Renewable Energy 177 (2021) 1001–1013

Contents lists available at ScienceDirect

Renewable Energy

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

Machine learning approaches for thermal updraft prediction in wind solar tower systems

Mostafa A. Rushdi [a, b, *], Shigeo Yoshida [a], Koichi Watanabe [c], Yuji Ohya [a]

[a] Research Institute for Applied Mechanics (RIAM), Kyushu University, Fukuoka, 816-8580, Japan
[b] Faculty of Engineering and Technology, Future University in Egypt (FUE), New Cairo, 11835, Egypt
[c] Platform of Inter/Transdisciplinary Energy Research (Q-PIT), Kyushu University, Kasuga, 816-8580, Japan

ARTICLE INFO

Article history:
Received 9 March 2021
Received in revised form
24 May 2021
Accepted 5 June 2021
Available online 9 June 2021

Keywords:
Wind solar tower
Renewable energy
Machine learning
Linear regression model

ABSTRACT

Wind solar towers constitute a fairly new scheme for harvesting renewable energy from solar and wind energy sources. In such a tower, solar radiation is collected and hot air is enforced to go fast through the tower, a process called thermal updraft, which fuels a wind turbine to generate power. Using vortex generators at the top of the tower creates a pressure difference, which increases the thermal updraft. In this work, we describe the setup of a wind solar tower system established at Kyushu University in Japan. Then, we demonstrate how data was collected from this system in order to train regression models for thermal updraft prediction. The feature selection process was guided by sensitivity analysis. After that, several machine learning models were investigated and the most suitable model was selected based on quality and time metrics. The linear regression model was particularly examined in detail, and was shown to have a satisfactory high accuracy of thermal updraft prediction graphically and numerically with a coefficient of determination of $R^2 = 0.981$. We also evaluated a reduced prediction model based on the six most essential features, which could be a reduced model description for the WST. This reduced model showed little performance degradation ($R^2 = 0.974$), with significant reduction in the needed effort and resources, as well as data collection requirements.

1. Introduction

With the imminent depletion of fossil fuel resources in the near future, significant research and development efforts have been made to create enhanced and efficient renewable energy systems [1]. Recently, hybrid systems with more than one source of renewable energy have been devised. In particular, hybridization of solar and wind energy sources has emerged as one of the most promising renewable energy schemes. Such a hybrid scheme can be realized using several designs including primarily wind solar towers [2].

1.1. Wind solar tower

The working idea of the solar chimney is based on the basic principle of physics, the fact that hot air rises to generate flow stream, as shown in Fig. 1a. Its configuration consists of 3 main components:

- Transparent solar thermal collector.
- High-rise vertical tower.
- Wind turbine.

Fig. 1 demonstrates the mechanism of power generation in wind solar towers. This mechanism could be divided into the following two components:

- Solar energy contribution: The air between the collector and ground is heated by the transparent solar collector, creating a green house effect. Then, a *thermal updraft* is induced within the vertical tower. This updraft turns a wind turbine at the base of the tower to produce electricity, as shown in Fig. 1a.

* Corresponding author. Research Institute for Applied Mechanics (RIAM), Kyushu University, Fukuoka, 816-8580, Japan.,
E-mail addresses: rushdimostafa@riam.kyushu-u.ac.jp, Mostafa.Roshdi@fue.edu.eg (M.A. Rushdi), yoshidas@riam.kyushu-u.ac.jp (S. Yoshida), koichi-watanabe@riam.kyushu-u.ac.jp (K. Watanabe), ohya@riam.kyushu-u.ac.jp (Y. Ohya).

https://doi.org/10.1016/j.renene.2021.06.033

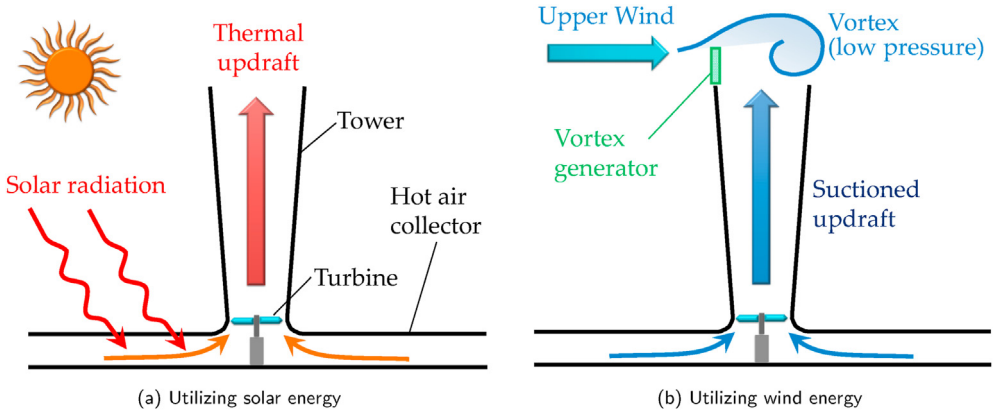


Fig. 1. A schematic for the power generation in the wind solar tower [3].

- Wind energy contribution: By adding a vortex generator at the top of the tower, the upper wind will hit the vortex generator resulting in a vortex flow pattern. This will create a low-pressure field that will urges the flow to go upward and cause a *suction updraft* that will rotate the turbine. This procedure is represented schematically in Fig. 1b.

Hybrid renewable energy systems that utilize solar and wind energy sources are fairly new. Such systems improve upon solar chimneys that generate power based on solar energy only [4–6]. In fact, the design of solar chimneys dates back 500 years ago when Leonardo Da Vinci envisioned the chimney tower as a smoke jack [7]. However, the use of solar towers for power generation was not proposed until the 19th century. Since then, several variations and enhancements of solar towers have been realized [4].

The first solar tower prototype was implemented in Manzanares, Spain by a team led by Jörg Schlaich [8,9]. This prototype achieved 50 kW successfully. According to the findings from this project, Haaf et al. [10,11] defined the basic physical concepts of electricity generation with solar chimney power plants (SCPP). This opened the doors for more projects in other countries [12–16] and research on how to improve the power generation [17,18]. From this point, new types of WST were introduced [19,20], which mainly differ on how to utilize wind power.

1.2. Machine learning applications in WST

Machine learning (ML) and deep learning (DL) methods have gained a lot of research momentum recently because of their capabilities in modeling nonlinear input–output relations when solving classification or regression problems. The power of these methods extends to multivariate problems where the number of input variables or features is large. Learning methods have been successfully applied in computer vision [21], pattern recognition [21], bioinformatics [22], medical diagnosis [23] etc. As well, common machine learning methods are currently included in hardware-optimized software libraries such as Scikit Learn [24], Pytorch [25] and TensorFlow [26].

Supervised learning methods represent a class of widely used methods. In such a method, pairs of the input variables $\mathbf{x}$ and the output variable y are used to learn the input–output mapping function $y = f(\mathbf{x})$. The goal of any supervised learning method is to approximate a mapping function by optimizing some objective function, such that when new input data samples $\mathbf{x}^*$ are available (without associated real output), we would be able to predict the outputs $y^*(\mathbf{x}^*)$ for these data samples. One-dimensional linear regression, for example, is the problem of fitting a line $y = ax + b$ to a number of n labeled points $\{\mathbf{x}, y\}_{i=1}^{n}$, by minimizing some loss function, such as the least-square error function.

Several approaches have been proposed for solar chimney system modeling, including primarily the work done by Pasumarthi et al. [27,28]. This model was tested and verified on a small-scale prototype built in Gainesville, Florida. However, on the large-scale, the model shows 20% accuracy regarding the exit velocity and 9.5% regarding the power output [28]. Machine learning models can show better generalization, tolerate weather variations, and represent good solutions for modeling and prediction in renewable energy systems, if the suitable data was measured. In this paper, we will use the data collected from the wind-solar tower developed and established at Kyushu University [3,16] for modeling and prediction of the thermal updraft in wind solar tower systems using a machine learning approach. To the best of our knowledge, this is the first attempt to employ machine learning algorithms with experimental results from a wind solar tower. Nevertheless, several machine learning methods have been reported for wind energy applications [29–33], solar energy applications [34,35], and renewable energy systems in general [36].

