Contents list available at IJRED website

International Journal of Renewable Energy Development

Journal homepage: https://ijred.undip.ac.id

Research Article

Consistent Regime-Switching Lasso Model of the Biomass Proximate Analysis Higher Heating Value

Akara Kijkarncharoensin* and Supachate Innet

Department of Computer Engineering and Financial Technology, School of Engineering, University of the Thai Chamber of Commerce, Bangkok, Thailand

Abstract. Prediction accuracy is crucial for higher heating value (HHV) models to promote renewable biomass energy, especially its consistency is crucial when retraining data and knowledge of the range are unavailable. Current HHV models lack consistency in accuracy and interpretability due to various reasons. Thus, this study aimed to construct an interpretable and consistent proximate-based biomass HHV model on a wide-range dataset. The model, regime-lasso, integrated the concepts of regime-switching, lasso regression, and federated averaging to construct a consistent HHV model. The regime-switching partitioned the dataset into optimal regimes, and the lasso trained the regime models. The regime-lasso model is a collection of these models. It provided root mean square error of 0.4430– 0.9050, mean absolute error of 0.2743–0.6867, and average absolute error of 1.512–4.5894% in the literature's wide-range datasets. The Kruskal–Wallis test confirmed the in-sample performance consistency at α=0.05, regardless of the training sets. In the out-of-sample situations without retraining, the model preserved its accuracy in six out of 11 datasets at α = 0.01. The interpretability of regime-lasso indicated the regime characteristic to be a factor of inconsistent prediction. The increase in FC had the maximum positive impact on HHV in the 2nd and 3rd regimes, while the increase in ASH negatively impacted the 1st and 2nd regimes. VM variation had neutral effects in all regimes. The regime-lasso solves the issues of accuracy declination and addresses the challenges in sensitivity analysis of the HHV model. The prediction accuracy issues of the model's direct implementation were fixed.

Keywords: clustering, consistency, federated averaging, HHV, interpretability, Kruskal-Wallis, prediction, regression, sensitivity

Received: 22nd July 2022; Revised: 8th September 2022; Accepted: 27th September 2022; Available online: 14th Oct 2022

1. Introduction

Biomass is a crucial sustainable energy source to reduce carbon dioxide levels, greenhouse gases, and dependency on fossil fuels (Core Writing Team 2015). Proximate and ultimate analyses are the main approaches to evaluate biomass energy yields; the former utilizes the percentage fraction of fixed carbon (FC), volatile matter (VM), and ash (ASH), while the latter applies the chemical element to estimate the higher heating value (HHV). The ultimate-based models have superior accuracy but require sophisticated equipment and are time consuming. The proximate-based models are more practical. Previous studies have tried to improve the HHV prediction accuracy of these models.

The mathematical proximate-based HHV model was extended from carbonaceous to lignocellulosic materials (Cordero *et al.* 2001) in specific locations. Then, the research field was expanded to several materials: solid fuels (Parikh *et al.* 2005), lignocellulosic compound biomass (Chun-Yang Yin 2011), non-woody biomass torrefaction char (Soponpongpipat *et al.* 2015), torrefied biomass (Nhuchhen and Afzal 2017), poultry waste (Qian *et al.* 2018), and biochar (Qian *et al.* 2020). These models were empirical correlations based on lower-order ordinary least squares regressions (OLS) and represent the interpretable relationship between sets {FC, VM, ASH}, called the *property* set, and HHV. These OLS models provided accurate prediction values. Their average absolute error (AAE) and R-squared (R^2) were 2.00–6.17% and 0.916–0.929, respectively. These models were practical, accurate, and interpretable; however, their validations were limited to a specific range of property sets.

Nhuchhen and Abdul Salam (2012) studied a wide-range proximate-based model. They constructed the fourth order OLS between sets {VM/FC, ASH/VM, FC/ASH}, called the *ratio* set, and HHV; the wide-range proximate-based models were documented by Mohammed *et al.* (2014). Further, the non-linear machine learning (ML) models successfully constructed wide-range proximate-based models.

Akkaya (2013) implemented an artificial neural network (ANN) to predict coal HHV. Subsequently, Ghugare *et al.* (2014) predicted a wide-range solid biomass HHV through genetic programming (GP). Various models have been developed to support wide-range materials: adaptive neuro-fuzzy inference system (ANFIS) (E. Akkaya 2016), ANN (Uzun *et al.* 2017), and genetic algorithms radial basis function (GA-RBF) (Dashti *et al.* 2019). Their AAE, R^2, and root mean square error (RMSE) were 2.64–5.0%, 0.8836–0.9852, and 0.375–1.3006, respectively. Estiati *et al.* (2016) confirmed the superior accuracy of ANN to lower-order OLS on wide-range data and reported the effects of sample number on the model accuracy. Recent studies have used superior ML models to predict HHV on specific materials: municipal solid waste (Taki and Rohani 2022).

* Corresponding author
Email: akara_kij@live4.utcc.ac.th (A. Kijkarncharoensin)

https://doi.org/10.14710/ijred.2023.47831
ISSN: 2252-4940/© 2023.The Author(s). Published by CBIORE

Some disagreements on the model's accuracy arose in comparison studies. Xing et al. (2019) used random forest (RF) models that benefited from bootstrap sampling to predict the HHV model and generated an R^2 of 0.94. They compared RF, ANN, support vector machine (SVM), and OLSs and insisted on the superiority of MLs to OLSs in terms of accuracy. They found that OLSs provided $R^2 < 0.7$, which was much lower than those reported in earlier studies. The ANN's R^2 and RMSE were 0.856 and 4.669, respectively, as compared to an R^2 of 0.963 and RMSE of 0.375 by Uzun et al. (2017). Thus, the decline in accuracy was apparent. The Monte-Carlo sensitivity analysis reported that ASH, VM, and FC have maximum, medium, and minimum effects. Samadi *et al.* (2021) were skeptical that the overfitting issues would render MLs to exhibit superior accuracy. Thus, they selected the gradient-boosted regression tree (GBRT), an unsupervised machine learning approach, to construct a proximate-based HHV model. Their AAE and R^2 were 3.783% and 0.93, compared to the OLS's AAE of 4.66–10.72%, R^2 of 0.43–0.89, and RMSE of 1.121–2.410 in their experiments. Their linear correlation indicated a more significant impact of FC on HHV than of ASH. The non-interpretation of ML models renders sensitivity analysis challenging, resulting in diversity. The linear models are interpretable but limited in their specific range of materials. Some factors affect the linear models' validity. The sensitivity analyses can be improved if an interpretable linear model is made in the wide-range material.

An accurate and consistent model to predict the HHV for biomass is necessary to promote renewable energy handling, especially in data-poor regions that must implement the model without retraining. The wide-range models are crucial as the range information is absent. Research focusing on consistency of HHV prediction, even in the ultimate-based model, is lacking. Kijkarncharoensin and Innet (2022a) examined and reported the decreasing accuracy, called *performance inconsistency*, of the ultimate-based models. The models can maintain accuracy in neither in-sample nor out-of-sample conditions and must be retrained before implementation. Boumanchar et al. (2019) also documented the ultimate-based models' accuracy declination. These studies also demonstrated that accuracy-decreasing issues emerged when implementing the models on the out-of-sample dataset.

This study's objective was to construct an interpretable and consistent proximate-based HHV model of wide-range biomass material. The interpretation refers to identifying the features' degree of sensitivity to the response. Consistency is the statistical indifference of the error distributions on distinct datasets. The strong definition of consistency is the similarity in accuracy ranking. This study is the first to demonstrate that an HHV model can statistically maintain the latter definition among the datasets. The research challenges are to identify the source of inconsistency that rendered the linear models unsuitable in a wide range of conditions and to prove the invariant accuracy of the proposed model.

Three procedures were conducted to achieve the goals. First was to integrate the concepts of regime-switching, least absolute shrinkage and selection operator (lasso), and federated averaging (FedAvg) into the regime-lasso model. The regime-switching, lasso, and FedAvg concepts were embedded into the biomass HHV models for the first time. Second was to statistically test the performance consistency. The article set the formal experiments and extended Samadi *et al.* (2021), covering 12 models from 2001–2020 Yr. presented in Section 3.2. Two definitions of performance consistency were defined based on the accuracy rankings and error distributions, as indicated in section 3.3. The Kruska–Wallis H-test determined these distributions' equivalence and verified the consistency of the performance. Third was to ensure the estimators' unbiasedness. The Kolmogorov–Sirmanov test examined the distributions' normality and guaranteed unbiased estimation of FedAvg within a number of samplings.

The remaining sections are organized as follows: Section 2 describes the regime-lasso model in detail and presents the schemes of k-means clustering, lasso, and FedAvg; Section 3 exhibits the experiments' schematic diagrams to evaluate the accuracy of the regime-lasso model, compares it with earlier studies, and tests its consistency; Section 4 indicates and discusses the performance evaluation of five measurements. The experimental results of Section 3 are expressed and discussed here; and lastly, the article ends with the research conclusion.

