons and unstable weathers, the ambient temperature sometimes is important and should not be neglected for modeling. Results show that without the solar radiation intensity and ambient temperature, our predictive models were still precise and robust enough [2]. Reducing the number of independent variables like these not only helps us dramatically reduce the required time for model training but also simplifies the input generation process at the following

HTS application. Another vital step is the scale and size of the database. Due to the complexity of the energy collection and transfer system, there are usually a large number of independent variables. To ensure a good training, a large and wide database should be used. If the size of the database for training is too small, it would generate high error rates during fittings; if the range of database is too narrow, the trained model would only have good performance in a very local data range, scarifying the precision of the data in a relatively remote region. In many previous cases, we can see that a large and wide database is crucial to ensure a good practical prediction [53]. In our case study, the ranges of the independent variables were wide enough to ensure a good predictive performance of the ANN [2]. Detailed descriptive statistics (maximum, minimum, data range, average value, and standard deviation) of the WGET-SWH database we used for training are shown in Table 1.

TABLE 1: Descriptive statistics of the variables for 915 samples of in service WGET-SWHs. Reproduced with permission from Liu et al. [2].

Items	Tube length (mm)	Number of tubes	TCD (mm)	Tank volume (kg)	Collector area (m ²)	Angle (°)	Final temp. (°C)	HCR
Maximum	2200	64	151	403	8.24	85	62	11.3
Minimum	1600	5	60	70	1.27	30	46	6.7
Data range	600	59	91	333	6.97	55	16	4.6
Average value	1811	21	76.2	172	2.69	46	53	8.9
Standard deviation	87.8	5.8	5.11	47.0	0.73	3.89	2.0	0.48

TCD: tube center distance; final temp.: final temperature; HCR: heat collection rate (MJ/m²). Tank volume was defined as the maximum mass of water in tank (kg).

TABLE 2: Number of selected values of different independent variables (extrinsic properties). Reproduced with permission from Liu et al. [23].

	Tube length (mm)	Number of tubes	TCD (mm)	Tank volume (kg)	Collector area (m ²)	Angle (°)	Final temp. (°C)
Number of selected values	5	30	5	111	50	5	17

TCD: tube center distance; final temp.: final temperature.

TABLE 3: Input variables of two newly designed WGET-SWHs. Reproduced with permission from Liu et al. [23].

	Tube length (mm)	Number of tubes	TCD (mm)	Tank volume (kg)	Collector area (m ²)	Angle (°)	Final temp. (°C)
Design A	1800	18	105.5	163	1.27	30	52–62
Design B	1800	20	105.5	307	1.27	30	52–62

TCD: tube center distance; final temp.: final temperature.

TABLE 4: Measured heat collection rates (MJ/m²) of the two novel designs. All the measurements were performed under the environmental conditions similar to those of the measurements for the previous database (Table 1). Reproduced with permission from Liu et al. [23].

	Day 1	Day 2	Day 3	Day 4	Average	Predicted	Error rate
Design A	11.38	11.26	11.34	11.29	11.32	11.47	1.35%
Design B	11.47	11.43	11.42	11.45	11.44	11.66	1.90%

4.2. A Rational Generation of Possible Inputs. A rational generation of inputs of the ANN during the HTS process is also crucial to ensure a quick HTS with less time consumption. Without a rational criterion, there will be infinite possible combinations, which will lead to infinite computational cost. In our current study, we found that a quick way is to generate the inputs according to the trained weights of each independent variables: the independent variable with a higher numerical weight of the model will be assigned more possible values as the input of ANN during prediction. The basic assumption is that a larger value of weight will lead to a more significant change to the predicted results. In Liu et al. [23], we show that the tank volume has the highest weight to determine the heat collection rate, which also qualitatively agreed with the empirical knowledge. Thus, we generated more inputs of tank volume with different numerical values for the HTS

