energies

MDPI

Article

Energy Potentials of Agricultural Biomass and the Possibility of Modelling Using RFR and SVM Models

Ivan Brandić [1], Alan Antonović [2,*], Lato Pezo [3], Božidar Matin [2], Tajana Krička [1], Vanja Jurišić [1], Karlo Špelić [1], Mislav Kontek [1], Juraj Kukuruzović [1], Mateja Grubor [1] and Ana Matin [1]

[1] Faculty of Agriculture, University of Zagreb, Svetošimunska Cesta 25, 10000 Zagreb, Croatia
[2] Faculty of Forestry and Wood Technology, University of Zagreb, Svetošimunska Cesta 23, 10000 Zagreb, Croatia
[3] Institute of General and Physical Chemistry, University of Belgrade, Studentski trg 12/V, 11000 Belgrade, Serbia
* Correspondence: alan.antonovic@gmail.com

Abstract: Agricultural biomass is one of the most important renewable energy sources. As a byproduct of corn, soybean and sunflower production, large amounts of biomass are produced that can be used as an energy source through conversion. In order to assess the quality and the possibility of the use of biomass, its composition and calorific value must be determined. The use of nonlinear models allows for an easier estimation of the energy properties of biomass concerning certain input and output parameters. In this paper, RFR (Random Forest Regression) and SVM (Support Vector Machine) models were developed to determine their capabilities in estimating the *HHV* (higher heating value) of biomass based on input parameters of ultimate analysis. The developed models showed good performance in terms of *HHV* estimation, confirmed by the coefficient of determination for the RFR (R^2 = 0.79) and SVM (R^2 = 0.93) models. The developed models have shown promising results in accurately predicting the *HHV* of biomass from various sources. The use of these algorithms for biomass energy prediction has the potential for further development.

Keywords: agricultural biomass; higher heating value; machine learning; estimation; energy potential

check for updates

Citation: Brandić, I.; Antonović, A.; Pezo, L.; Matin, B.; Krička, T.; Jurišić, V.; Špelić, K.; Kontek, M.; Kukuruzović, J.; Grubor, M.; et al. Energy Potentials of Agricultural Biomass and the Possibility of Modelling Using RFR and SVM Models. *Energies* **2023**, *16*, 690. https://doi.org/10.3390/en16020690

Academic Editor: David Herak

Received: 30 November 2022
Revised: 3 January 2023
Accepted: 4 January 2023
Published: 6 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With increasing population growth and negative climate change trends, there is a need to create sustainable systems of energy production and bioeconomy [1]. The use of green technologies, i.e., biofuels, is among the most effective ways for reducing greenhouse gasses that directly affect global warming. Biomass is a renewable energy source and has high potential in energy production [2]. Large amounts of lignocellulosic biomass in the world allow its utilization and conversion into an alternative fuel source [3]. Calorific value is the most important parameter in assessing the possibility of using biomass as a fuel [4]. Higher heating value (*HHV*) is an important factor in describing the quality of fuel and the possibility of using biomass for energy conversion. The use of biomass as an alternative fuel source is considered environmentally and economically viable and offers the possibility of replacing current fossil fuels. One of the fundamental characteristics of biomass is its chemical composition, i.e., the ultimate analysis, which includes a percentage of carbon (C), hydrogen (H), nitrogen (N), sulfur (S) and oxygen (O) [5]. Large amounts of biomass are produced as a by-product in agricultural production and most of it is unused. Biomass from agricultural products such as corn residues, straw and sunflower stalks represents an easily available energy source [6]. The authors also state that corn cobs are an economically viable and environmentally accessible source of biomass that can be used for agricultural energy production. Soybean is a legume that has economic value in terms of seed, straw and biomass yield. In the study by Krisnawati and Adie [7], the soybean was mentioned for its great prospect as a biomass energy source and is especially recognized as a biomass

source for fuel production due to its favorable energy properties [8]. The sunflower is an extremely important feedstock grown on a large area that can produce a significant amount of biomass per hectare [9]. The authors also note that biomass from sunflower cultivation is a suitable feedstock for the production of second-generation biofuels.