2. Theoretical models

The fundamentals of the regime-lasso concepts are the dataset's clustering, model's training, and aggregation of local to global model. Economics implement K-means-clustering to cluster the data into economic states (Liao 2017), called *regimes*. Kijkarncharoensin and Innet (2022b) embedded this idea with the correlation distance metric to classify biomass datasets into uncorrelated regimes treated as different populations. The HHV models can be separately trained in these regimes; however, the consistency improvement should solve the feature collinearity issue and the estimated coefficient error.

The collinearity in the models' feature set affects the OLS's variance (Mela and Kopalle 2002). However, correlation distance clustering increases the degree of collinearity within the regime. Therefore, the OLS-based model cannot be considered as these regimes' unique minimum variance unbiased estimator.

The lasso conditionally searches the OLS's hypothesis space solutions regardless of the collinearity (Tibshirani 1996). This model intelligently sacrifices some bias for robustness by shrinking some coefficients to zero. The model's variance reduction improves its overall accuracy. Thus, the lasso can efficiently solve the collinearity issue and train the HHV model in the regimes.

The FedAvg, a federated learning algorithm, can handle the estimated coefficient error. It distributes the training across the decentralized database and aggregates them into a new model through averaging (Hard *et al.* 2018). According to the central limit theorem, the normality of the linear model's error distribution leads to the mean of the estimated coefficient being the true value within a finite number of samples. If the errors follow a normal distribution, FedAvg guarantees that the average linear model coefficient is an unbiased estimator within a finite number of bootstrapped.

2.1 K-means clustering

K-means clustering is a type of unsupervised learning that groups the dataset into k similar clusters. Let $\boldsymbol{x}\epsilon\mathbb{R}^{\boldsymbol{p}}$ be a p-dimension data and $\boldsymbol{c}\epsilon\mathbb{R}^{\boldsymbol{p}}$ be a centroid of those k clusters. The correlation distance metric, as given in Equation (1), measures the interval between $\boldsymbol{x}$ and the k clusters' centroids $\boldsymbol{c}$.

$$d(x,c) = 1 - \frac{(x-\vec{x})'(c-\vec{c})}{\sqrt{(x-\vec{x})'(x-\vec{x})}\sqrt{(c-\vec{c})'(c-\vec{c})}} \quad (1)$$

Let a_i be the average intracluster distance between the i[th] point and its centroid, while b_i be the average inter-cluster distance of that point and other centroids. Silhouette value, S_i, in Equation (2), defined as the ratio of a_i and b_i, measures the distance between clusters. S_i has a [–1, 1] range; the higher, the better. The excellent partitions should have S_i above 0.8.

$$S_i = 1 - \frac{a_i}{b_i} \tag{2}$$

The k-means clustering starts by randomly selecting the initial k centroids and using Equation (1) to group the data. Then, the algorithm calculates new k centers' positions and iteratively groups the data based on these centroids. These iterative processes run until all the cluster's centers remain stable. .

2.2 Least absolute shrinkage and selection operator (Lasso)

The lasso model minimizes the sum of square errors subject to its absolute coefficients' summation (3), equivalent to the Lagrangian objective function (4).

$$\min_{\beta} \frac{1}{2N}\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 \quad subject\ to\ \sum_{j=1}^{p}|\beta_j| \le t \tag{3}$$

$$\mathcal{L}(\beta) = \frac{1}{2N}\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{4}$$

Suppose x_{ij} are standardized: $\sum_{i=1}^{N} x_{ij} = 0, \frac{1}{N}\sum_{i=1}^{N} x_{ij}^2 = 1$, so $\beta_0 = \bar{y}$. Then, coordinate descent fast algorithms (Friedman *et al.* 2010) solve Equation (4) through Equation (5)–(7). $\hat{y}_i^{OLS}$ is the OLS response, thus $\tilde{y}_i^{(j)}$ in Equation (6) is the OLS estimator excluded x_{ij}'s contribution. Function $(\cdot)^+$refers to the positive value. Choose the pre-determine λ such that minimizes MSE in the testing set.

$$\tilde{\beta}_j \leftarrow S\left(\frac{1}{N}\sum_{i=1}^{N} x_{ij}\left(y_i - \tilde{y}_i^{(j)}\right), \lambda\right) \tag{5}$$

$$\tilde{y}_i^{(j)} = \hat{y}_i^{OLS} - x_{ij}\tilde{\beta}_j \tag{6}$$

$$S(z,\lambda) = sign(z)(|z| - \lambda)^+ \tag{7}$$

The coordinate-wise update determines the lasso coefficient β instead of explicitly using the OLS estimation. Therefore, the lasso can handle the collinearity issues in the predictor x and reduces the estimator's variability.

2.3 Federated averaging

FedAvg is a distributed optimization algorithm that trains the models through the decentralized dataset and aggregates these models into one global model. FebAvg is an improved version, robust to unbalanced and non-IID distributed local datasets (McMahan *et al.* 2017).

Let $\tilde{\beta}^{r,k} \epsilon \mathbb{R}^p$; k = 1 to K, be the k[th] fold lasso coefficient of the regime; r = 1 to R, and $\tilde{\beta}^r$ be the aggregated coefficient of that regime. The regime-lasso coefficients of the r[th] regime are the aggregation of $\tilde{\beta}^{r,k}$, as follows in Equation (8).

$$\tilde{\beta}_j^r \leftarrow \frac{1}{K}\sum_{k=1}^{K} \tilde{\beta}_j^{r,k} \tag{8}$$

The k-mean clustering divides the article dataset into R regimes. Then, the lasso model runs on all R regimes to generate R equations. The K-fold cross-validation iteratively evaluates these R models and develops K × R models. Finally, FedAvg uses Equation (8) to integrate them into the lasso models for these R regimes.

The clustered datasets may be the non-IID distribution. However, if the $\tilde{\beta}^r$ is normally distributed, Equation (8) guarantees the unbiased estimator on the finite number of cross-validations. The regime-lasso model is the set of these R regressions.

3. The experimental design

This study classified the article dataset (Kijkarncharoensin 2022b) through the k-means clustering algorithm on correlation distance [Equation (1)]. The maximum silhouette value from Equation (2) was a proxy for the optimal cluster number. The lasso model was separately trained on these clusters and aggregated to be the regime-lasso model. Then, two experiments measured the model performances through Equations (9)–(13) and compared them with literature.

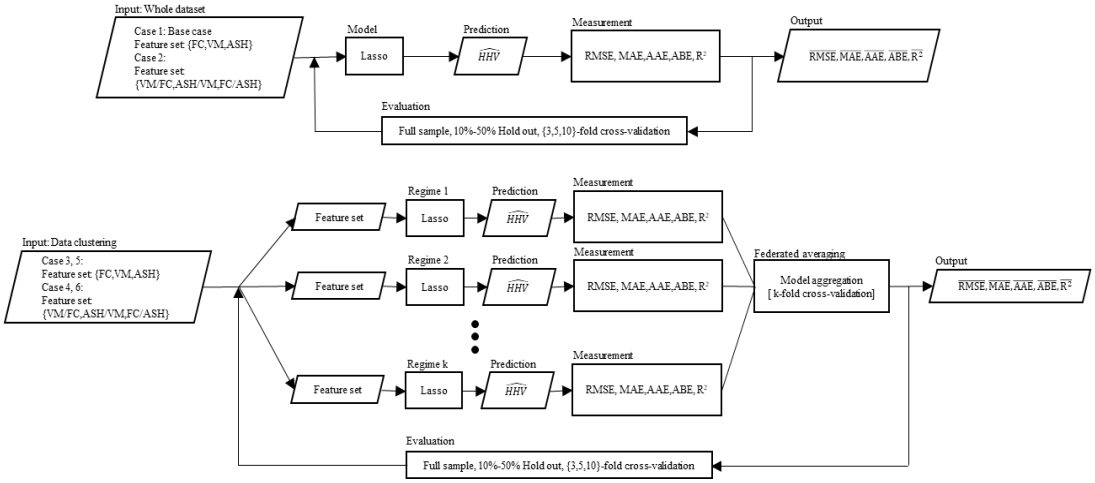

Fig. 1 Schematic flowchart representing the experimental design to evaluate the regime-lasso model; fixed carbon (FC), volatile matter (VM), ash (ASH), higher heating value (HHV), root mean square error(RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), and R-squared (R^2).

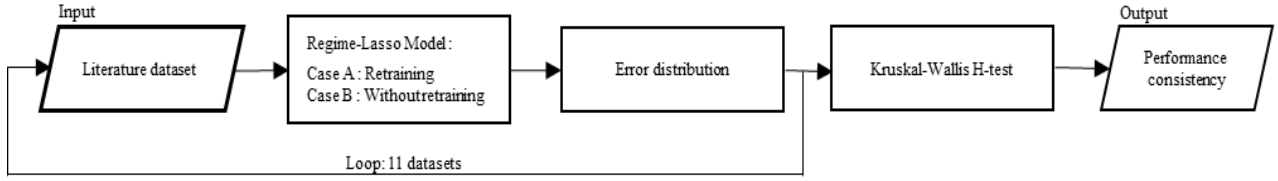

Fig. 2 Schematic flowchart representing the statistical testing of indifference in the error distributions. The distribution equivalence is a proxy of performance consistency in the article.