1.3. Organization and contribution

In this paper, we present a fairly new renewable energy technique that combines solar and wind powers. It's called the wind-solar tower or wind-solar chimney power plant. We used the data collected from the wind-solar tower developed at Kyushu University [3,16]. We performed sensitivity analysis to choose the important features. The aim was to come up with a model that describes the system or predicting the thermal updraft. We evaluated several machine learning models based on key quality metrics to choose the best model that balances accuracy and speed.

The paper is divided into five main sections. Section 2 presents the wind-solar tower system and the collected data. In Section 3, we presented the data and performed sensitivity analysis that guided us through the feature selection. Section 4 shows the machine learning construction and how the quality assessment of the model was performed. In Section 5, we present the detailed results of a simple and efficient machine learning technique, which is a linear regression model. Finally, Section 6 concludes the paper and points out future research directions.

2. System setup and data collection

2.1. System components

The data was collected from a WST system built in Kyushu University - Chikuchi Campus. The details of the system could be found in the previous work of Ohya et al. [3,16]. The data was collected using the following sensors:

- Temperature sensors for measuring the temperature at several points at the base or inside the collector. Also, for measuring the ambient air temperature and the temperature of the thermal updraft inside the tower.
- A pyranometer for measuring solar radiation.
- An ultrasonic speedometer for measuring the thermal updraft (i.e. the internal ascending wind speed).
- A 3-cup anemometer for measuring the wind speed of the ambient air.
- Sensors to measure the wind turbine speed (rpm) and output power (mW).

The system setup was originally symmetric, i.e. the tower is exactly in the middle of the collector base. Currently, the collector was moved a little to the right (as represented by the green dashed line in Fig. 2), which makes the two temperate sensors [CH1-3 & CH1-4] outside the collector. The reason for the new configuration is that it was expected it will cause a higher power generation due to the effect of the heat transfer representative length [37].

2.2. Data description and statistics

Data collection was performed with two setups:

- **No wind turbine:** The output in this case is the *thermal updraft* (CH2-7). Note that, the channels of the WT are still connected in this setup but return no useable/meaningless data, as shown and proved through the sensitivity analysis.
- **With wind turbine:** The output in this case is the *power generated* from WT (CH3-2).

The two cases are represented in two folders, each folder contains several ".csv" files. Each file contain 28 columns, consists of 27 channels represented in Table 1 in addition to a "DATETIME" column that represent the exact date and time of the measured sample. The locations of the sensors are presented in Fig. 2 and the data statistics are presented in Table 2. Note that there are some quantities that are measured by two sensors, then we use the stable one.

In this paper, we will focus on the case of "No wind turbine" and it consists of 65 days of measurements, each day represented in one ".csv" file. The measurement is done between 4:50 p.m. and 10:40 a.m. on random days within the period between November and April, which belongs to Autumn and Winter semesters and considered to have less solar radiation.

2.3. Added features

Two more features were also investigated for enhancing the thermal updraft prediction performance. The first one is the difference between the ambient temperature [CH1-12] and the temperature inside the wind solar tower [CH1-13].

The second feature is a categorical variable for identifying day and night times. This feature was initially calculated based on the current time and the sun position relative to the WST system location using the PyEphem Python library for astronomical algorithms. Contrary to our expectations, this feature turned to be not highly correlated with the output thermal updraft. This is possibly because of frequent cloudy days. While we may have picked another feature to reflect weather, we chose to categorize this feature based on the amount of solar radiation. Later, the

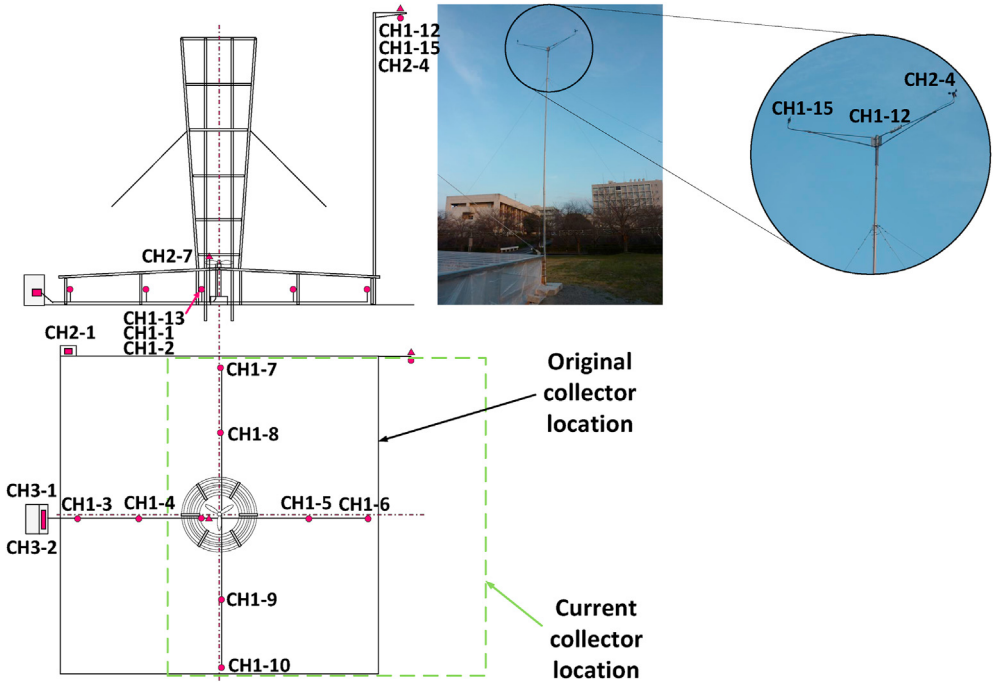


Fig. 2. Sensor locations within the wind solar tower system.

Table 1
Description and units of the collected data.

Channels	Description	Unit
CH1-1: CH1-10	Temperature (at the base/collector)	°C
CH1-11	Wind speed [Not Used - replaced by CH2-4]	m/s
CH1-12	Temperature of the outside air	°C
CH1-13	Temperature inside the tower	°C
CH1-14	[−]	–
CH1-15	Wind direction (outside the tower)	°
CH1-16	Wind turbine speed [Not Used - replaced by CH3-1]	rpm
CH2-1	Solar radiation (by pyranometer)	kW/m^2
CH2-2	Electric generating capacity (inside the collector)	kW
CH2-3	Electric generating capacity (outside)	kW
CH2-4	Wind speed of the outside air "u"	m/s
CH2-5	Wind turbine output [Not Used - replaced by CH3-2]	kW
CH2-6	Wind turbine rotation speed [Not Used - replaced by CH3-1]	rpm
CH2-7	Internal ascending wind speed "w" (at the rotor)	m/s
CH2-8	Wind direction (at the rotor)	°
CH3-1	Wind turbine speed	rpm
CH3-2	Wind turbine power output	mW
CH3-3	Voltage after rectification	V

Table 2
Basic statistics of the collected data.

CH	mean	std	median	min	max
CH1-1	7.434557	4.926402	5.233700	2.345050	23.859300
CH1-2	14.485885	5.028030	12.490100	9.710620	30.764100
CH1-3	−55.717631	15.743660	−59.065200	−59.577300	20.777300
CH1-4	−7.978556	28.897641	7.325270	−59.594100	30.846200
CH1-5	12.858084	6.706170	10.528900	8.004400	36.033000
CH1-6	11.327303	6.192751	8.629300	6.115750	32.773600
CH1-7	13.107311	6.942188	10.101100	8.027840	35.455700
CH1-8	14.829164	8.376205	11.367000	8.721610	44.804400
CH1-9	14.486690	8.061751	11.066700	8.797800	40.088600
CH1-10	21.125678	8.409371	21.671800	8.906230	41.099600
CH1-11	0.435379	0.038580	0.419927	0.386007	0.544762
CH1-12	11.635292	4.412510	9.934800	6.576560	26.411000
CH1-13	13.266569	4.950520	11.226900	9.090820	29.602600
CH1-14	0.007975	0.001109	0.008066	0.005289	0.011612
CH1-15	135.461883	60.934473	124.607000	7.905580	339.600000
CH1-16	−9.159765	4.746319	−11.980500	−12.000000	2.442000
CH2-1	0.074719	0.190329	−0.007750	−0.028270	0.660562
CH2-2	0.000523	0.000164	0.000509	0.000122	0.000997
CH2-3	0.002564	0.001272	0.003093	0.000305	0.004660
CH2-4	0.472834	0.376623	0.281543	0.266172	2.201430
CH2-5	0.367155	0.190773	0.411945	0.070859	0.622304
CH2-6	0.011046	0.012899	0.004212	0.001852	0.043773
CH2-7	0.786003	0.965832	0.356561	−0.009530	3.651520
CH2-8	118.174867	57.351833	145.852000	8.652930	252.247000
CH3-1	0.000000	0.000000	0.000000	0.000000	0.000000
CH3-2	1.028876	3.007783	0.100000	0.000000	35.300000
CH3-3	0.073904	0.082253	0.052933	0.005133	0.642300

categorical feature was encoded using a one-hot encoder into a one-hot numeric array. So, each category is represented by a binary vector.