In order to assess the possibility of using biomass as a fuel, its composition must be determined by various laboratory analyses. The elements of organic matter in biomass are carbon (C), hydrogen (H), oxygen (O), nitrogen (N) and sulfur (S) which can be measured by the ultimate analysis [10,11]. Higher heating value (*HHV*) is a specific characteristic of biomass that can be used to evaluate the possibility of using biomass for energy through conversion (in the form of heat, fuel, etc.) [12]. In the study of Roman et al. [13], the investigation aimed to evaluate the mechanical and energetic properties of shredded pine forest residues during the briquetting process. The shredded fragments of the forest residues were compacted by the principal stresses with a determination of the energy value consumed during the briquetting process.

In the conducted research by Ibikunle et al. [14], a model was developed to predict the *HHV* of municipal waste using the input data of the ultimate analysis. The models used to estimate *HHV* were either in linear or quadratic form. The best-fitting model in the study showed good performance in terms of prediction and the coefficient of determination (R^2 = 0.97) was used as the main evaluation parameter. SVM models for estimating *HHV* are applicable to different types of biomass, thus providing a good solution to the problem of estimating *HHV* [15]. A machine learning (ML) model created to estimate *HHV* of biomass was based on the input parameters of proximate analysis data (percentage of fixed carbon, ash and volatile matter). An extreme learning machine (ELM) method proved to be very practical in estimating the *HHV*, as evidenced by the high coefficient of determination for the input parameters of fixed carbon (0.972), volatiles (0.989) and ash (0.968) [12]. Bychkov et al. [16] investigated developed models for predicting the *HHV* of plant biomass from ultimate analysis data. In the conducted study, authors used 150 models of which 8 were selected for model testing, with 3 showing good performance in estimating the *HHV* of biomass with small deviations from actual values.

The aim of this paper is to determine the possibility of the mathematical modelling of *HHV* for corn, soybean and sunflower biomass using machine learning techniques, such as the support vector machine (SVM) and random forest regression (RFR) models for regression, where all data are divided into two parts, such that one part is for training and one part for testing the model in a ratio of 70–30(%). The statistical test "goodness of fit" is used as the main evaluation method for the model performance. A comparison of the SVM and RFR models will show which model is better for predicting *HHV* biomass based on the input parameters of the ultimate analysis.

2. Materials and Methods

2.1. Data Collection

The data for creating the SVM and RFR models were obtained from the literature [17–28]. The data of the ultimate analysis and *HHV* for 51 biomass samples, including 27 samples for corn, 15 for soybean and 19 for sunflower biomass, including the biomass and biomass products, were collected for model development.

2.2. Nonlinear Modelling

After the collection dataset, the data were divided into a part for training and a part for testing the model, in the ratio of 70% and 30%, for SVM and RFR models. Statistical analysis was performed using TIBCO STATISTICA 13.3.0 software (StatSoft TIBCO Software Inc., Palo Alto, CA, USA). The analyzed data were presented in the form of mean and standard deviation. Analysis of variance (ANOVA) and Tukey's HSD (honestly significant difference) test were used to compare the samples to show the variability of the observed data.

SVM, as a supervised learning model, is based on statistical theories and can be used for clustering and regression [15]. The SVM model created is based on the input data

from the ultimate analysis, as type 1 regression models with Kelner type (Radial Basis Function—RDF) and 9 support vectors. The total number of iterations of the SVM model is 10,000.

As nonlinear models, RFRs are suitable for predictions with medium and large data sets [28]. The RFR models were also based on the input data from the final analysis with the data split into 70% for training and 30% for testing the model. The models were built using 10,000 random trees. For each internal node within the decision trees, entropy is calculated using the formula (Equation (1)) [29]:

$$E = -\sum_{i=1}^{c} p_i \times \log(p_i) \tag{1}$$

where c represents the number of unique classes and p_i prior probability of each given class. In order to solve the nonlinear problem, the Kelner function is used to map the input vectors into a multidimensional vector proctor that is used to find the hyperplane [30]. The equation used to create the SVM model is shown in the following equation (Equation (2)) [31]:

$$\gamma = \omega^T \theta(\chi) + b \tag{2}$$

where γ is target value, w is the weight vector, b is the threshold, θ is the nonlinear function of the model and χ is input vector.