Table 1
Literature models and datasets for the regime-lasso model's accuracy comparison and consistency examination.

	HHV model	Dataset Size	Dataset Source
I	$-30.3FC^2+62.5ASH^2+55.4FC-48.5ASH+9.591$	33	C. Qian *et al.* (2020)
II	$140.2-1.167FC-0.210VM-1.558ASH-0.02739VM^2+0.000191VM^3+0.00104FC\times ASH$	37	X. Qian *et al.* (2018)
III	$0.1846VM+0.3525FC$	207	Nhuchhen and Afzal (2017)
IV.A	$0.3368FC+0.1646VM+0.0113ASH$	12	Estiati *et al.* (2016)
IV.B	$0.3241FC+3.7667\times10^{-4}FC^2-4.1530\times10^{-4}FC\times VM+0.1947VM-2.8207\times10^{-4}VM^2-0.0025ASH$		
V	$35.4879-0.3023ASH-0.1905VM$	28	Soponpongpipat *et al.* (2015)
VI	$0.365FC+0.131VM+\frac{1.397}{FC}+\frac{328.568VM}{10283.138+0.531ASH\times FC^3-6.863\times ASH\times FC^2}$	382	Ghugare *et al.* (2014)
VII	$1.9999+0.248FC+0.162VM-0.137ASH$	-	Mohammed *et al.* (2014)
VIII	$20.7999-0.3214\frac{VM}{FC}+0.0051\left(\frac{VM}{FC}\right)^2-11.2277\frac{ASH}{VM}+4.4953\left(\frac{ASH}{VM}\right)^2-0.7223\left(\frac{ASH}{VM}\right)^3+0.0383\left(\frac{ASH}{VM}\right)^4+0.0076\frac{FC}{ASH}$	250	Nhuchhen and Abdul Salam (2012)
IX	$0.2521FC+0.1905VM$	53	Chun-Yang Yin (2011)
X	$0.3536FC+0.1559VM-0.0078ASH$	100	Parikh *et al.* (2005)
XI	$354.3FC+170.8VM$	24	Cordero *et al.* (2001)
XII	This study regime-lasso model	802	Kijkarncharoensin (2022b)

3.1 Regime-lasso analysis

Two feature sets were fitted into the model. The first one was the proximate analysis data, {FC, VM, ASH}, called *property*; and the other was the ratio of these data, {VM/FC, ASH/VM, FC/ASH}, called the *ratio* implemented in Nhuchhen and Abdul Salam, 2012. The experiment base case was the lasso model of the *property* set trained on the entire sample.

The experiment exhibited in Fig. 1 implements the *ratio* set on the base case, called case 2, to identify the improvement. Next, k-mean clustering divided the article dataset into k regimes. Two sets of {FC, VM, ASH, HHV} and {VM/FC, ASH/VM, FC/ASH, HHV} classified the dataset through the correlation distance metric. Then, the lasso models with the *property* set were applied to the clusters generated from {FC, VM, ASH, HHV} and named case 3 (with two regimes) and case 5 (with three regimes). Similarly, case 4 (with two regimes) and case 6 (with three regimes) were the lasso with the *ratio* set on the regimes generated by {VM/FC, ASH/VM, FC/ASH, HHV}. Cases 3–6 were examined for accuracy improvement after embedding the regime-switching concepts.

These experiments involved 10 evaluation schemes, the full sample substitution, p% hold-out, and k-fold cross-validation. The average of RMSE, MAE, AAE, ABE, and R^2 measure the prediction accuracy. The following section compares the most accurate model with the literature models.

3.2 Literature comparison

The experiments compared the regime-lasso model, evaluated in the previous section, with 12 literature models (Chun-Yang Yin 2011, Cordero *et al.* 2001, Estiati *et al.* 2016, Ghugare *et al.* 2014, Mohammed *et al.* 2014, Nhuchhen and Abdul Salam 2012, Nhuchhen and Afzal 2017, Parikh *et al.* 2005, C. Qian *et al.* 2020, X. Qian *et al.* 2018, Soponpongpipat *et al.* 2015) listed in Table 1. The 11 examination datasets were the models' training data reported in the literature. The regime-lasso should outperform these literature models in their environments.

The experiment ended with the performance comparison conducted on the article dataset. The performance comparison relied on a total sample, which promoted reproducibility and improved the model performance as high as possible. The performance consistency of the proposed model was verified as a result.

3.3 Testing of the performance consistency

The article defined two definitions of performance consistency. The first was the similar ranking in prediction accuracy, while the other was the indifference in the error distributions.

> *Definition 1 Performance consistency in the ranking:* The performance of an HHV prediction model is consistent if its accuracy rankings in all testing datasets are the same.
>
> *Definition 2 Performance consistency in the error distribution:* The performance of an HHV prediction model is consistent if its error distributions in all testing datasets are statistically indifferent.

The consistency examinations of these definitions were conducted differently. The study assumed that the early study models were the most accurate in their training sets. The proof of definition 1 was straightforward if the model outperformed them in their datasets. Definition 2's verification proceeded through the hypothesis testing of the distribution similarity. For example, the Kruskal–Wallis H-Test, a nonparametric ANOVA, inspects the prediction errors based on the null of the distribution equivalence. Fig. 2 presents the schematic diagram to prove the 2nd definition.

The proof of definition 2 was conducted in two situations. The first situation retrained the model in Table 1 before performance measuring. A total of 11 datasets were involved in

the experiments, except the dataset of Mohammed *et al.* (2014), as it was unpublished. The other situation was to implement the model in Section 3.1 without retraining. The hypothesis testing results in Section 4.4 emphasize the accuracy and consistency of the regime-lasso model. The author publishes the source code (Kijkarncharoensin 2022a) to promote the experiment's reproducibility.

3.4 Evaluation indicators

The article measures prediction accuracy through five indicators; RMSE, mean absolute error (MAE), AAE, average bias error (ABE), and R^2.

$$RMSE = \sqrt{MSE} = \sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (9)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \quad (10)$$

$$AAE = \frac{100}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \quad (11)$$

$$ABE = \frac{100}{N}\sum_{i=1}^{N}\frac{y_i - \hat{y}_i}{y_i} \quad (12)$$

$$R^2 = 1 - \frac{SSE}{SST} \quad (13)$$

RMSE is the standard error. MAE is the absolute value, while AAE is the relative one. ABE measures relative bias. R^2 refers to the ratio of error sum of squares (SSE) and total sum of squares (SST). In addition, the study implemented the Kolmogorov-Smirnov test to inspect the residual distribution.

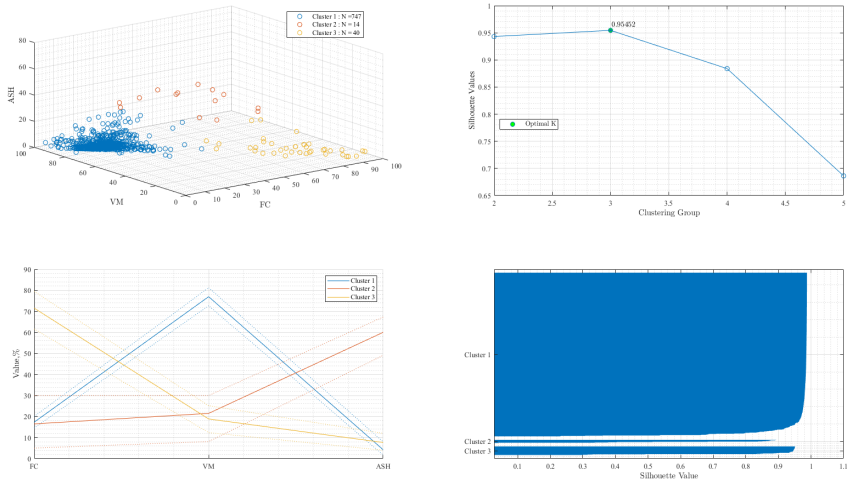

Fig. 3 K-means clustering analysis on the set of {FC, VM, ASH, HHV} and correlation distance metric. The correlation-based centroids are c = [c1 c2 c3]^T, where c1 = [–0.30 0.80 –0.50]^T, c2 = [–0.33 –0.35 0.69]^T, and c3 = [0.76 –0.26 –0.50]^T; fixed carbon (FC), volatile matter (VM), ash (ASH), and higher heating value (HHV).