3. Data analysis and pre-processing

In this section, we will provide details of our data analysis including data visualization, and correlation analysis. We also carry out sensitivity analysis to decide suitable parameter settings and significant features for prediction model construction. Numerous sensitivity analysis methods have been proposed for the design and realization of wind solar tower systems [38–41], as well as investigating the effects of parameter variations on the WST system output [14,15]. In the current work, we perform sensitivity analysis based on the Pearson correlation coefficient, and select the features accordingly.

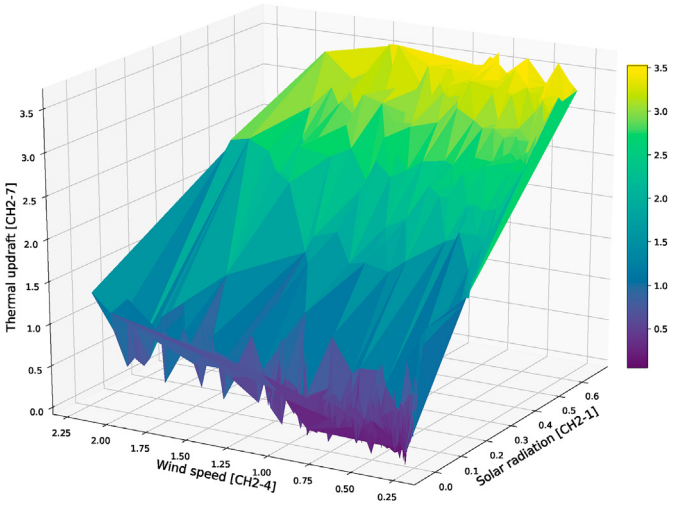


Fig. 3. A 3D mesh plot of the thermal updraft as a function of wind speed and solar radiation.

3.1. Data visualization

For exploring the collected data patterns and gaining better insights on them, we examined the variations in the output thermal updraft [CH2-7] as a function of the two most important WST parameters, namely the solar radiation [CH2-1] and the wind speed [CH2-4]. Figs. 3 and 4 show 3D mesh and scatter plots of this functional relationship, respectively. The two figures show that the effect of solar radiation is higher than that of the wind speed. Also, Fig. 4 shows that solar radiation is highly correlated with the thermal updraft, according to the projection on the plane containing both variables. This correlation is further demonstrated in Fig. 6.

3.2. Correlation analysis

Bivariate correlation analysis seeks to measure the strength and direction of association between two variables. The correlation strength is represented by the value of the correlation coefficient which varies between +1 and −1 [42].

A perfect degree of association between two variables is indicated by correlation coefficient values of 1 or -1. The association between the two variables will be weaker, as the value of the correlation coefficient goes towards zero. The sign of the correlation coefficient indicates the direction of the relationship: +1 indicates a direct linear proportionality, while −1 indicates inverse proportionality.

The most widely used statistical measures of correlation are Pearson correlation, Kendall rank correlation and Spearman correlation. Pearson correlation is the most common measure of dependence between two variables. This measure is defined as the ratio of the co-variance of the two variables to the square root of their variance product. Based on expected value computations, a Pearson product-moment correlation coefficient can be used to create a best-fit line for a dataset of the two variables. The resulting Pearson correlation coefficient shows how far apart the real data

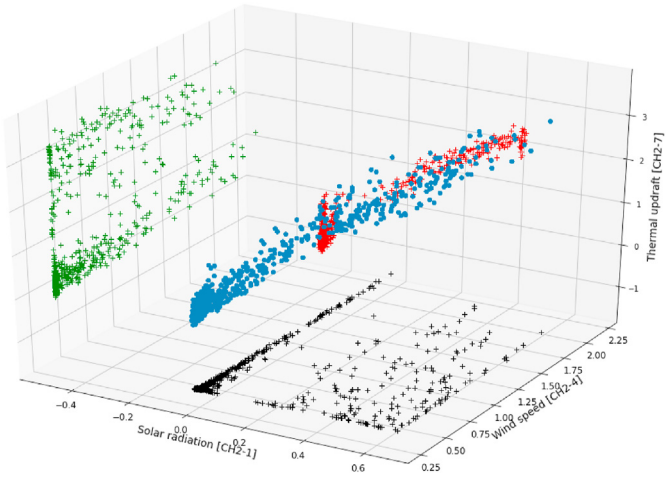


Fig. 4. A 3D scatter plot of the thermal updraft as a function of wind speed and solar radiation.

measurements are from the expected values. We will end up with either negative or positive correlation (depending on the sign of Pearson's correlation coefficient) if there is some kind of interaction between variables.

The correlation coefficient ρ for two random variables X and Y with expected values μ_X & μ_Y and standard deviations σ_X & σ_Y is defined as

$$\rho_{X,Y} = \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{1}$$

where E is the expected value operator, and cov() denotes the covariance function. The Pearson correlation is defined only if both standard deviations are finite and positive. This correlation coefficient can be defined alternatively as:

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}} \tag{2}$$

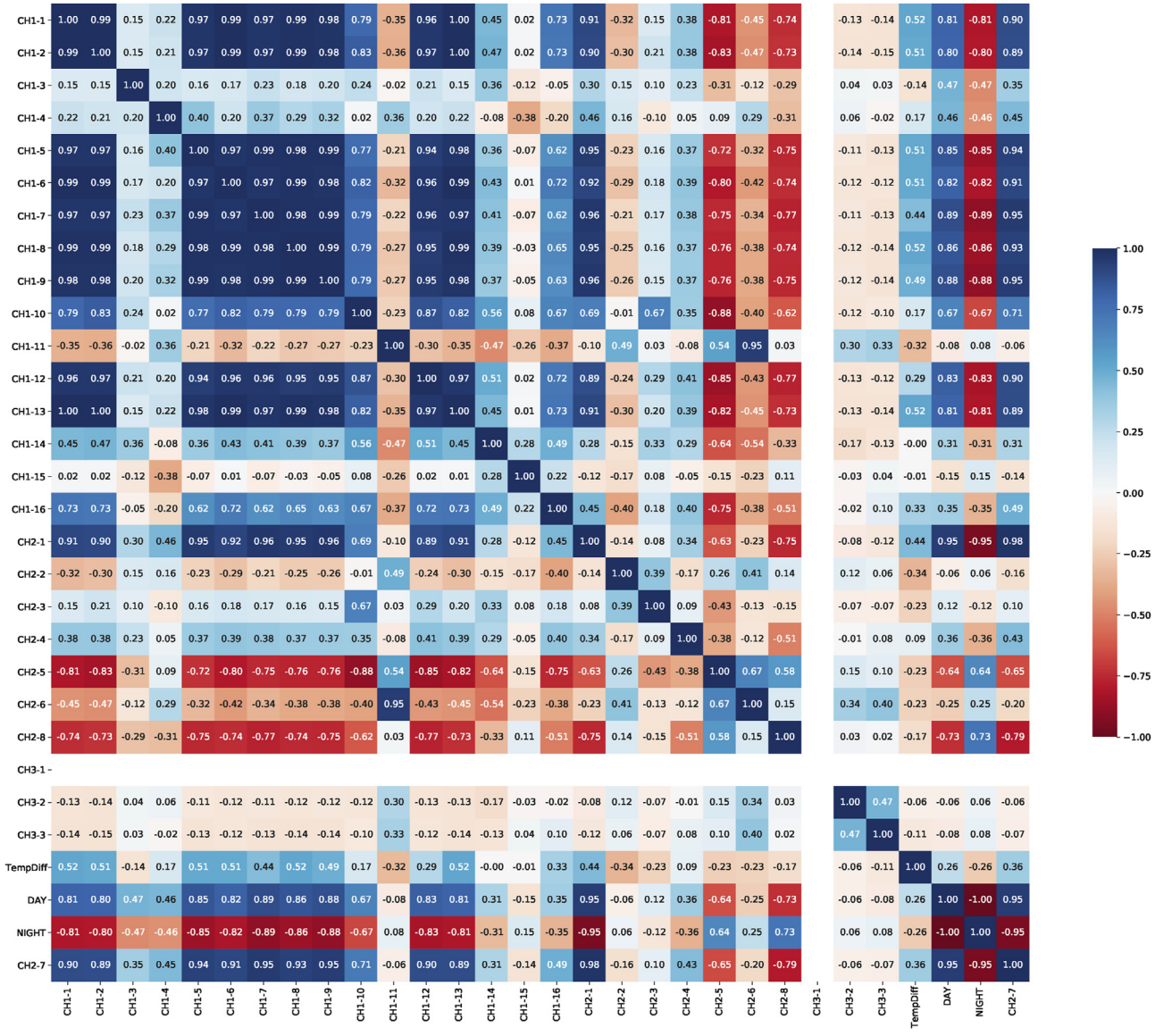


Fig. 5. Heatmap of the pairwise Pearson correlation coefficient between each pair of the measured data.