process. Table 2 shows the numbers of selected values of independent variables for screening the optimized WGET-SWHs via an HTS process [23]. Except for the final temperature, the number of values of all the independent variables was assigned according to their sequences of weight after a typical and robust ANN training. In terms of the inputs of final temperature, since it is not a part of the SWH installation, we consider all its possible integers shown in the database (Table 1) as the inputs for HTS. It should be noted that the weight values of a trained ANN do not contain exact physical meanings because the initial weights for an ANN training were usually selected randomly. Multiple trainings of ANN will lead to different final weight values. Thus, in addition to referring to the trained weight values, sometimes we should artificially assign more possible input values for the independent variables that are physically more influential to the predicted results. For other weight-free algorithms (e.g., SVM), artificial choices for inputs are particularly important.

4.3. Experimental Validation. With the inputs of the generated independent variable values, the machine learning model is able to output their predicted heat collection rates in an extremely short timescale. After screening, those designs with high predicted heat collection rates can be recorded as the candidates for future applications. In our recent studies, two typical designs after an HTS process were selected for experimental installations, with their independent variables summarized in Table 3.

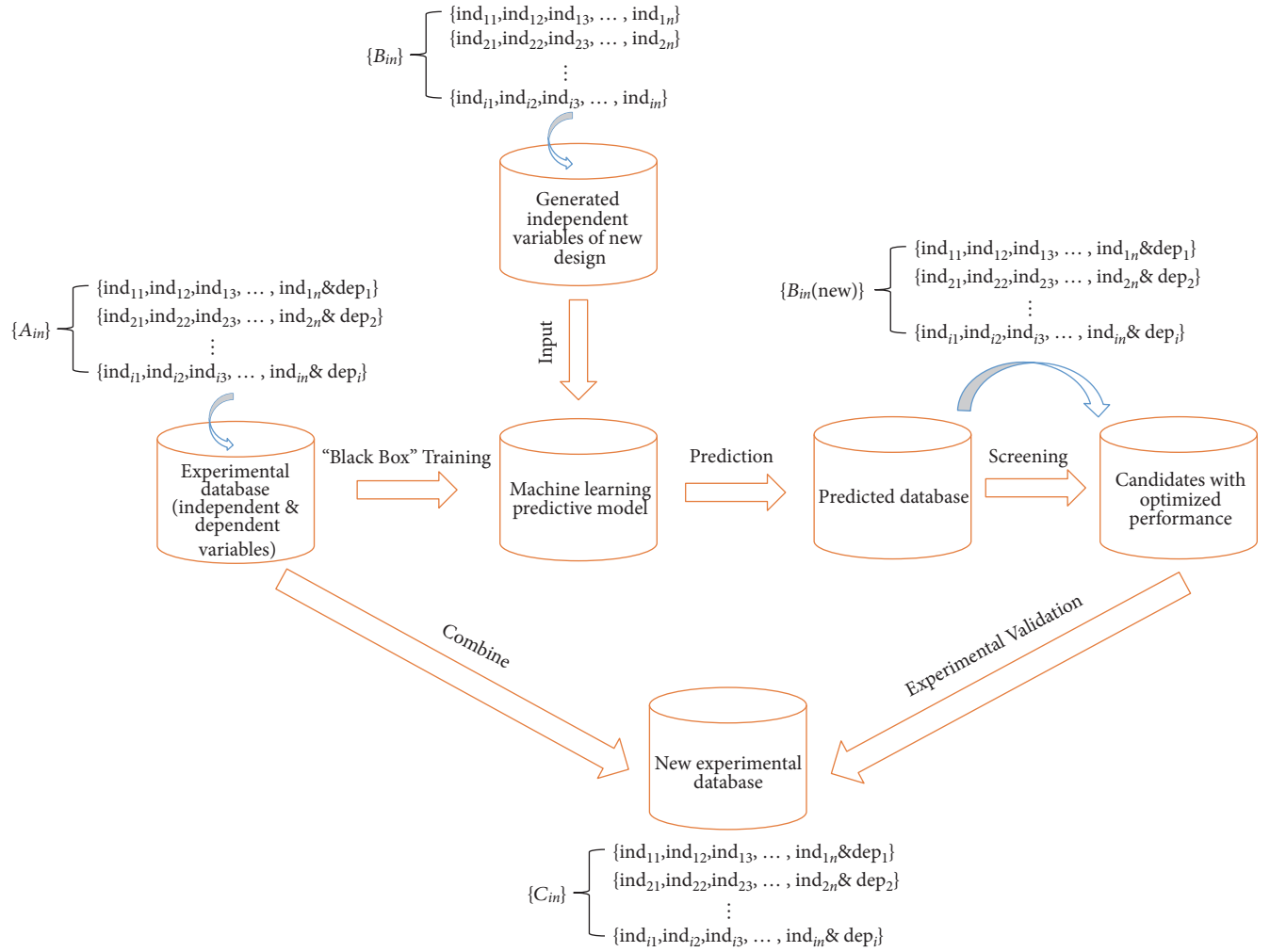


FIGURE 5: A proposed framework of machine learning-assisted HTS process for target performance optimization. Independent variables are assigned as “ind.” Dependent variables are assigned as “dep.” $\{A_{in}\}$ represents the original experimental database. $\{B_{in}\}$ represents the generated independent variables as the inputs. $\{B_{in}(\text{new})\}$ represents the generated independent variables and their predicted dependent variables. $\{C_{in}\}$ represents the new experimental database combining the original experimental database and the experimental validation results of the screened candidates.

Rigorous experimental measurements on these two new designs validated that both of them outperformed all our 915 WGET-SWHs in the previous database under similar environmental conditions (Table 4). More comparative results are shown in [23].

4.4. A Framework for HTS-Based Optimization. The proposed framework for HTS-based optimization mainly consists of two parts: (i) developing a predictive model and (ii) screening possible candidates. The machine learning model is described as a “black box” in this framework since we do not need to know what really happens inside the training for real applications (and usually we care more about the fitting results). The concrete algorithmic and experimental processes of the proposed framework can be summarized as follows:

Step 1: Select the independent and dependent variables for the machine learning model.

Step 2: Train and test a predictive machine learning model with a proper experimental database.

Step 3: Generate a large number of the combinations of independent variable values.

Step 4: Input the generated independent variables into the well-trained predictive model.

Step 5: Screen and record the outputted dependent variable values and their corresponding independent variable values that fulfill all the screening criterions.

Step 6: Select the candidates from the results of Step 5 for experimental validation.

Step 7: Record the experimental results from Step 6.

To sum up, the proposed framework is shown in Figure 5. It can be seen that once all the preconditions of the

“cylinders” discussed above are fulfilled, a completed machine learning-assisted process can be achieved. The ultimate goal of the screening is to find out better candidates with optimized target performance. These candidates will have the independent variables different (or partially different) from the previous experimental database. Combining the previous experimental database with the experimental validation on new designed candidates, we can construct a new experimental database with more informative knowledge for future applications. It should be noted that this framework not only works for solar energy systems but also works for the optimization cases of other devices. We expect that this framework can be expanded to other optimization demands in the future.

5. Conclusions

In this paper, we have summarized our recent studies on the predictive performance of machine learning on an energy system and proposed a framework of SWH design using a machine learning-based HTS method. This framework consists of (i) developing a predictive model and (ii) screening possible candidates. A combined computational and experimental case study on WGET-SWH shows that this framework can help efficiently design new WGET-SWH with optimized performance without knowing the complicated knowledge of the physical relationship between the SWH settings and the target performances. We expect that this study can fill the blank of the HTS applications on optimizing energy systems and provide new insight on the design of high-performance energy systems.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Hao Li proposed and studied the overall HTS framework and wrote the manuscript. Zhijian Liu provided the experimental and financial supports. Kejun Liu provided relevant programming supports. Zhien Zhang participated in the discussions and revised the manuscript.