2.3. Models Verification

To show the performance of the developed SVM and RFR models with respect to the input variables of the ultimate analysis, the following statistical parameters must be calculated: x^2 (reduced chi-square) from Equation (3), RMSE (root mean square error) from Equation (4), MBE (mean bias error) from Equation (5), MPE (mean percentage error) from Equation (6), and SSE (sum of squared estimate error) from Equation (7). "Goodness of fit" is calculated using the above statistical parameters to find the model with the lowest error, and they are represented by the following equations [32]:

$$x^2 = \frac{\sum_{i=1}^{N} (x_{pre,i} - x_{\exp,i})^2}{N - n} \tag{3}$$

$$RMSE = \left[\frac{1}{N} \cdot \sum_{i=1}^{N} (x_{pre,i} - x_{\exp,i})^2 \right]^{1/2} \tag{4}$$

$$MBE = \frac{1}{N} \cdot \sum_{i=1}^{N} (x_{pre,i} - x_{\exp,i}) \tag{5}$$

$$MPE = \frac{100}{N} \cdot \sum_{i=1}^{N} \left(\frac{\left| x_{pre,i} - x_{\exp,i} \right|}{x_{\exp,i}} \right) \tag{6}$$

$$SSE = \sum_{i=1}^{N} (x_{pre,i} - x_{\exp,i})^2 \tag{7}$$

3. Results

Table 1 shows the average values of the ultimate analysis and the *HHV* of the observed biomass.

Table 1 shows the mean and standard deviation of the variables of the ultimate analysis of corn, soybean and sunflower biomass. Sunflower biomass has the highest value for C (average 54.16%), H (average 6.51%), N (2.22%) and *HHV* (average 22.46 MJ/kg), and the lowest value for S (average 0.06%) and O (35.62%). Higher proportions of C and H influenced the increase in the *HHV* value [33]. Soybean biomass has the lowest average

value of C (47.37%), H (4.53%) and *HHV* (17.93 MJ/kg), while it has the highest average value of S (0.22%). The corn biomass sample has an average value of C (45.45%), H (5.29%), N (0.70%), S (0.08%), O (38.11%) and *HHV* (19.42 MJ/kg). All observed samples showed a significant difference. Thus, the variables C and S are statistically significant at a significance level of $\alpha = 0.05$, while the observed variables H, N, O and *HHV* are statistically significant at a significance level of $\alpha = 0.01$.

Table 1. Ultimate analysis of corn, soybean and sunflower biomass.

Sample	C (%)	H (%)	N (%)	S (%)	O (%)	*HHV* (MJ/kg)
Corn biomass	48.45 ± 7.3^{a}	5.29 ± 1.11^{b}	0.70 ± 0.50^{a}	0.08 ± 0.09^{a}	38.11 ± 11.22^{a}	19.42 ± 2.72^{a}
Soybean biomass	47.37 ± 3.48^{a}	4.53 ± 0.73^{a}	1.38 ± 2.32^{b}	0.22 ± 0.3^{b}	46.5 ± 6.83^{b}	17.93 ± 1.04^{a}
Sunflower biomass	54.16 ± 8.9^{b}	6.51 ± 0.9^{c}	2.22 ± 1.64^{c}	0.06 ± 0.13^{a}	35.62 ± 9.25^{a}	22.46 ± 4.41^{b}
Significance	**	*	*	**	*	*
Minimum	47.37	4.53	0.70	0.06	35.62	17.93
Maximum	54.16	6.51	2.22	0.22	46.50	22.46
Average	49.99	5.44	1.44	0.12	40.08	19.94

C—concentration of carbon; H—concentration of hydrogen; N—concentration of nitrogen; S—concentration of sulfur; O—concentration of oxygen; *HHV*—higher heating value. The means in the same column (various samples) with different lowercase superscripts are statistically different ($p < 0.05$), according to Tukey's HSD test. Statistical significance; * $p \leq 0.01$; ** $p \leq 0.05$.

Figure 1 shows the correlation diagram of the observed variables of the ultimate analysis and *HHV* of the average values of corn, soybean and sunflower biomass. The correlation of the observed values is shown in the range −1 to 1, which corresponds to the color intensity. Variables C (r = 0.98), H (r = 0.99) and N (r = 0.70) are positively correlated with the *HHV* value, while variables S (r = −0.83) and O (r = −0.88) are negatively correlated with the *HHV* value.