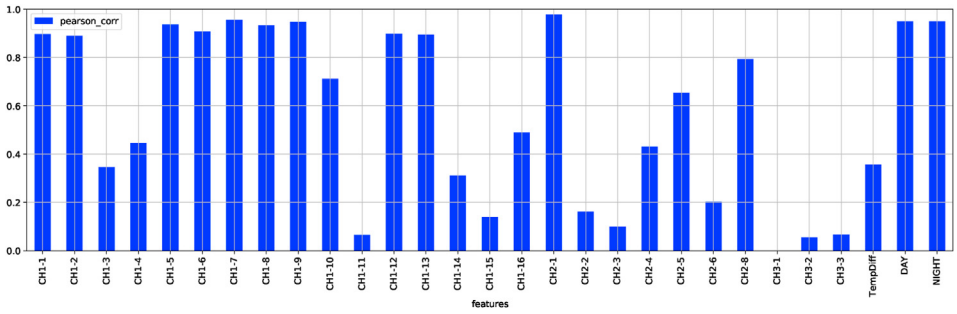


Fig. 6. Absolute value of Pearson correlation coefficient between input features and the output thermal updraft [CH2-7].

Fig. 5 shows a heatmap representation of the Pearson correlation coefficients for all variables. Further, Fig. 6 shows the correlation of each possible input variable of the WST system with the thermal updraft output. Indeed, Fig. 6 is just an alternative visual representation of the last row of the heatmap in Fig. 5. The correlation information in these two figures is exploited for feature selection.

Figs. 7–10 show the correlation between thermal updraft and each of the following key input variables respectively: the solar radiation, the wind speed, the temperature outside the tower, and the temperature inside the tower. For example, Fig. 7 shows that the data samples are concentrated around the red line, indicating that an increase in the solar radiation will lead to a strongly proportional increase in the thermal updraft, especially after 0.2 kW/m2. This strong correlation between the solar radiation and thermal updraft is quantified by the high Pearson correlation coefficient of 0.98. On another hand, Fig. 8 shows that wind speed and thermal updraft have a moderate value of the Pearson correlation coefficient of 0.43, which agrees with the observation that the data is scattered and hence the red line couldn't describe a strong relationship.

3.3. Feature selection

Feature selection is the process in which we select, automatically or manually, the features that contribute the most to the prediction variable or output that we are interested in. If there are irrelevant features in the data, these features can decrease the prediction model accuracy, robustness, and generalization. Feature selection and data cleaning should be the first and most essential step of prediction model construction [43].

In our work, we will manually select features based on the correlation analysis results shown in Fig. 5. According to our

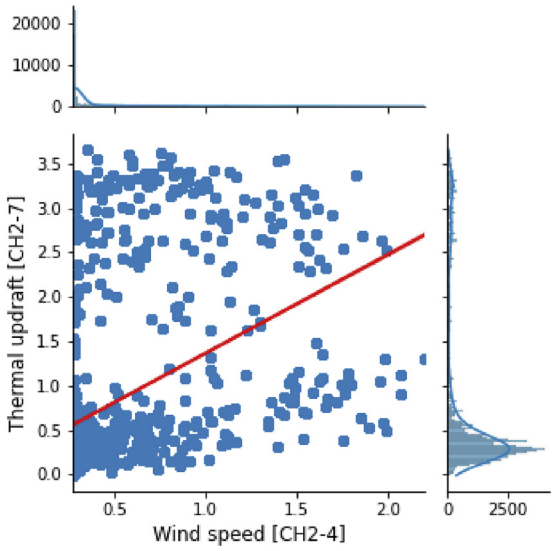


Fig. 8. Correlation between thermal updraft and wind speed.

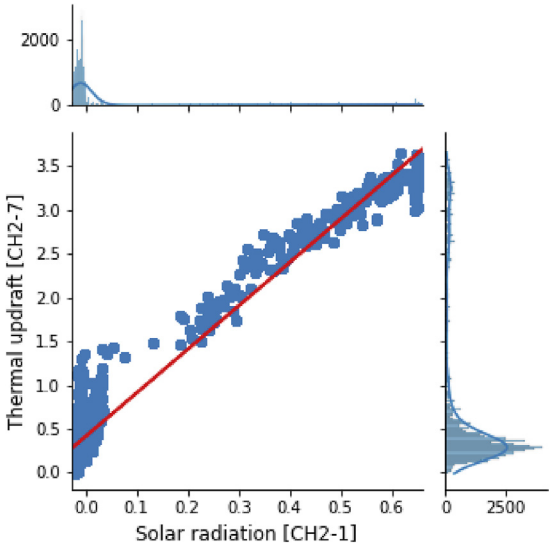


Fig. 7. Correlation between thermal updraft and solar radiation.

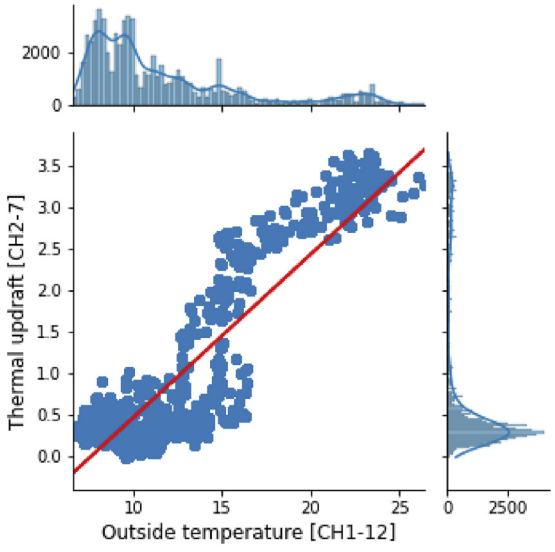


Fig. 9. Correlation between thermal updraft and outside temperature.

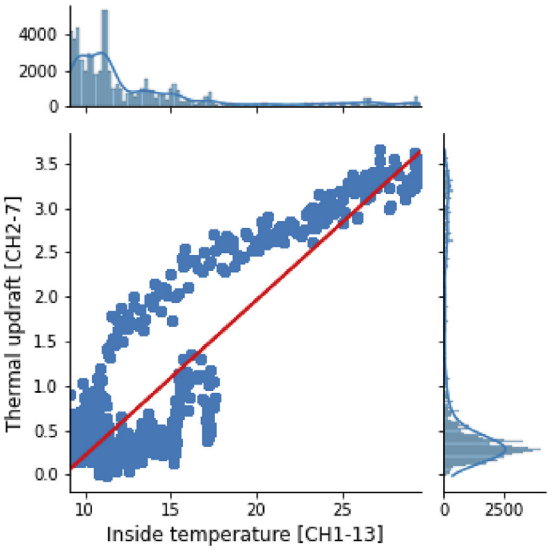


Fig. 10. Correlation between thermal updraft and inside temperature.

intuitive understanding of the WST system, we expect that the output thermal updraft will mainly be affected by solar radiation. As expected, Fig. 6 shows that the thermal updraft has the highest correlation with the solar radiation. This agrees with the temperature readings from the sensors distributed at the base/collector [CH1-5 - CH1-10] as well as inside and outside the tower [CH1-12 & CH1-13].