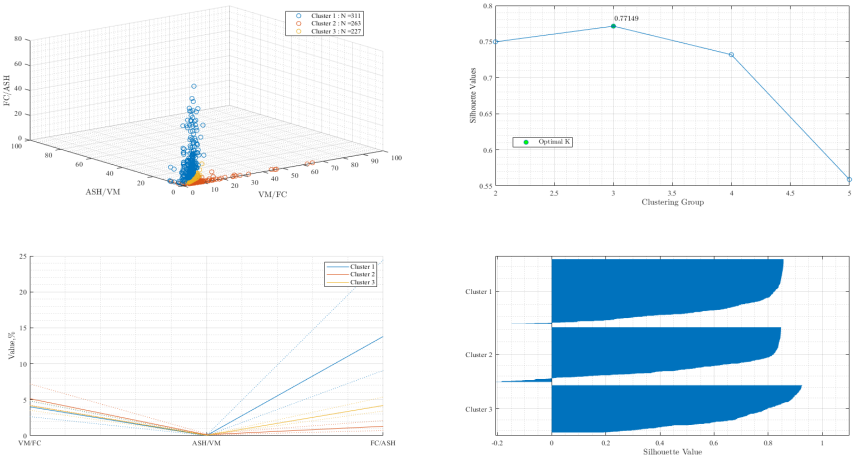

Fig. 4 K-means clustering analysis on the set of {VM/FC, ASH/VM, FC/ASH, HHV} and the correlation distance metric. The correlation-based centroids are c = [c1 c2 c3]^T, where c1 = [–0.22 –0.53 0.75]^T, c2 = [0.75 –0.53 –0.21]^T, and c3 = [0.39 –0.80 0.41]^T; fixed carbon (FC), volatile matter (VM), ash (ASH), and higher heating value (HHV).

4. Results and Discussion

The experimental results are composed of four parts. The first Section 4.1 indicates and inspects the practical techniques to classify the data into thermal regimes. Section 4.2 constructs the regime-lasso model and evaluates its accuracy based on the schematic flowchart in Section 3. The sensitivities of the regime's thermal properties on the HHV are compared with those of previous studies. The sources of the inconsistency are identified. Section 4.3 compares the regime-lasso's in-sample accuracy with that of the literature models. The out-of-sample accuracies of the earlier study's models are examined. Section 4.4 presents the statistical consistency in the regime-lasso's accuracy. The regime-lasso's invariant accuracies are proved, and the study's objectives on the consistency and interpretability of the HHV model are addressed.

4.1 Regime clustering

The data clustering indicates that three regimes were optimal for the present study dataset. The set of {FC, VM, ASH, HHV} provided 0.9545 silhouette value, which was higher than that of {VM/FC, ASH/VM, FC/ASH, HHV}. The cluster analysis of the former set is expressed in Fig. 3, while that of the latter is shown in Fig. 4.

The left and right plots in the lower panel of Figs 3 & 4 are the parallel coordinate and silhouette plots, respectively. The former plot presents the 25% quantile of the properties, while the latter indicates the regimes' data similarity through the silhouette value; its range is [–1, 1]; the higher the better. The silhouette values of all the data in each cluster were higher than 0.8, which was the criterion. Therefore, {FC, VM, ASH, HHV} can distinctly cluster the article dataset effectively. The negative silhouette value in Fig. 4 indicates the dissimilarity in the cluster data. Additionally, overlaps in the parallel coordinates insisted on the inappropriate clustering received from {VM/FC, ASH/VM, FC/ASH}. The results supported Kijkarncharoensin and Innet (2022b) using {FC, VM, ASH, HHV} to cluster the data into regimes.

Table 2
Correlation matrix of the whole article dataset and its regime, indicated in Fig. 2; Sample size (N), fixed carbon (FC), volatile matter (VM), ash (ASH), and higher heating value (HHV).

Full Sample	FC	VM	ASH	HHV
FC	1.0000	-0.7844	-0.0706	0.7405
VM	-0.7844	1.0000	-0.5601	-0.3007
ASH	-0.0706	-0.5601	1.0000	-0.5009
HHV	0.7405	-0.3007	-0.5009	1.0000
Cluster 1: N = 747				
	FC	VM	ASH	HHV
FC	1.0000	-0.4411	-0.3382	0.3331
VM	-0.4411	1.0000	-0.6855	0.3842
ASH	-0.3382	-0.6855	1.0000	-0.6745
HHV	0.3331	0.3842	-0.6745	1.0000
Cluster 2: N = 14				
	FC	VM	ASH	HHV
FC	1.0000	-0.6363	-0.4480	0.9020
VM	-0.6363	1.0000	-0.4046	-0.2919
ASH	-0.4480	-0.4046	1.0000	-0.7311
HHV	0.9020	-0.2919	-0.7311	1.0000
Cluster 3: N = 40				
	FC	VM	ASH	HHV
FC	1.0000	-0.7572	-0.5734	0.8618
VM	-0.7572	1.0000	-0.0932	-0.4873
ASH	-0.5734	-0.0932	1.0000	-0.6663
HHV	0.8618	-0.4873	-0.6663	1.0000

Table 2 presents the correlation matrix of the regime's feature sets and indicates high collinearity among them. These highly linear relationships arose from clustering by correlation distance metric, which generated a high correlation between the response HHV and the *property* set. The entire sample exhibited a correlation of 0.7405 between FC and HHV. The correlation increased to 0.9020 and 0.8618 in the 2nd and 3rd regimes, respectively, after clustering.

The lasso regression was implemented coordinate-wise as given in Section 2.2, to handle the feature collinearities. Accuracy improvement from training the lasso into these regimes can be expected.

4.2 Model evaluation

A total of 10 evaluation schemes were conducted on the lasso model to measure the prediction accuracy. The average value of 10 schemes, a complete substitution, six sets of hold-out, and three sets of k-fold cross-validation were considered proxies of the model performance. Table 3 indicates the base and second cases' mean values of the five indicators (RMSE, MAE, AAE, ABE, R^2). The base case provided an average RMSE, MAE, AAE, ABE, and R^2 as 1.449, 1.0647, 5.8088%, 0.0166%, and 0.7985, respectively. Likewise, these parameters for the second case were 2.5821, 1.6505, 9.0535%, –0.1208%, and 0.3603, respectively. The *ratio* set did not improve prediction accuracy because the base case provided a lower magnitude of RMSE, MAE, AAE, and ABE and a higher one of R^2 than did the second one. The KS-test informed the residual non-normality in both cases with a 0.0000 p-value.

The accuracy improvement due to regime-switching is expressed in Table 4 for two regimes and in Table 5 for three, the optimal number of clusters. Case 3 in Table 4 generated mean values of 1.4168, 1.0455, 5.7543%, –0.0050%, and 0.8076 for RMSE, MAE, AAE, ABE, and R^2, respectively, while these values for case 4 were 1.7495, 1.2386, 6.6772%, 0.1168%, and 0.7065, respectively. Case 5 in Table 5 presents the mean RMSE, MAE, AAE, ABE, and R^2 of 1.3933, 1.0249, 5.5786%, 0.0300%, and 0.8139, respectively. Case 6 in Table 5 generated average RMSE, MAE, AAE, ABE, and R^2 of 1.7313, 1.2208, 6.5565%, 0.0323%, and 0.7125, respectively. Cases 3 and 4 in Table 4 outperformed the base and second cases. The accuracies of cases 5 and 6 in Table 5 were better than those in Table 4. The optimal regime cases provided a lower mean of these accuracy indicators and a higher mean of R^2. The regime-switching concept improved the prediction power of the *property* and *ratio* sets of the lasso model. Case 5 provided a lower values for RMSE of 0.338, MAE of 0.1959, AAE of 0.9779%, ABE of 0.0023%, whereas a higher R^2 of 0.1014 compared to those of case 6. The *ratio* set did not improve the prediction accuracy, contrasting with Nhuchhen and Abdul Salam (2012).

The regime-lasso models exhibited the highest accuracies at the optimal number of regimes on the *property* set. Case 5, constructed on the *property* set and optimal regime number, presented the highest accuracies in the experiments. An average p-value of 0.0281 of its KS-test indicates the residual normality at $\alpha = 0.01$, indicating that with a finite fold number of the cross-validation, $\tilde{\beta}$s in Equation (5) aggregated from the FedAvg of Equation (8) is the unbiased estimator of β. The benefits of FedAvg implementation and normality guarantee are superior to the bootstrap sampling Xing *et al.* (2019)'s RF received. Their bootstrap samplings on non-normal distributions require a number of samples $n \to \infty$ for an unbiased β estimator.

Case 5's 20% hold-out evaluation provided the highest accuracy among all other evaluations. The 5-fold cross-validation, composed of 20% hold-out for five times, provided the highest accuracy of the k-fold cross-validation. The 5-fold cross-validation's RMSE, MAE, AAE, ABE, and R^2 were 1.3831, 1.0204, 5.5523%, –0.0547%, and 0.8166, respectively, which equivalently improved a lower RMSE of 0.0668, lower AAE of 0.2565%, and a higher R^2 of 0.0181 from the base case. The case 5's $\tilde{\beta}$s were benefited by the FedAvg to be the unbiased estimator with the normality guaranteed at a KS-test's p-value of 0.0335, while the 20% hold-out did not. In this study, we selected a 3-cluster regime-lasso model generated with the *property* set on 5-fold cross-validation to compare the model's accuracy with that of published models, as presented in the following sections. Table 6 and Fig. 5 exhibit $\tilde{\beta}$s of this model; the FedAvg aggregated them based on 5-fold cross-validation. These coefficients indicate the features' sensitivities to the response HHV. The regime-lasso's interpretable characteristics of sensitivity analyses lead to a better understanding of the regimes' thermal properties.