Acknowledgments

This work was supported by the Major Basic Research Development and Transformation Program of Qinghai province (no. 2016-NN-141) and Natural Science Foundation of Hebei (no. E2017502051).

References

- [1] S. Mekhilef, R. Saidur, and A. Safari, “A review on solar energy use in industries,” *Renewable and Sustainable Energy Reviews*, vol. 15, pp. 1777–1790, 2011.
- [2] Z. Liu, H. Li, X. Zhang, G. Jin, and K. Cheng, “Novel method for measuring the heat collection rate and heat loss coefficient of water-in-glass evacuated tube solar water heaters based on artificial neural networks and support vector machine,” *Energies*, vol. 8, pp. 8814–8834, 2015.
- [3] Z. Wei, T. M. Lim, M. Skyllas-Kazacos, N. Wai, and K. J. Tseng, “Online state of charge and model parameter co-estimation based on a novel multi-timescale estimator for vanadium redox flow battery,” *Applied Energy*, vol. 172, pp. 169–179, 2016.
- [4] Z. Wei, K. J. Tseng, N. Wai, T. M. Lim, and M. Skyllas-Kazacos, “Adaptive estimation of state of charge and capacity with online identified battery model for vanadium redox flow battery,” *Journal of Power Sources*, vol. 332, pp. 389–398, 2016.
- [5] Z. Wei, S. Meng, K. J. Tseng, T. M. Lim, B. H. Soong, and M. Skyllas-Kazacos, “An adaptive model for vanadium redox flow battery and its application for online peak power estimation,” *Journal of Power Sources*, vol. 344, pp. 195–207, 2017.
- [6] Z. Wang and Y. Li, “Layer pattern thermal design and optimization for multistream plate-fin heat exchangers—a review,” *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 500–514, 2016.
- [7] Z. Wang, B. Sundén, and Y. Li, “A novel optimization framework for designing multi-stream compact heat exchangers and associated network,” *Applied Thermal Engineering*, vol. 116, pp. 110–125, 2017.
- [8] Z. Wang and Y. Li, “Irreversibility analysis for optimization design of plate fin heat exchangers using a multi-objective cuckoo search algorithm,” *Energy Conversion and Management*, vol. 101, pp. 126–135, 2015.
- [9] J. Xu and J. Tang, “Modeling and analysis of piezoelectric cantilever-pendulum system for multi-directional energy harvesting,” *Journal of Intelligent Material Systems and Structures*, vol. 28, pp. 323–338, 2017.
- [10] J. Xu and J. Tang, “Linear stiffness compensation using magnetic effect to improve electro-mechanical coupling for piezoelectric energy harvesting,” *Sensors and Actuators A: Physical*, vol. 235, pp. 80–94, 2015.
- [11] J. W. Xu, Y. B. Liu, W. W. Shao, and Z. Feng, “Optimization of a right-angle piezoelectric cantilever using auxiliary beams with different stiffness levels for vibration energy harvesting,” *Smart Materials and Structures*, vol. 21, p. 65017, 2012.
- [12] S. Kalogirou, “The potential of solar industrial process heat applications,” *Applied Energy*, vol. 76, pp. 337–361, 2003.
- [13] S. A. Kalogirou, S. Panteliou, and A. Dentsoras, “Artificial neural networks used for the performance prediction of a thermosiphon solar water heater,” *Renewable Energy*, vol. 18, pp. 87–99, 1999.
- [14] S. A. Kalogirou, “Artificial neural networks and genetic algorithms in energy applications in buildings,” *Advances in Building Energy Research*, vol. 3, pp. 83–119, 2009.
- [15] S. A. Kalogirou, “Applications of artificial neural-networks for energy systems,” *Applied Energy*, vol. 67, pp. 17–35, 2000.
- [16] S. A. Kalogirou, “Solar thermal collectors and applications,” *Progress in Energy and Combustion Science*, vol. 30, pp. 231–295, 2004.
- [17] S. Kalogirou, “Artificial neural networks for the prediction of the energy consumption of a passive solar building,” *Energy*, vol. 25, pp. 479–491, 2000.
- [18] S. A. Kalogirou, E. Mathioulakis, and V. Belessiotis, “Artificial neural networks for the performance prediction of large solar systems,” *Renewable Energy*, vol. 63, pp. 90–97, 2014.