In the current system configuration, the two channels [CH1-3 & CH1-4] are outside the base/collector and should be correlated with the outside temperature [CH1-12]. However, the heatmap didn't support this correlation. So, we made further investigations and found some irregularities in the sensor measurements. As a result, those two features will not be selected because their measurements are questionable and could be replaced instead by [CH1-12].

The sensors of the two channels [CH1-1 & CH1-2] are located inside the tower, and so they are highly correlated with [CH1-13], although their sensor types are different. The channel [CH1-2] is located just under the location of [CH1-13], and thus the two channels are highly correlated, as shown in Fig. 5. For now, we will keep all these channels. However, [CH1-1 & CH1-2] could be replaced by [CH1-13].

In the current no-wind-turbine setup, the wind-turbine-related channels are irrelevant and have very low correlation as shown in Fig. 6. The following channels will be removed: [CH1-16, CH2-5, CH2-6, CH3-1, CH3-2, CH3-3].

4. Construction of prediction models

4.1. Performance assessment

For performance assessment of the machine learning algorithm, the data set was split, as shown in Fig. 11, as follows:

- Training Set: 70% of the samples were used to train the model.
- Test Set: 20% of the samples used to test the model using the quality metrics.
- Validation Set: 10% of the data was separated from the beginning and the model doesn't have any access or information about them. This set will be used for testing the ability of generalization of the model. This set could be chosen based on some data that have a certain meaning. In our case, we choose this set as the first day of measurements.

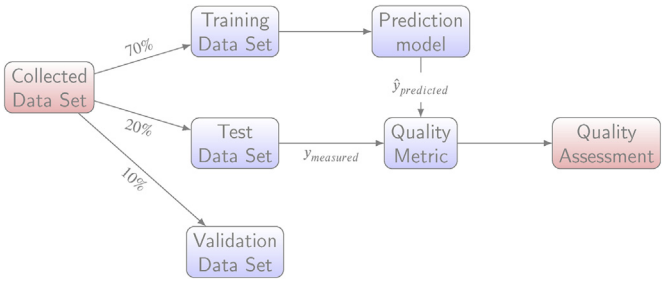


Fig. 11. Experimental setup for prediction model training and testing (adapted from Ref. [33]).

After training, multiple models were developed for thermal updraft prediction. However, all of the obtained models are complex and not easily interpretable, except for the linear regression model which we investigate further. For the testing data, the model performance is evaluated by comparing the model prediction and the ground-truth thermal updraft measurements.

Quality metrics were used as cost functions to be minimized during model training. Then, any optimization algorithm such as the gradient descent one could be used to minimize these cost functions. The quality metrics were used for quantitative comparison on the test data. The employed quality metrics in this work are reviewed next.

4.2. Quality metrics

In this subsection, we denote $\hat{y}_i$ as the predicted value of the i-th sample, y_i as the corresponding true value, m as the number of samples and Var as the variance.

- **Mean Square Error:** The expected value of the squared (quadratic) error

$$\mathrm{MSE}(y,\hat{y}) = \frac{1}{m}\sum_{i=0}^{m-1}(y_i-\hat{y}_i)^2. \tag{3}$$

- **Coefficient of Determination** (R^2): represents the proportion of variance (of y) that has been explained by the independent variables in the model, providing an indication of goodness of fit and therefore a measure of how well unseen samples are likely to be predicted by the model, through the proportion of explained variance

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{m}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{m}(y_i-\bar{y})^2}, \tag{4}$$

where $\bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i$ and $\sum_{i=1}^{m}(y_i-\hat{y}_i)^2 = \sum_{i=1}^{m}\epsilon_i^2$. The best possible score is 1 and it can be negative, because the model can be arbitrarily worse. A constant model that always predicts the expected value of y, disregarding the input features, would get a score of 0.

- **Maximum Residual Error:** captures the worst-case error between the predicted value and the true value. In a perfectly fitted single output regression model, it would be 0 on the training set. This metric shows the extent of error that the model had when it was fitted

$$\text{Max Error}(y,\hat{y}) = \max(|y_i - \hat{y}_i|). \tag{5}$$

- **Explained Variance:** The best possible score is 1, while lower values are worse.

$$\text{Explained Variance}(y,\hat{y}) = 1 - \frac{\text{Var}\{y-\hat{y}\}}{\text{Var}\{y\}}. \tag{6}$$

- **Mean Absolute Error:** The expected value of the absolute error loss or l_1-norm loss

$$\text{MAE}(y,\hat{y}) = \frac{1}{m}\sum_{i=0}^{m-1}|y_i - \hat{y}_i|. \tag{7}$$

5. ML results

In this section, we will show the results of different machine learning algorithms and we will compare them according to the quality metrics and training time in seconds, as shown in Table 3. Then the chosen machine learning model will be represented in detail. Note that all used models are based on the standard Scikit learn implementations [24].

According to the quality metrics and training time, presented in Table 3, the linear regression model shows high/satisfactory accuracy and second-lowest training time. In the next two subsections, we will present the details of the linear regression model and we will visualize its accuracy.

5.1. Model representation

In this subsection, we will represent the linear regression model [44] in detail. It could be represented mathematically as follows:

$$\hat{y} = h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}^{\boldsymbol{T}} \cdot \boldsymbol{x} = \begin{bmatrix}\theta_0 & \theta_1 & \dots & \theta_n\end{bmatrix}\begin{bmatrix}x_0\\x_1\\\vdots\\x_n\end{bmatrix} \tag{8}$$

where $\boldsymbol{\theta}$ is the model's parameter column vector, containing the bias term θ_0 and the feature wights θ_1 to θ_n, where n is the number of features which is 20 in our case. $\boldsymbol{x}$ is the instance's feature column vector, containing x_0 to x_n, with x_0 always equal to 1. Finally, h_θ is the hypothesis function or prediction formula and $\hat{y}$ is the predicted value.

If we select the mean square error among the quality metrics as our performance measure or cost function. Then our goal will be finding the values of the feature wights $\boldsymbol{\theta}$ that minimizes the mean square error. This goal could be represented mathematically as follows:

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \underbrace{\frac{1}{2m}\sum_{i=1}^{m}\Big(\overbrace{h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})}^{\text{Predicted value}} - \overbrace{y^{(i)}}^{\text{True value}}\Big)^2}_{\text{cost function } J(\theta)} \tag{9}$$

To achieve this goal, we could use an algorithm such as gradient descent [45] that will be able to minimize the quality metrics. The gradient descent algorithm could be represented mathematically as follows:

$$\text{repeat until convergence } \{\theta_j := \theta_j - \alpha \underbrace{\frac{\partial}{\partial\theta_j}J(\boldsymbol{\theta})}_{\text{Derivative term}} \} \tag{10}$$

where α is the learning rate which represents the steps. θ_j is the j-th feature and j is from 0 to number of features m. Noting that the sign $:=$ means simultaneous update. Simply, if the derivative term has -ve then θ_j will decrease approaching the local minimum point, and if the derivative term has + ve then θ_j will increase approaching the local minimum point. We don't have to worry about getting stuck in the local minimum point as our cost function is a convex function and has one global minimum point.

Fig. 12 represents the coefficients $\boldsymbol{\theta}$ of the linear regression model. Notice that, it can be confusing and misleading to look at the coefficients plot to gauge feature importance. The reason is there maybe some of the features vary on a small scale, while others vary a lot more. To avoid this, we should reduce all the coefficients to the same unit of measure, by multiplying the coefficients by the standard deviation of the related feature [46]. This is equivalent to normalize numerical variables to their standard deviation, as

$$y = \sum coef_i \times X_i = \sum (coef_i \times std_i) \times (X_i / std_i) \tag{11}$$

In that way, we emphasize that the greater the variance of a feature, the larger the weight of the corresponding coefficient on the output, all else being equal. Fig. 13 represents the coefficient importance, it could be seen that it is quite similar to Fig. 6.

Table 3
Quality metrics of different regression models.