The HHV in the regimes was distinctly reacted to {FC, VM, ASH}. ASH decreased the HHV in the 1st and 2nd regimes but had a neutral effect in the 3rd regime. A percentage increase in ASH decreased the HHV by 18.1 kJ/kg and 67.9 kJ/kg in the 1st and 2nd regimes, but that in the 3rd did not change. The negative effect of ASH on the HHV partially supports the findings of Soponpongpipat *et al.* (2015) and Qian *et al.* (2020), because ASH can have a neutral or even positive effect on the HHV. The ASH–FC and ASH–VM interactions can decrease, neutralize, or increase the HHV depending on the regime characteristics. An increase of 1% in ASH–FC interactions leads to a 1.8 kJ/kg decline in the 1st regime's HHV, a 0.1 kJ/kg change in the 2nd regime's HHV, and a 2.8 kJ/kg increase in the 3rd regime's HHV. The ASH–VM interactions had similar effects. A 1.9 kJ/kg drop in the 1st regime's HHV, zero change in the 2nd regime's HHV, and 9.9 kJ/kg increase in the 3rd regime's HHV from a 1% change in ASH–VM interactions. FC exhibited the most positive impact on HHV except in the 1st regime. An increase of 1% in FC increased 99.8 kJ/kg and 295.6 kJ/kg HHV in the 2nd and 3rd regimes, but with no HHV change in 1st regime. Thus, the positive effects of FC depend on the regime, contrasting with Samadi *et al.* (2021), who implemented a relevancy factor to report the highest positive impact of FC on HHV. The FC-square rendered the HHV to increase in all regimes; its 1% increase resulted in 1.7, 2.7, and 1.8 kJ/kg increases in the HHV in the 1st, 2nd, and 3rd regimes, respectively. This study found that VM solely had neutral effects on HHV in all regimes; however, VM-square led to a 0.2 kJ/kg increase in the 1st regime. VM had a positive impact on HHV upon interaction with FC. An increase in 1% FC–VM resulted in a 5.5 kJ/kg increased HHV in the 3rd regime.

Table 3
Performance of the lasso model training on the entire dataset. It indicates the accuracy improvement of the ratio set over the properties set. The KS-test exhibits the p-value of the Kolmogorov-Smirnov test, in which the null refers to the standardized residual's normality. The abbreviations are root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), and R-squared (R^2).

	Case 1 (*Property*)						Case 2 (*Ratio*)					
Evaluation	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test
Full - sample	1.4462	1.0602	5.7835	-0.0330	0.7995	0.0030	2.4917	1.6131	9.0670	0.2685	0.4048	0.0000
10% hold-out	1.4470	1.0631	5.8058	0.0972	0.7993	0.0050	2.6700	1.7190	9.3962	0.0166	0.3165	0.0000
15% hold-out	1.4474	1.0591	5.7735	-0.1545	0.7992	0.0018	2.4955	1.6222	8.9948	0.0630	0.4030	0.0000
20% hold-out	1.4474	1.0573	5.7716	-0.1107	0.7991	0.0014	2.6479	1.6566	8.9648	-0.2068	0.3278	0.0000
25% hold-out	1.4466	1.0650	5.8151	0.0548	0.7994	0.0086	2.5097	1.6076	9.2362	-0.2699	0.3962	0.0000
30% hold-out	1.4497	1.0720	5.8543	0.1896	0.7985	0.0226	2.6839	1.6876	9.1228	-0.2077	0.3094	0.0000
50% hold-out	1.4672	1.0771	5.8443	0.3166	0.7936	0.0055	2.6639	1.6826	8.9213	-0.6554	0.3197	0.0000
3-fold cross-validation	1.4487	1.0629	5.8051	-0.0509	0.7988	0.0064	2.5717	1.6466	9.0851	0.0345	0.3659	0.0000
5-fold cross-validation	1.4466	1.0606	5.7879	-0.0427	0.7994	0.0046	2.5547	1.6437	8.8736	-0.1834	0.3743	0.0000
10-fold cross-validation	1.4520	1.0699	5.8470	-0.1005	0.7979	0.0117	2.5321	1.6257	8.8734	-0.0676	0.3853	0.0000
mean	1.4499	1.0647	5.8088	0.0166	0.7985	0.0071	2.5821	1.6505	9.0535	-0.1208	0.3603	0.0000

Table 4
Performance of the 2-cluster regime-lasso generated from the article dataset. The KS-test exhibits the p-value of the Kolmogorov-Smirnov test, in which the null refers to the standardized residual's normality. The abbreviations are root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), and R-squared (R^2).

	Case 3 (*Property*)						Case 4 (*Ratio*)					
Evaluation	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test
Full - sample	1.3959	1.0308	5.6717	-0.0194	0.8132	0.0226	1.7129	1.2225	6.5661	0.0788	0.7187	0.0001
10% hold-out	1.4505	1.0743	5.9116	-0.2886	0.7983	0.0425	1.7436	1.2516	6.8000	-0.1058	0.7086	0.0001
15% hold-out	1.4230	1.0486	5.7510	0.0300	0.8059	0.0203	1.7557	1.2420	6.6808	0.3204	0.7045	0.0003
20% hold-out	1.4111	1.0385	5.6980	0.0284	0.8091	0.0223	1.7175	1.2236	6.5859	0.1455	0.7172	0.0004
25% hold-out	1.4160	1.0518	5.8083	0.0769	0.8078	0.0501	1.7884	1.2529	6.6706	0.2386	0.6934	0.0001
30% hold-out	1.4140	1.0464	5.7749	0.0882	0.8083	0.0471	1.7381	1.2440	6.9074	0.1835	0.7104	0.0013
50% hold-out	1.4298	1.0390	5.7370	0.2583	0.8040	0.0084	1.7611	1.2360	6.7541	0.3037	0.7026	0.0000
3-fold cross-validation	1.4078	1.0382	5.7106	-0.0608	0.8100	0.0177	1.8034	1.2523	6.6512	-0.0186	0.6882	0.0000
5-fold cross-validation	1.4091	1.0420	5.7308	-0.0830	0.8097	0.0253	1.7330	1.2285	6.5700	0.0157	0.7121	0.0005
10-fold cross-validation	1.4104	1.0451	5.7493	-0.0798	0.8093	0.0325	1.7415	1.2326	6.5861	0.0058	0.7092	0.0001
mean	1.4168	1.0455	5.7543	-0.0050	0.8076	0.0289	1.7495	1.2386	6.6772	0.1168	0.7065	0.0003

Table 5
Performance of the 3-cluster regime-lasso generated from the article dataset. The KS-test exhibits the p-value of the Kolmogorov-Smirnov test, in which the null refers to the standardized residual's normality. The abbreviations are root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), and R-squared (R^2).

	Case 5 (*Property*)						Case 6 (*Ratio*)					
Evaluation	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test
Full - sample	1.3712	1.0088	5.4867	-0.0088	0.8197	0.0334	1.6814	1.1907	6.3576	0.0295	0.7290	0.0001
10% hold-out	1.4063	1.0350	5.6168	-0.0384	0.8104	0.0201	1.7396	1.2150	6.7912	0.3482	0.7099	0.0003
15% hold-out	1.3973	1.0232	5.5559	0.2037	0.8128	0.0179	1.8017	1.2707	6.8206	-0.0570	0.6888	0.0002
20% hold-out	1.3826	1.0188	5.5775	0.0643	0.8167	0.0255	1.7131	1.2086	6.4774	0.1854	0.7187	0.0001
25% hold-out	1.4096	1.0354	5.6139	-0.0671	0.8095	0.0198	1.7221	1.2208	6.5746	-0.0760	0.7157	0.0002
30% hold-out	1.3972	1.0337	5.6446	0.0851	0.8128	0.0410	1.7062	1.2103	6.4882	0.2555	0.7209	0.0006
50% hold-out	1.4010	1.0222	5.5858	0.2670	0.8118	0.0296	1.7618	1.2355	6.5732	-0.3089	0.7024	0.0001
3-fold cross-validation	1.3910	1.0224	5.5676	-0.0905	0.8145	0.0257	1.7425	1.2227	6.5104	0.0392	0.7089	0.0000
5-fold cross-validation	1.3831	1.0204	5.5523	-0.0547	0.8166	0.0335	1.7179	1.2101	6.4293	-0.0257	0.7171	0.0002
10-fold cross-validation	1.3935	1.0293	5.5846	-0.0603	0.8138	0.0348	1.7271	1.2233	6.5424	-0.0672	0.7140	0.0005
mean	1.3933	1.0249	5.5786	0.0300	0.8139	0.0281	1.7313	1.2208	6.5565	0.0323	0.7125	0.0002

Table 6
Regime-lasso's coefficient with federated averaging on the 5-fold cross-validation

	Regime 1[1]	Regime 2[2]	Regime 3[3]
Intercept	17.8847	12.5947	-12.1354
FC	0.0000	0.0998	0.2956
VM	0.0000	0.0000	0.0000
ASH	-0.0181	-0.0679	0.0000
FC*VM	0.0000	0.0001	0.0055
FC*ASH	-0.0018	0.0001	0.0028
VM*ASH	-0.0019	0.0000	0.0099
$(FC)^2$	0.0017	0.0027	0.0018
$(VM)^2$	0.0002	0.0000	0.0000
$(ASH)^2$	0.0000	-0.0004	-0.0011

[1] Sample size = 747
[2] Sample size = 14
[3] Sample size = 40

Figure 5 indicates the positive effects of FC on HHV, while ASH contributed to the adverse effects. VM tiny advocates for HHV via the interaction term on ASH. Xing *et al.* (2019) conducted the sensitivity analysis via Monte-Carlo simulations and reported ASH's maximum, VM's medium, and FC's minimum effects on HHV. Their time-consuming investigations disagreed with those of Samadi *et al.* (2021) and this study. The interpretable regime-lasso model is superior to their Monte-Carlo as the regime-lasso provides sensitivity analyses in all regimes without additional cost. The clusters' distinct thermal sensitivities also insist on the regime-lasso concept's necessity. The regime's thermal properties are a source of performance inconsistency. The HHV model should be trained through them, and not via geographical (Chun-Yang Yin, 2011) or materials (Parikh *et al.* 2005).