- [19] S. Kalogirou, A. Designing, and Modeling Solar Energy Systems, *Solar Energy Engineering*, pp. 583–699, Elsevier, Oxford, UK, 2014.
- [20] Z. Liu, K. Liu, H. Li, X. Zhang, G. Jin, and K. Cheng, “Artificial neural networks-based software for measuring heat collection rate and heat loss coefficient of water-in-glass evacuated tube solar water heaters,” *PLoS One*, vol. 10, article e0143624, 2015.
- [21] Z. Liu, H. Li, X. Tang, X. Zhang, F. Lin, and K. Cheng, “Extreme learning machine: a new alternative for measuring heat collection rate and heat loss coefficient of water-in-glass evacuated tube solar water heaters,” *SpringerPlus*, vol. 5, 2016.
- [22] H. Peng and X. Ling, “Optimal design approach for the plate-fin heat exchangers using neural networks cooperated with genetic algorithms,” *Applied Thermal Engineering*, vol. 28, pp. 642–650, 2008.
- [23] Z. Liu, H. Li, K. Liu, H. Yu, and K. Cheng, “Design of high-performance water-in-glass evacuated tube solar water heaters by a high-throughput screening based on machine learning: a combined modeling and experimental study,” *Solar Energy*, vol. 142, pp. 61–67, 2017.
- [24] W. F. An and N. Tolliday, “Cell-based assays for high-throughput screening,” *Molecular Biotechnology*, vol. 45, pp. 180–186, 2010.
- [25] T. Colbert, “High-throughput screening for induced point mutations,” *Plant Physiology*, vol. 126, pp. 480–484, 2001.
- [26] J. Bajorath, “Integration of virtual and high-throughput screening,” *Nature Reviews Drug Discovery*, vol. 1, pp. 882–894, 2002.
- [27] D. Wahler and J. L. Reymond, “High-throughput screening for biocatalysts,” *Current Opinion in Biotechnology*, vol. 12, pp. 535–544, 2001.
- [28] R. P. Hertzberg and A. J. Pope, “High-throughput screening: new technology for the 21st century,” *Current Opinion in Chemical Biology*, vol. 4, pp. 445–451, 2000.
- [29] J. Greeley, T. F. Jaramillo, J. Bonde, I. B. Chorkendorff, and J. K. Nørskov, “Computational high-throughput screening of electrocatalytic materials for hydrogen evolution,” *Nature Materials*, vol. 5, pp. 909–913, 2006.
- [30] J. Greeley and J. K. Nørskov, “Combinatorial density functional theory-based screening of surface alloys for the oxygen reduction reaction,” *Journal of Physical Chemistry C*, vol. 113, pp. 4932–4939, 2009.
- [31] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, “Finding natures missing ternary oxide compounds using machine learning and density functional theory,” *Chemistry of Materials*, vol. 22, pp. 3762–3767, 2010.
- [32] G. L. Morrison, N. H. Tran, D. R. McKenzie, I. C. Onley, G. L. Harding, and R. E. Collins, “Long term performance of evacuated tubular solar water heaters in Sydney, Australia,” *Solar Energy*, vol. 32, pp. 785–791, 1984.
- [33] R. Tang, Z. Li, H. Zhong, and Q. Lan, “Assessment of uncertainty in mean heat loss coefficient of all glass evacuated solar collector tube testing,” *Energy Conversion and Management*, vol. 47, pp. 60–67, 2006.
- [34] Y. M. Liu, K. M. Chung, K. C. Chang, and T. S. Lee, “Performance of thermosyphon solar water heaters in series,” *Energies*, vol. 5, pp. 3266–3278, 2012.
- [35] G. L. Morrison, I. Budihardjo, and M. Behnia, “Water-in-glass evacuated tube solar water heaters,” *Solar Energy*, vol. 76, pp. 135–140, 2004.