Model	MSE	R^2	Max Error	Exp. Var.	MAE	Train time
Decision Tree Regressor	0.000993	0.998941	0.264864	0.998941	0.017059	0.433970
Gradient Boosting	0.005623	0.994002	0.269524	0.994002	0.056275	9.272727
Non-linear Regression	0.009588	0.989772	0.425258	0.989772	0.073504	0.636417
Adaptive Boosting	0.017238	0.981611	0.343100	0.983065	0.109200	5.219409
Linear Regression	0.017779	0.981034	0.616641	0.981035	0.099711	0.039288
Ridge Regression	0.020639	0.977983	0.669888	0.977983	0.107422	0.020396
Elastic Net Regression	0.083757	0.910650	0.995346	0.910661	0.216492	0.340950
Lasso Regression	0.109775	0.882894	1.090966	0.882909	0.249860	0.466636
Neural Network	0.121896	0.869964	0.881781	0.976042	0.317456	12.782129

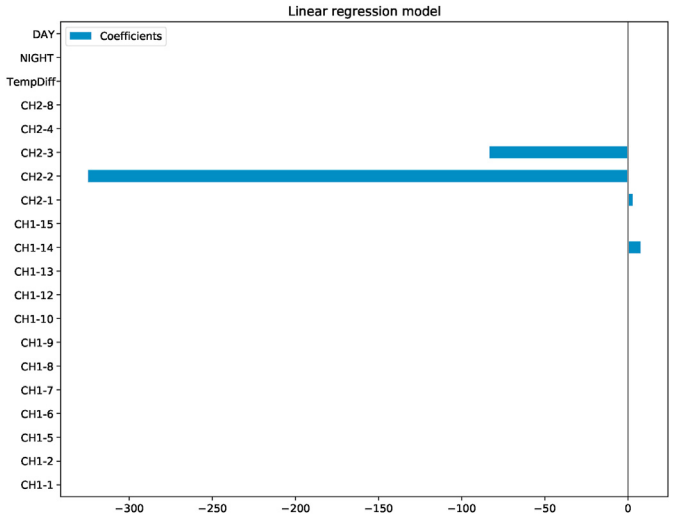


Fig. 12. Linear regression model coefficients (different units).

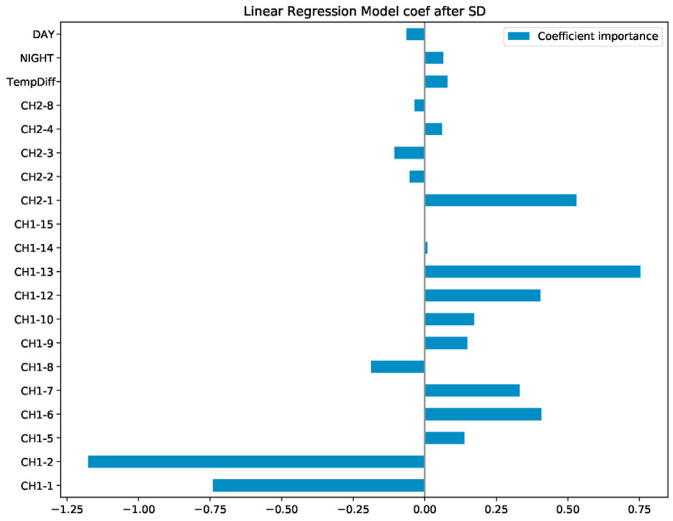


Fig. 13. Coefficients importance (scaled with SD of each feature).

5.2. Model accuracy

Fig. 14 represented the true values vs the predicted values. If the model is perfect, it will result in a straight line with 45° slope. It could be noticed that the regression red line and the 45° slope green line is almost identical.

Fig. 15 represents the histogram of the prediction error, which is the difference between the predicted value and the true value. For a good model, this histogram should be following the normal distribution centered around zero without a heavy tail, as in our case. Fig. 16 represents a box plot for the prediction error, which is another representation for the normal distribution. Finally, the most important representation of the model accuracy is using the validation set by drawing the time history of the predicted and true values for this set, as shown in Fig. 17a, and show how identical they are. The importance of Fig. 17a comes from the point that the model have not any clue about the data in the validation set, however, the model was able to predict the time history in a great way.

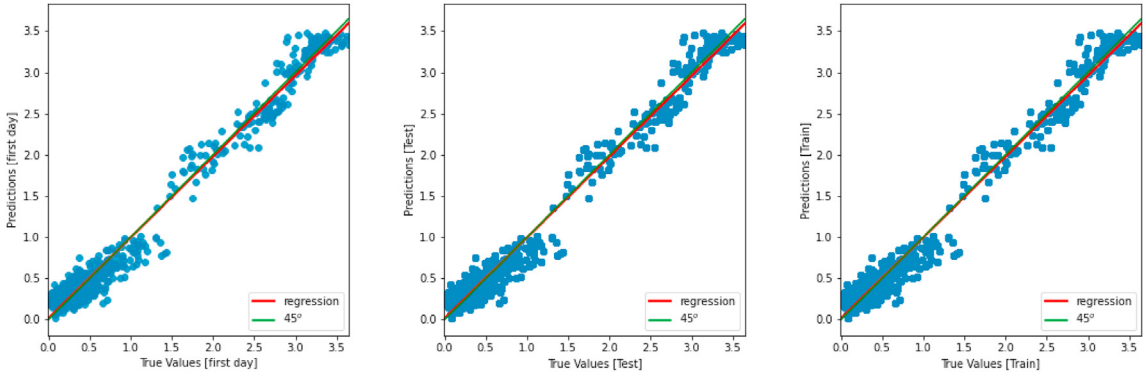
5.3. Reduced model

For further investigation of the wind solar tower system, we make the predictions using essential features only, which are stated below.

- Temperature of the outside air [CH1-12].
- Temperature inside the tower [CH1-13].
- Solar radiation [CH2-1].
- Wind speed of the outside air [CH2-4].
- Wind direction (at the rotor) [CH2-8].
- Time categorical feature.

Fig. 18 represents the comparison between the full-features model (19 features) and reduced model (6 features), according to the scores of quality metrics. It shows that the reduced model caused a small reduction in accuracy. But it will save huge efforts and resources by measuring fewer data.

The results of the linear regression model for the reduced model are represented in Fig. 19, then the coefficients importance are shown in Fig. 20. The visual measure of the model accuracy are represented by a true versus predicted values curve for the validation set, shown in Fig. 21. It is noticed that the main difference with the full-features model results, shown in Fig. 14a, is that the high values are not predicted well and there is a notable deviation



(a) Validation set (b) Test set (c) Train set

Fig. 14. True values vs predicted values of the linear regression model.

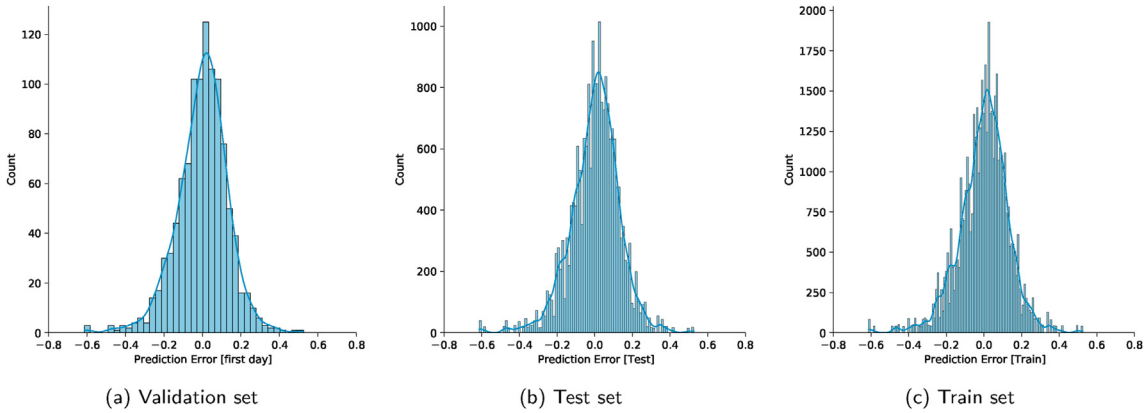


Fig. 15. Prediction Error of the linear regression model.

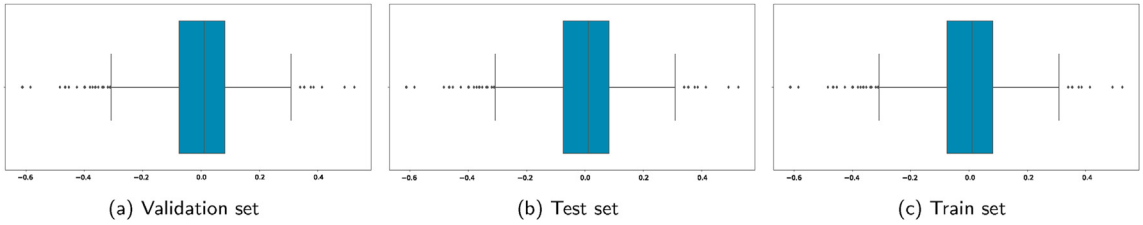


Fig. 16. Box plot for the prediction Error of the linear regression model.

between the green and red lines, which represents the perfect model with 45° slope and the regression line of the current data, respectively.