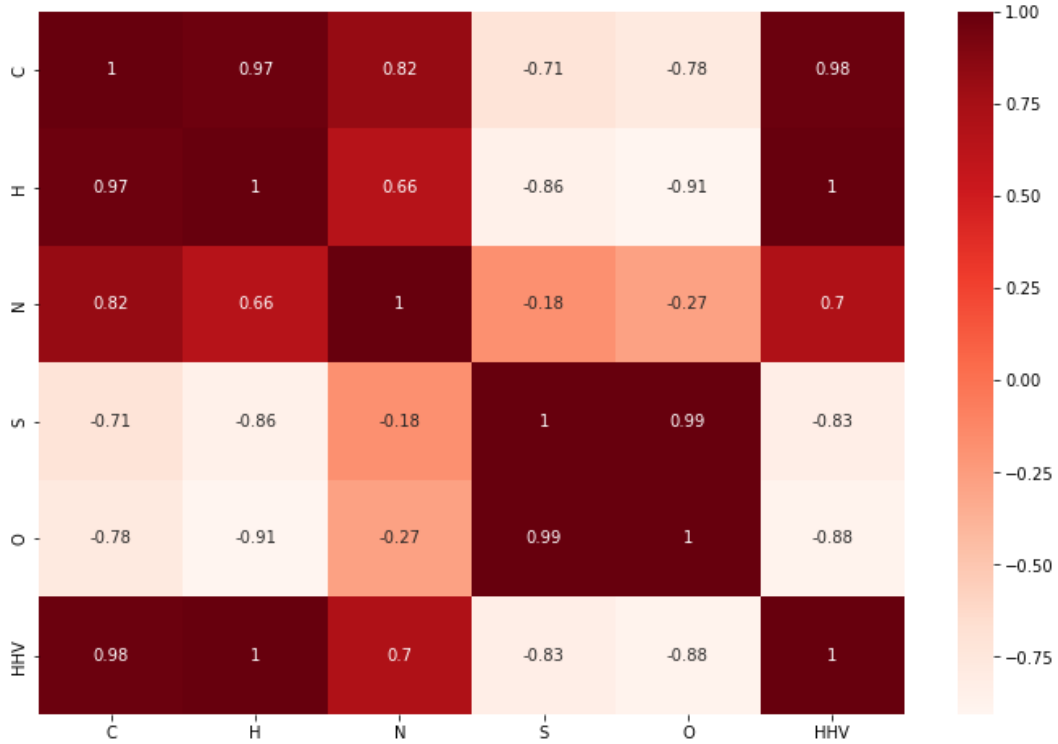


Figure 1. Correlation plot of observed variables of ultimate analysis and *HHV*.

Figure 2 shows the PCA of the observed biomass samples in relation to the variables of the ultimate analysis and the *HHV*. The PCA method is used to simplify a complex array of

data into more understandable groups for presentation by grouping all data with minimal losses into meaningful units [34,35]. The sunflower biomass group has the highest value for H, C, N and *HHV*, while the soybean biomass group has the lowest value for C, H and *HHV*. The observed corn biomass group has the lowest value for the variable N.

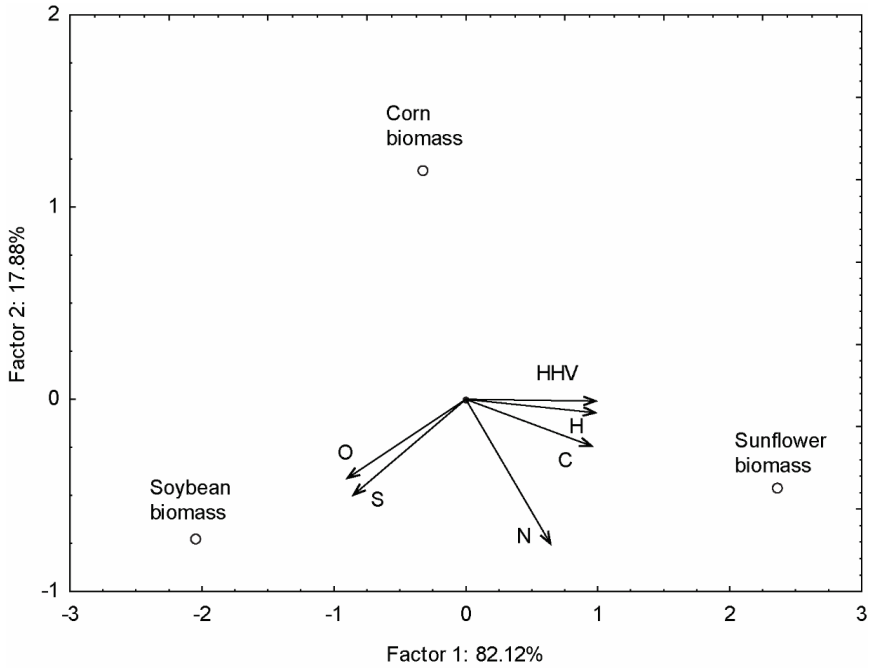


Figure 2. Principal component analysis (PCA) of observed corn, soybean and sunflower biomass.