4.3 Performance comparison

This section discusses the regime-lasso's in-sample and literature models' out-of-sample performance. Subsection (a) allocated ten literature datasets to train the regime-lasso and compared its accuracy with that of the 12 earlier studies mentioned previously; and (b) implemented the earlier studies' models to the article set and measured the accuracy. The case of the regime-lasso's out-of-sample performance has been explored later in the next section.

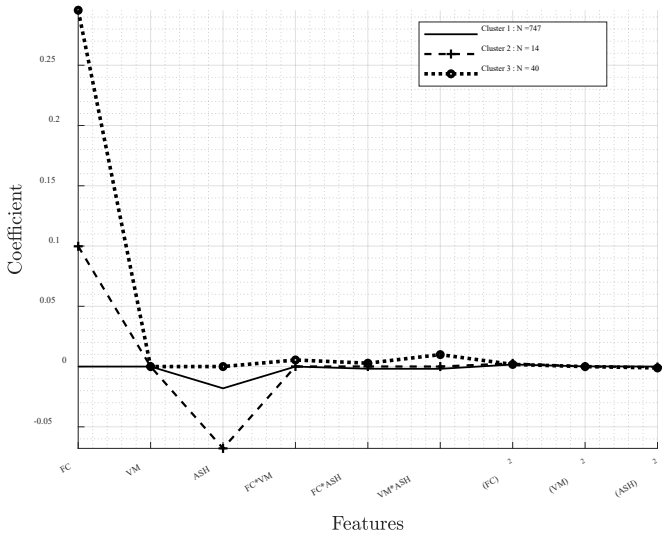

Fig. 5 The plots of the regime-lasso coefficients when implementing the article dataset as the training set. Set {FC, VM, ASH, HHV} classified three clusters based on the correlation distance metric; fixed carbon (FC), volatile matter (VM), ash (ASH), and higher heating value (HHV).

a) Literature dataset

The regime-lasso's accuracy with *property* features was compared with those of the models from literature using their original training sets. The regime-lasso was retrained on these datasets prior to the comparison. Table 7 expresses the comparison results based on RMSE, MAE, AAE, ABE, R^2, and KS-test via a complete sample substitution approach. As the literature did not report all these indicators, this study computes the missing indicators based on the literature model and its dataset.

The models in Table 7 are classified into two groups. The first group comprises the wide-range models: IV.A, IV.B, VI, VII, and VIII. The remaining are the specific material HHV models. Among the wide-range models, model IV.B provided lowest RMSE (0.6885) and MAE (0.5918), while model IV.A provided the lowest AAE (3.24%) and ABE (0.3724%). Their accuracies were inferior to that of the regime-lasso's, which generated RMSE of 0.4430, MAE of 0.2743, AAE of 1.5122%, and ABE of –0.0038% in dataset IV. The regime-lasso also outperformed the remaining wide-range models, VI and VIII. The accuracy indicators of the regime-lasso were as follows: 0.9050 RMSE, 0.6867 MAE, 3.7862% AAE, and 0.0003% ABE; whereas, the highest values of these accuracy indicators of models VI and VIII were as follows: 0.9410 RMSE, 0.8623 MAE, 3.8000% AAE, and 0.8000% ABE. The regime-lasso generated a higher accuracy than did all the wide-range models. The study excluded model VII from the comparison because its original training set was not revealed.

The specific materials of the models in Table 7 are biochar (I), poultry waste (II), torrefied biomass (III), non-woody biomass & torrefaction char (V), geographical material (IX), solid fuel (X), and lignocellulosic & carbonaceous (XI). The regime-lasso generated (i) 0.3527, 0.0242, 0.1080, 0.3064, 0.1893, 0.1059, and 0.0837 lower RMSE; (ii) 0.2639, 0.0425, 0.0467, 0.2744, 0.1611, 0.1004, and 0.0975 lower MAE; (iii) 1.4030%, 0.0833%, 0.2229%, 1.1829%, 0.8665%, 0.8320%, and 0.4989% lower AAE; and (iv) a 0.0756%, 0.3223%, 0.5871%, 0.5166%, 0.6624%, 1.8246%, and 0.0023% lower ABE in datasets I, II, III, V, IX, X, and XI, respectively. The highest R^2 of the specific range model was 0.9875, while that of regime-lasso was 0.9897. Thus, the interpretable regime-lasso, constructed as a wide-range HHV model, can replace the earlier specific range models, with superior accuracy.

b) Dataset from the current study

Table 8 presents the accuracy comparisons between the regime-lasso and literature models executed on the article dataset. The performance of the earlier models on this study's dataset represented their out-of-sample performance. The decreases in accuracy confirmed the necessity and importance of the regime-lasso. Table 8's highest accuracy model was model VI, while that in Table 7 is XI. Model XI had a 1.2616, 0.9277, 5.5020%, and 3.9210% increase in RMSE, MAE, AAE, and ABE, respectively, implying performance inconsistency. This issue was more pronounces in model I (biochar). The increases in AAE and ABE were 26.7204% and 25.5261% higher than the values computed in their original training set. These inaccuracies emerge from the improper model implementation of wide-range data. The accuracy declined even in the wide-range models. The dataset allocation increased in model IV.B's RMSE, MAE, AAE, and ABE by 0.8728, 0.5302, 2.6573%, and 0.4098%, respectively. As model IV.B could not maintain its accuracy ranking, the first definition of performance consistency is not held. The accuracy measurements with equations (9)–(13) cannot infer the model's consistent performance.

Table 7
Accuracy comparison between the regime-lasso and the literature models performed on the literature dataset. {FC, VM, ASH} is the regime-lasso's feature set. The correlation distance metric divides the literature dataset into the regimes through {FC, VM, ASH, HHV}. The KS-test exhibits the p-value of the Kolmogorov-Smirnov test, in which the null refers to the standardized residual's normality. The abbreviations are root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), R-squared (R^2), and regime-lasso (RGL).

Model	Size	K	RMSE		MAE		AAE(%)		ABE(%)		R^2		KS-test	
			Reported	RGL	Reported	RGL	Reported	RGL	Reported	RGL	Reported	RGL	Reported	RGL
I	33	2	0.7943*	0.4416	0.6008*	0.3369	2.5900	1.1870	0.0800	-0.0044	0.9820	0.9897	0.6053*	0.9324
II	37	2	1.0518*	1.0276	0.8446*	0.8021	5.9800	5.8967	-0.3500	-0.0277	0.9162	0.8748	0.9108*	0.6419
III	207	2	1.9673*	1.8593	1.3800	1.3333	6.1700	5.9471	0.6000	-0.0129	0.5847*	0.6290	0.0011*	0.0063
IV.A	12	2	0.7209*	0.4430	0.6048*	0.2743	3.4200	1.5122	0.3724*	-0.0038	0.9081	0.8375	0.6604*	0.5972
IV.B	12	2	0.6885*	0.4430	0.5918*	0.2743	3.4700	1.5122	0.4849*	-0.0038	0.9047	0.8375	0.3758*	0.5972
V	28	2	0.7453*	0.4389	0.5653*	0.2909	2.7400	1.5571	0.5200	-0.0034	0.9810	0.8052	0.4637*	0.3946
VI	382	3	0.9410	0.9050	0.8623*	0.6867	3.8000	3.7862	3.2956*	0.0018	0.9081*	0.9446	0.1477*	0.4943
VII							2.3200		-0.7600		0.9289			
VIII	250	2	1.3266*	1.0518	0.9600	0.7948	5.8800	4.5894	0.8000	0.0003	0.7918*	0.8691	0.0097*	0.2127
IX	53	2	1.3261*	1.1368	1.0283*	0.8672	5.4900	4.6235	0.6700	-0.0076	0.4217**,***	0.5516	0.3887*	0.3440
X	100	2	0.9010*	0.7951	0.7287*	0.6283	4.1933*	3.3613	1.8479*	-0.0233	0.9740*	0.9798	0.9551*	0.8772
XI	24	2	0.5586*	0.4749	0.4235*	0.3260	1.541*	1.0421	-0.0070*	-0.0047	0.9875*	0.9557	0.9873*	0.5657

* Computed in this study based on the literature's own training sets.
** Computed in this article. The literature reported a wrong 0.9976 R^2 computed by setting "Constant is zero" during data analysis in Excel. This incorrect operation made an overshoot in the sum square total because the actual mean of HHV is non-zero. Therefore, R^2 is a highly positive bias. Thus, the author recomputes the correct 0.4217 R^2 through the literature model and training set.