The selection of the essential features is based on the understanding of the WST system and the Pearson correlation coefficient represented in Fig. 6. All the temperature channels could be represented by two channels; one inside the tower [CH1-12] and the other one is outside the tower [CH1-13]. The channels for solar radiation [CH2-1] and wind speed [CH2-4] are essential for the WST concept. Finally, the two channels [CH2-8] & [TIME] have a high correlation with the output.

We tried to make more reductions in the number of features by reducing one feature at a time, in other words, we will reduce the number of features from 6 to 5 by different combinations. It was found that removing the solar radiation [CH2-1] from the features to be used with the ML algorithm, caused the most reduction in quality metrics, as expected. Then we tried to remove two features at a time (reduce the number of features from 6 to 4), it was found that any combination includes removing the solar radiation [CH2-1] causes a huge reduction in quality metrics values.

6. Conclusion

In this work, we demonstrated an approach to employ machine learning methods for the prediction of the thermal updraft of the wind solar tower system. We used the experimental data from the wind solar tower which designed and built at Kyushu University to make sensitivity analysis and select the suitable features that will be used within the regression model. We applied multivariate regression models, then we choose a suitable model based on the quality metrics and training time. Linear regression model was illustrated and its results were visualized to show its high prediction accuracy. Finally, we reduced the number of used features

Fig. 17. History of the true (blue line) and predicted (orange line) thermal updraft [CH2-7] values of the linear regression model.

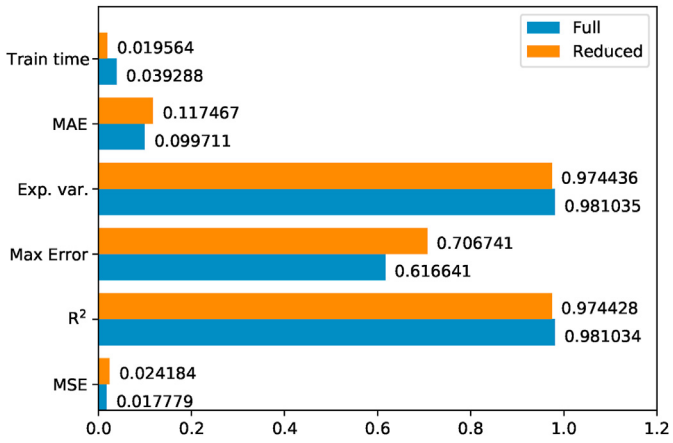


Fig. 18. Comparison between full-features model and reduced model, according to the scores of quality metrics.

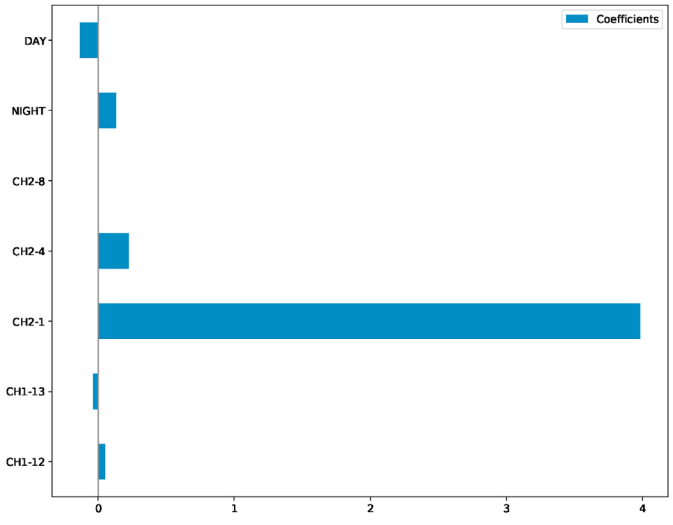


Fig. 19. Linear regression coefficients for reduced model.

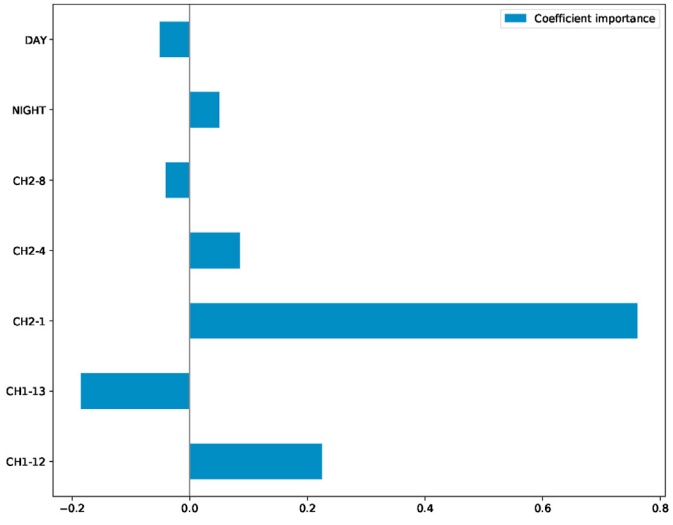


Fig. 20. Coefficient importance for the reduced model.

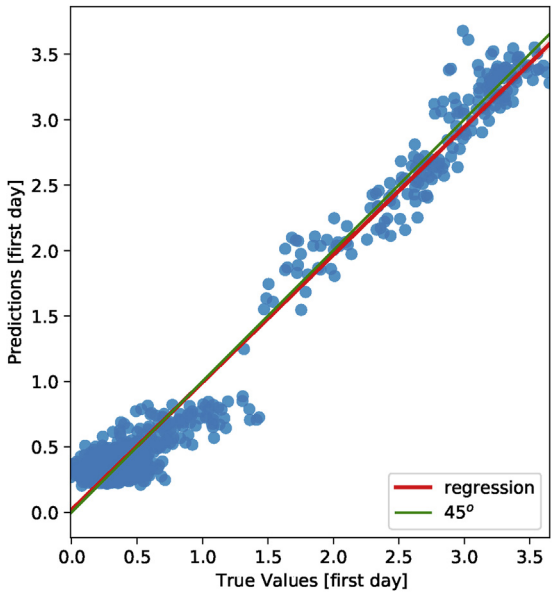


Fig. 21. True values vs predicted values of the linear regression for the reduced model validation set.

which cause a small accuracy reduction but with great efforts and resource savings.

In the future work, we will add set of features that describe the weather and that will help us to determine the suitable deployment locations for WST systems which will lead to power maximization. Also, measurements to distinguish between rainy and dry days, and measurements of natural precipitation factors inside the collector will enhance the model accuracy all over the whole year. Machine learning algorithms could be used for the system design of the wind solar tower by defining features like: inclination angle of the tower, tower height, diameter of the tower base, turbine location from the base, etc. and perform several experiments with different values for each feature and different combination between the features. Finally, the proposed reduced model shows that machine learning not only can assist future designs but also it could inform the development of simplified deterministic models.

CRediT authorship contribution statement

Mostafa A. Rushdi: Writing – original draft, Initial Draft, Writing – review & editing, Formal analysis, Formal analysis, Feature Selection, Implementation, Software, Software Development. **Shigeo Yoshida:** Feature Selection, Writing – review & editing, Supervision, Funding acquisition. **Koichi Watanabe:** System Design, Data curation, Data Collection, Feature Selection. **Yuji Ohya:** System Design, Data curation, Data Collection, Supervision, Funding acquisition.

Declaration of competing interest

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

Acknowledgement

The first author would like to acknowledge Dr. Ahmad Rushdi (Sandia National Laboratories) and Mahmoud Ayyad (Stevens Institute of Technology) for their help and guidance through the

machine learning part. We gratefully acknowledge Keiji Matsushima, Kenichiro Sugitani, and Kimihiko Watanabe, the staffs of Research Institute for Applied Mechanics, Kyushu University, for their great cooperation in the experiments.

The research was financially supported by the Grant-in-Aids for Scientific Research (A), No. 24246161, sponsored by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan (for the last author).