- [36] L. J. Shah and S. Furbo, “Theoretical flow investigations of an all glass evacuated tubular collector,” *Solar Energy*, vol. 81, pp. 822–828, 2007.
- [37] Z. H. Liu, R. L. Hu, L. Lu, F. Zhao, and H. S. Xiao, “Thermal performance of an open thermosyphon using nanofluid for evacuated tubular high temperature air solar collector,” *Energy Conversion and Management*, vol. 73, pp. 135–143, 2013.
- [38] M. Souliotis, S. Kalogirou, and Y. Tripanagnostopoulos, “Modelling of an ICS solar water heater using artificial neural networks and TRNSYS,” *Renewable Energy*, vol. 34, pp. 1333–1339, 2009.
- [39] W. Sun, Y. He, and H. Chang, “Forecasting fossil fuel energy consumption for power generation using QHSA-based LSSVM model,” *Energies*, vol. 8, pp. 939–959, 2015.
- [40] H. C. Jung, J. S. Kim, and H. Heo, “Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach,” *Energy and Buildings*, vol. 90, pp. 76–84, 2015.
- [41] K. Mohammadi, S. Shamshirband, P. L. Yee, D. Petković, M. Zamani, and S. Ch, “Predicting the wind power density based upon extreme learning machine,” *Energy*, vol. 86, pp. 232–239, 2015.
- [42] S. Kalogirou, “Applications of artificial neural networks in energy systems,” *Energy Conversion and Management*, vol. 40, pp. 1073–1087, 1999.
- [43] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, pp. 489–501, 2006.
- [44] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Transactions on Systems Man and Cybernetics, Part B (Cybernetics)*, vol. 42, pp. 513–529, 2012.
- [45] G. Huang, G. B. Huang, S. Song, and K. You, “Trends in extreme learning machines: a review,” *Neural Networks*, vol. 61, pp. 32–48, 2015.
- [46] M.-C. Lee and C. To, “Comparison of support vector machine and back propagation neural network in evaluating the enterprise financial distress,” *International Journal of Artificial Intelligence & Applications*, vol. 1, pp. 31–43, 2010.
- [47] J. Z. Wang, J. J. Wang, Z. G. Zhang, and S. P. Guo, “Forecasting stock indices with back propagation neural network,” *Expert Systems with Applications*, vol. 38, pp. 14346–14355, 2011.
- [48] N. M. Nawi, A. Khan, and M. Z. Rehman, “A new back-propagation neural network optimized,” *ICCSA 2013*, pp. 413–426, 2013.
- [49] D. F. Specht, “A general regression neural network,” *IEEE Transactions on Neural Networks*, vol. 2, pp. 568–576, 1991.
- [50] C.-M. Hong, F.-S. Cheng, and C.-H. Chen, “Optimal control for variable-speed wind generation systems using general regression neural network,” *International Journal of Electrical Power & Energy Systems*, vol. 60, pp. 14–23, 2014.
- [51] H. Li, X. Tang, R. Wang, F. Lin, Z. Liu, and K. Cheng, “Comparative study on theoretical and machine learning methods for acquiring compressed liquid densities of 1,1,1,2,3,3,3-heptafluoropropane (R227ea) via song and Mason equation, support vector machine, and artificial neural networks,” *Applied Sciences*, vol. 6, p. 25, 2016.

- [52] E. O. Pyzer-Knapp, C. Suh, R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, and A. Aspuru-Guzik, "What is high-throughput virtual screening? A perspective from organic materials discovery," *Annual Review of Materials Research*, vol. 45, pp. 195–216, 2015.
- [53] S. A. Kalogirou, "Artificial neural networks in renewable energy systems applications: a review," *Renewable and Sustainable Energy Reviews*, vol. 5, 2000.