Table 8

Accuracy comparison of the three-regime-lasso and literature models perform on the article dataset without retraining the models. {FC, VM, ASH} is the regime-lasso's feature set. The correlation distance metric divides the literature dataset into the regimes through {FC, VM, ASH, HHV}. The KS-test exhibits the p-value of the Kolmogorov-Smirnov test, in which the null refers to the standardized residual's normality. The abbreviations are root mean square error (RMSE), mean absolute error (MAE), average absolute error (AAE), average bias error (ABE), R-squared (R^2), and regime-lasso (RGL).

Model	RMSE	MAE	AAE(%)	ABE(%)	R^2	KS-test
RGL	1.3933	1.0249	5.5786	0.0300	0.8139	0.0281
I	3.9907	3.0995	29.3104	25.6061	-0.5268	0.0003
II	3.4914	2.6303	13.3436	-7.4008	-0.1687	0.0001
III	2.4219	2.0265	9.9970	-8.5861	0.4377	0.0003
IV.A	1.6117	1.1582	6.2923	-0.4039	0.7510	0.0019
IV.B	1.5613	1.1220	6.1273	-0.0751	0.7663	0.0051
V	1.8625	1.3646	7.1895	-4.0763	0.6674	0.0025
VI	1.5901	1.1634	6.6294	3.0763	0.7576	0.0013
VII	2.6285	1.6537	22.7209	8.2613	0.3376	0.0000
VIII	4.7162	1.8411	10.5754	3.2241	-1.1324	0.0000
IX	2.2190	1.4495	7.7010	-1.0488	0.5280	0.0000
X	1.6853	1.2236	6.9828	2.4101	0.7277	0.0003
XI	1.8202	1.3512	7.0430	-3.9280	0.6824	0.0004

The regime-lasso expectedly provided the highest accuracy in this study's dataset, which was its original training set. The regime-lasso statistically generated the residuals' normal distribution in this study's dataset, with a p-value of 0.0281. The results indicated that if the regime-lasso is retrained, it can outperform all models in all datasets listed in Table 1. The superior in-sample accuracy of regime-lasso cannot infer the unique minimum unbiased estimator (UMVUE), because the lasso sacrificed some biasness for robustness and interpretability. In this study, as focused on the consistency and interpretability of the model, the UMVUE property of an HHV model has to be further researched.

4.4 The performance consistency

The comparison of performance with the models from earlier studies presented in Tables 7 and 8 indicates the performance consistency of the regime-lasso model on definition 1. In the case of the in-sample performance, the regime-lasso provided higher accuracy than those in the literature, regardless of the training data. Therefore, the regime-lasso model provides consistent performance in ranking over the different datasets. To the best of our knowledge, this is the first document demonstrated that the model maintained accuracy ranking among the experimental datasets. On the contrary, the increase in the earlier models' RMSE, MAE, AAE, and ABE in Table 8 exhibited inconsistent performance. The R^2s were –0.5268, –0.1687, and –1.1325 in models I, II, and XIII, respectively, inferring unsuitable implementation of the specific range material to the wide-range data.

Figure 6 and Table 9 report definition 2's conformability in the case of model retraining, Section 3.3's case A. The figure exhibits the error distributions of the regime-lasso model generated from 11 datasets. These residuals had a non-normal distribution with an approximate zero median. Nonparametric ANOVA played an important role in examining the distribution equivalence. A 0.9943 p-value of the Kruskal–Wallis H-test was reported in Table 9, which confirms the indifference of these distributions at $\alpha = 0.05$. The regime-lasso model statistically guaranteed the in-sample performance and supported definition 2's conformability in the experimental datasets, regardless of the material range.

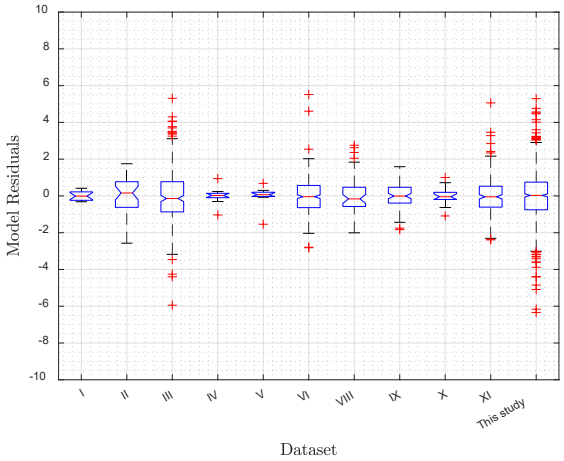

Fig. 6 Box plot presenting the error distributions equivalence of the regime-lasso in the case of model retraining

Table 9
Kruskal–Wallis H-test of the regime-lasso's error distributions in the case of model retraining

Source	SS.	d.f.	MS	Chi-square	p-value
Treatments	717,896.20	10	71,790	2.23	0.9943
Error	632,523,325.8	1,955	323,541		
Total	633,241,222.0	1,965			

The experiment in Table 7 and test in Table 9 represent cases of retraining. The regime-lasso model generated consistent accuracy in the literature datasets, supporting the retraining recommendation of Kijkarncharoensin and Innet (2022a). The normality test in Table 7 insists on the in-sample distribution invariability of the proposed models. Xing *et al.* (2019) reported the superiority of the ultimate-based ML models over the proximate-based with a 2.395 RMSE and 0.651 MAE on their wide-range training set. The literature's OLSs had a 2.00–6.17% AAE, while the MLs provided a 0.375–1.3006 RMSE and 2.64%–5.00% AAE. The results showed that: (i) the regime-lasso generated a 0.4430–0.9050 RMSE, 0.2743–0.6867 MAE, and 1.512%–4.5894% AAE on the wide-range datasets (IV, VI, and VIII); and (ii) a consistent performance at $\alpha = 0.05$. Therefore, it could be suggested that a proximate-based model can provide more accurate and consistent results than do the ultimate-based models.

Figure 7 and Table 10 exhibit error distribution if implementing Table 6 without retraining, Section 3.3's case B. The regime-lasso can maintain its accuracy on 6 out of 11 datasets at a 0.0128 p-value or $\alpha = 0.01$. The experiment statistically confirmed the regime-lasso's out-of-sample performance on the datasets I (biochar), IX (geographical materials), X (solid fuels), and wide-range datasets (VI, VIII, this study).

The model could not maintain the error distribution on the other five datasets for two reasons. First, dataset III was the torrefied biomass, while the article dataset was an ordinary material. The HHV of the former was higher than that of the latter (Nhuchhen and Afzal 2017). Second, the other datasets had a small sample size. Table 1 reports the 12–37 samples in datasets II, IV, V, and XI. Figures S1-S5 in the supplementary material present the scatter plot comparison of these five datasets with the present study's dataset. As the regime-lasso preserved its accuracy in more than 50% of the datasets involved in the experiment, we believe that the model is practical to implement in data-poor regions, such as Thailand, which lack native datasets to retrain the model.

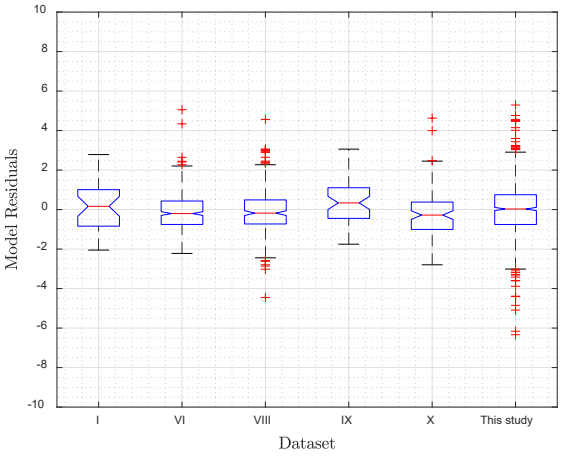

Fig. 7 Box plot presents the error distribution equivalence of regime-lasso model without retraining

Table 10
Kruskal–Wallis H-test of the regime-lasso's error distributions without retraining

Source	SS.	d.f.	MS	Chi-square	p-value
Treatments	3,167,690	5	633,538.0	14.49	0.0128
Error	350,470,000	1,613	217,278.5		
Total	353,638,000	1,618			

5. Conclusion

The regime-lasso's prediction values were statistically interpretable, unbiased, and consistent across a wide-range of data. Its in-sample accuracies were 0.4430–0.9050 RMSE, 0.2743–0.6867 MAE, and 1.512%–4.5894% AAE in wide-range literature datasets and 1.3831 RMSE, 1.0204 MAE, 5.5523% AAE, and –0.0547% ABE in our dataset. The unbiasedness was obtained from the $\boldsymbol{\beta}$ coefficients' FedAvg aggregation through the 5-fold bootstrap sampling. The error distribution from the KS-test's 0.0281 p-values and central limit theorem confirmed the model's unbiasedness. Its consistency was statistically achieved with 0.9943 and 0.0128 Kruskal–Wallis test's p-values for in and out-of-sample data points, respectively. When retraining, the model maintained the highest accuracy and preserved the error distributions in all datasets. The definitions of performance and consistency were maintained.