References

[1] J. Twidell, T. Weir, Renewable Energy Resources, Routledge, 2015.
[2] T.T. Chow, G. Tiwari, C. Menezo, Hybrid solar: a review on photovoltaic and thermal power integration, Int. J. Photoenergy (2012).
[3] K. Watanabe, S. Fukutomi, Y. Ohya, T. Uchida, An ignored wind generates more electricity: a solar updraft tower to a wind solar tower, Int. J. Photoenergy (2020).
[4] A. Dhahri, A. Omri, A review of solar chimney power generation technology, Int. J. Eng. Adv. Technol. 2 (3) (2013) 1–17.
[5] X. Zhou, F. Wang, R.M. Ochieng, A review of solar chimney power technology, Renew. Sustain. Energy Rev. 14 (8) (2010) 2315–2338.
[6] J. Schlaich, The Solar Chimney: Electricity from the Sun, Edition Axel Menges, 1995.
[7] R. Wengenmayr, T. Bührke, Renewable Energy: Sustainable Energy Concepts for the Future, John Wiley & Sons, 2011.
[8] J.r. Schlaich, R. Bergermann, W. Schiel, G. Weinrebe, "Design of commercial solar updraft tower systems—utilization of solar induced convective flows for power generation, J. Sol. Energy Eng. 127 (1) (2005) 117–124.
[9] R. Richards, et al., Spanish solar chimney nears completion, MPS Rev. 6 (1981) 21–23.
[10] W. Haaf, K. Friedrich, G. Mayr, J. Schlaich, Solar chimneys part i: principle and construction of the pilot plant in manzanares, Int. J. Sol. Energy 2 (1) (1983) 3–20.
[11] W. Haaf, Solar chimneys: part ii: preliminary test results from the manzanares pilot plant, Int. J. Sustain. Energy 2 (2) (1984) 141–161.
[12] A. Asnaghi, S. Ladjevardi, Solar chimney power plant performance in Iran, Renew. Sustain. Energy Rev. 16 (5) (2012) 3383–3390.
[13] M.R. Ahmed, S.K. Patel, Computational and experimental studies on solar chimney power plants for power generation in pacific island countries, Energy Convers. Manag. 149 (2017) 61–78.
[14] K.M. Shirvan, S. Mirzakhanlari, M. Mamourian, N. Abu-Hamdeh, Numerical investigation and sensitivity analysis of effective parameters to obtain potential maximum power output: a case study on zanjan prototype solar chimney power plant, Energy Convers. Manag. 136 (2017) 350–360.
[15] N. Jafarifar, M.M. Behzadi, M. Yaghini, The effect of strong ambient winds on the efficiency of solar updraft power towers: a numerical case study for orkney, Renew. Energy 136 (2019) 937–944.
[16] Y. Ohya, M. Wataka, K. Watanabe, T. Uchida, Laboratory experiment and numerical analysis of a new type of solar tower efficiently generating a thermal updraft, Energies 9 (12) (2016) 1077.
[17] S. Okada, T. Uchida, T. Karasudani, Y. Ohya, Improvement in solar chimney power generation by using a diffuser tower, J. Sol. Energy Eng. 137 (3) (2015).
[18] M. Motoyama, K. Sugitani, Y. Ohya, T. Karasudani, T. Nagai, S. Okada, et al., Improving the power generation performance of a solar tower using thermal updraft wind, Energy Power Eng. 6 (11) (2014) 362.
[19] J. Li, P. Guo, Y. Wang, Preliminary investigation of a novel solar and wind energy extraction system, Proc. IME J. Power Energy 226 (1) (2012) 73–85.
[20] L. Zuo, L. Ding, J. Chen, X. Zhou, B. Xu, Z. Liu, Comprehensive study of wind supercharged solar chimney power plant combined with seawater desalination, Sol. Energy 166 (2018) 59–70.
[21] C.M. Bishop, Pattern Recognition and Machine Learning, Springer-Verlag, New York, 2006.
[22] P. Baldi, S. Brunak, Bioinformatics: the Machine Learning Approach, 2 ed., MIT press, Cambridge, MA, 2001.
[23] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Comput. Struct. Biotechnol. J. 13 (2015) 8–17.
[24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (Oct) (2011) 2825–2830.
[25] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in pytorch, in: 31st Conference on Neural Information Processing Systems (NIPS 2017), (Long Beach, CA, USA), 2017.
[26] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: a system for large-scale machine learning, in: 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), 2016, pp. 265–283.
[27] N. Pasumarthi, S. Sherif, "Experimental and theoretical performance of a demonstration solar chimney model—part i: mathematical model development, Int. J. Energy Res. 22 (3) (1998) 277–288.
[28] N. Pasumarthi, S. Sherif, "Experimental and theoretical performance of a demonstration solar chimney model—part ii: experimental and theoretical results and economic analysis, Int. J. Energy Res. 22 (5) (1998) 443–461.
[29] A. Clifton, L. Kilcher, J.K. Lundquist, P. Fleming, Using machine learning to predict wind turbine power output, Environ. Res. Lett. 8 (2) (2013), 024009.
[30] J. Heinermann, O. Kramer, Machine learning ensembles for wind power prediction, Renew. Energy 89 (2016) 671–679.
[31] N.A. Treiber, J. Heinermann, O. Kramer, Wind power prediction with machine learning, in: Computational Sustainability, Springer, 2016, pp. 13–29.
[32] Z. Ti, X.W. Deng, H. Yang, Wake modeling of wind turbines using machine learning, Appl. Energy 257 (2020) 114025.
[33] M.A. Rushdi, A.A. Rushdi, T.N. Dief, A.M. Halawa, S. Yoshida, R. Schmehl, Power prediction of airborne wind energy systems using multivariate machine learning, Energies 13 (9) (2020) 2367.
[34] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: a review, Renew. Energy 105 (2017) 569–582.
[35] N. Sharma, P. Sharma, D. Irwin, P. Shenoy, Predicting solar generation from weather forecasts using machine learning, in: 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, 2011, pp. 528–533.
[36] S. Kalogirou, Artificial Intelligence in Energy and Renewable Energy Systems, Nova Publishers, 2007.
[37] K. Watanabe, Y. Ohya, Performance of wind solar tower that utilizes solar heat and wind simultaneously, Proc. Japan Wind Energy Sympos. 40 (2018) 436–439.
[38] J.-y. Li, P.-h. Guo, Y. Wang, Effects of collector radius and chimney height on power output of a solar chimney power plant with turbines, Renew. Energy 47 (2012) 21–28.
[39] A. Koonsrisuk, T. Chitsomboon, Effects of flow area changes on the potential of solar chimney power plants, Energy 51 (2013) 400–406.
[40] P. Guo, J. Li, Y. Wang, Y. Wang, Evaluation of the optimal turbine pressure drop ratio for a solar chimney power plant, Energy Convers. Manag. 108 (2016) 14–22.
[41] S. Hu, D.Y. Leung, J.C. Chan, Impact of the geometry of divergent chimneys on the power output of a solar chimney power plant, Energy 120 (2017) 1–11.
[42] J. Benesty, J. Chen, Y. Huang, I. Cohen, Pearson correlation coefficient, in: Noise Reduction in Speech Processing, Springer, 2009, pp. 1–4.
[43] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (Mar) (2003) 1157–1182.
[44] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, vol. 821, John Wiley & Sons, 2012.
[45] S. Ruder, An Overview of Gradient Descent Optimization Algorithms, 2016 arXiv preprint arXiv:1609.04747.
[46] Common pitfalls in interpretation of coefficients of linear models. https://scikit-learn.org/stable/auto_examples/inspection/plot_linear_model_coefficient_interpretation.html. Accessed 25 Jan 2021.

Mostafa A. Rushdi, Ph.D. in Kyushu University, BSc, and MSc in (Aerospace Engineering, Cairo University). Postdoctoral researcher, Research Institute for Applied Mechanics (RIAM), Kyushu University. Instructor in Mechanical Engineering, Future University (FUE) - On leave. Research Areas: System Dynamics, Optimization, and AI (ML-DL).

Dr. Engineer **Shigeo Yoshida**. Professor, Research Institute for Applied Mechanics (RIAM), Kyushu University. Research Area: Wind Energy.

Dr. Engineer **Koichi Watanabe**. Associate Professor, Platform of Inter/Transdisciplinary Energy Research (Q-PIT), Kyushu University. Research Areas: Wind Energy, Fluid Dynamics.

Dr. Engineer **Yuji Ohya**. Professor, Research Institute for Applied Mechanics (RIAM), Kyushu University. Research Areas: Wind Energy, Fluid Dynamics.