This article successfully proves that the data cluster's distinct thermal properties render the linear model's performance inconsistent with the wide-range data. The interpretability of the model indicated different sensitivities between HHV and {FC, AV, ASH} regimes. Of note, training the model through the regimes advocates the prediction accuracy of the entire dataset. The sensitivity analysis through the regime-lasso coefficients provides the following interpretations:

1. Increasing FC had the maximum positive impact on the 2nd and 3rd regimes, but slightly on the 1st through the FC-square.
2. Increasing ASH had a medium negative effect on the 1st and 2nd regimes, but none on the 3rd; while the interactions between ASH and others had minor positive effects on the 3rd regime's HHV.
3. Increasing VM had a neutral effect in all regimes; however, interactions with the FC and ASH had a minor positive impact on the 3rd regime.

The article presents three significant contributions. First is the consistent and unbiased proximate-based HHV model. Second is the definition of consistent performance with the capability of statistical testing. Third is the proof of the source of inconsistency that rendered linear models unsuitable for a wide range of applications. The increase in the prediction power of the HHV models promotes the practical application of biomass. Further, it encourages the use of sustainable energy, reducing greenhouse gas emissions, and mitigating the climate change crisis. The article assumed that the earlier studied were the most accurate model in its dataset. If this assumption is violated, the proof of definition 1 might be completed with other approaches. Further research should study the different factors that affect performance consistency. This study clustered and labelled the data into three regimes; however, the implications of these regimes must be investigated further.

Author Contributions: Akara Kijkarncharoensin—Conceptualization, methodology, formal analysis, writing—original draft, resources, review, editing, project administration, and validation. Supachate Innet: Advisor.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

Akkaya, E. (2016). ANFIS based prediction model for biomass heating value using proximate analysis components. *Fuel*, 180, 687–693; https://doi.org/10.1016/j.fuel.2016.04.112

Akkaya, A. V. (2013). Predicting coal heating values using proximate analysis via a neural network approach. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 35(3), 253–260; https://doi.org/10.1080/15567036.2010.509090

Boumanchar, I., Charafeddine, K., Chhiti, Y., M'hamdi Alaoui, F. E., Sahibed-dine, A., Bentiss, F., Jama, C., & Bensitel, M. (2019). Biomass higher heating value prediction from ultimate analysis using multiple regression and genetic programming. *Biomass Conversion and Biorefinery*, 9(3), 499–509; https://doi.org/10.1007/s13399-019-00386-5

Chun-Yang Yin. (2011). Prediction of higher heating values of biomass from proximate and ultimate analyses. *Fuel*, 90(3), 1128–1132; https://doi.org/10.1016/j.fuel.2010.11.031

Cordero, T., Marquez, F., Rodriguez-Mirasol, J., & Rodriguez, J. (2001). Predicting heating values of lignocellulosics and carbonaceous materials from proximate analysis. *Fuel*, 80(11), 1567–1571; https://doi.org/10.1016/S0016-2361(01)00034-5

Core Writing Team, R. K. P. and L. A. M. (2015). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *In IPCC*. https://doi.org/10.1016/S0022-0248(00)00575-3

Dashti, A., Noushabadi, A. S., Raji, M., Razmi, A., Ceylan, S., & Mohammadi, A. H. (2019). Estimation of biomass higher heating value (HHV) based on the proximate analysis: Smart modeling and correlation. *Fuel*, 257, 115931; https://doi.org/10.1016/j.fuel.2019.115931

Estiati, I., Freire, F. B., Freire, J. T., Aguado, R., & Olazar, M. (2016). Fitting performance of artificial neural networks and empirical correlations to estimate higher heating values of biomass. *Fuel*, 180, 377–383; https://doi.org/10.1016/j.fuel.2016.04.051

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22;

https://doi.org/10.18637/jss.v033.i01
Ghugare, S. B., Tiwary, S., Elangovan, V., & Tambe, S. S. (2014). Prediction of Higher Heating Value of Solid Biomass Fuels Using Artificial Intelligence Formalisms. *Bioenergy Research*, 7(2), 681–692; https://doi.org/10.1007/s12155-013-9393-5
Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2018). *Federated learning for mobile keyboard prediction*. arXiv preprint arXiv, 1811.03604, Nov 8, 2018; https://doi.org/10.48550/arXiv.1811.03604
Kijkarncharoensin, A. (2022a). A regime-switching lasso model of the biomass higher heating value on the proximate analysis. *Codeocean*. https://doi.org/https://doi.org/10.24433/CO.9945805.v2
Kijkarncharoensin, A. (2022b). *Proximate_Analysis. Mendeley Data*, V2; https://doi.org/10.17632/g36dhg826s.2
Kijkarncharoensin, A., & Innet, S. (2022a). Performance Inconsistencies in Biomass Higher Heating Value Models for Ultimate Analysis. *The Journal of King Mongkut's University of Technology North Bangkok*, 1–13 (Inpress).
Kijkarncharoensin, A., & Innet, S. (2022b). An unsupervised learning technique to classify the biomass thermal properties on the proximate analysis. Proc. of the International Conference on Electrical, *Computer and Energy Technologies (ICECET 2022)*, July, 20–22.
Liao, Y. (2017). Machine Learning in Macro-Economic Series Forecasting. *International Journal of Economics and Finance*, 9(12), 71; https://doi.org/10.5539/ijef.v9n12p71
McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. Y. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of Machine Learning Research*, 54, 10. https://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf
Mela, C. F., & Kopalle, P. K. (2002). The impact of collinearity on regression analysis: The asymmetric effect of negative and positive correlations. *Applied Economics*, 34(6), 667–677; https://doi.org/10.1080/00036840110058482
Mohammed, I. Y., Kazi, F. K., Suzana, Y., Alshareef, I. M., & Chin, S. A. (2014). Higher Heating Values (HHV) Prediction Model from Biomass Proximate Analysis Data. Conference: *International Conference & Exhibition on Clean Energy*, October; https://doi.org/10.13140/RG.2.1.1979.9525
Nhuchhen, D. R., & Abdul Salam, P. (2012). Estimation of higher heating value of biomass from proximate analysis: A new approach. *Fuel*, 99, 55–63; https://doi.org/10.1016/j.fuel.2012.04.015
Nhuchhen, D. R., & Afzal, M. T. (2017). HHV predicting correlations for torrefied biomass using proximate and ultimate analyses. *Bioengineering*, 4(1); https://doi.org/10.3390/bioengineering4010007
Parikh, J., Channiwala, S. A., & Ghosal, G. K. (2005). A correlation for calculating HHV from proximate analysis of solid fuels. *Fuel*, 84(5), 487–494; https://doi.org/10.1016/j.fuel.2004.10.010
Qian, C., Li, Q., Zhang, Z., Wang, X., Hu, J., & Cao, W. (2020). Prediction of higher heating values of biochar from proximate and ultimate analysis. *Fuel*, 265(September 2019), 116925; https://doi.org/10.1016/j.fuel.2019.116925
Qian, X., Lee, S., Soto, A. M., & Chen, G. (2018). Regression model to predict the higher heating value of poultry waste from proximate analysis. *Resources*, 7(3); https://doi.org/10.3390/resources7030039
Samadi, S. H., Ghobadian, B., & Nosrati, M. (2021). Prediction of higher heating value of biomass materials based on proximate analysis using gradient boosted regression trees method. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, 43(6), 672–681; https://doi.org/10.1080/15567036.2019.1630521
Soponpongpipat, N., Sittikul, D., & Sae-Ueng, U. (2015). Higher heating value prediction of torrefaction char produced from non-woody biomass. *Frontiers in Energy*, 9(4), 461–471; https://doi.org/10.1007/s11708-015-0377-3
Taki, M., & Rohani, A. (2022). Machine learning models for prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for waste-to-energy evaluation. *Case Studies in Thermal Engineering*, 31(January), 101823; https://doi.org/10.1016/j.csite.2022.101823
Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288.
Uzun, H., Yıldız, Z., Goldfarb, J. L., & Ceylan, S. (2017). Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *Bioresource Technology*, 234, 122–130; https://doi.org/10.1016/j.biortech.2017.03.015
Xing, J., Luo, K., Wang, H., Gao, Z., & Fan, J. (2019). A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy*, 188, 116077; https://doi.org/10.1016/j.energy.2019.116077