4. Discussion

4.1. Support Vector Machine (SVM)

SVM models can be used for prediction in the form of regression due to their ability to generalize to different sample sizes and the possibility of nonlinear modeling [36]. The standard form of SVM belongs to the supervised form of learning and offers numerous advantages in terms of optimization and solution finding in nonlinear modeling [30]. Table 2 shows the values of the vector, weighting coefficients and decision constant with respect to the input variables in the SVM model.

Table 2. Vector values of the developed SVM model.

Vector No.	Weights	Support Vector					Decision Constant
		C	H	N	S	O	
1	9.00	0.22	0.63	0.07	0.07	0.73	−0.09
2	−7.34	0.20	0.68	0.07	0.12	0.70	
3	−0.21	0.95	0.00	0.21	0.02	0.00	
4	−1.01	0.35	0.49	0.00	0.32	0.96	
5	0.34	0.25	0.71	0.11	0.47	0.83	
6	−9.00	0.28	0.67	0.17	0.09	0.72	
7	6.14	0.51	0.70	0.16	0.00	0.64	
8	−1.95	0.77	0.96	0.57	0.00	0.36	
9	4.02	1.00	0.97	0.59	0.00	0.20	

Table 2 presents vector values of the developed SVM model.

4.2. Random Forest Regression (RFR)

Random Forest Regression (RFR) can easily adapt to nonlinear relationships between data and shows better predictive ability than linear regression models. RFRs are considered a reliable tool for predicting performance [29,37].

Figure 3 shows the importance of the predictor in the value 0 to 1 on the output value of *HHV* in the RFR model. In modeling, the highest predictor value is for O (1.00), followed by C (0.99), N (0.95), S (0.78) and H (0.57). With regard to the presented Figure 3, it can be concluded that the input parameters O, C and N have the greatest influence on the output value of *HHV*.

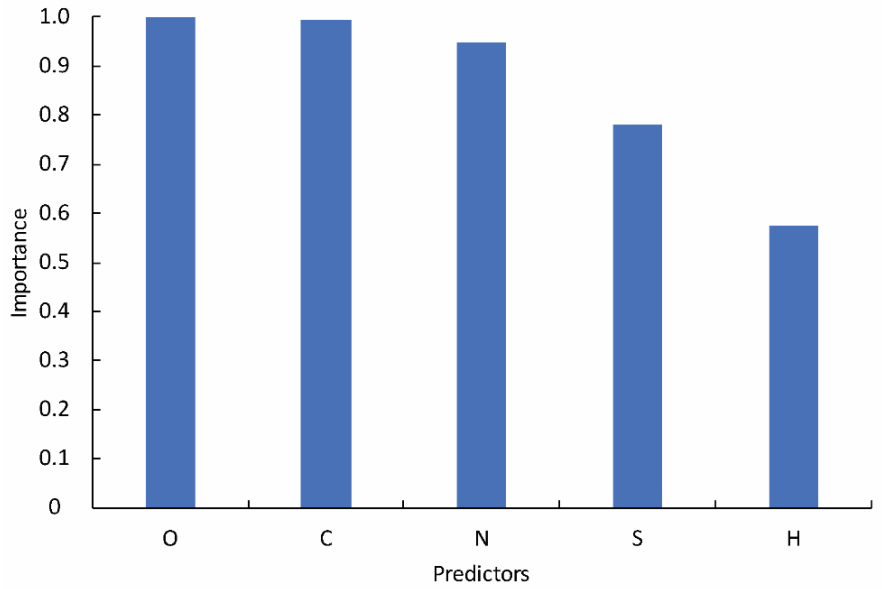


Figure 3. Importance of predictors (ultimate analysis) on *HHV* (RFR model).

4.3. Goodness of Fit

Table 3 shows the statistical analysis, "goodness of fit", which presents the ability of the developed models to predict the *HHV* values.

Table 3. "Goodness of fit" for developed SVM and RFR models.

								Residual Analysis			
Model	**x^2**	**RMSE**	**MBE**	**MPE**	**SSE**	**AARD**	**R^2**	**Skewness**	**Kurtosis**	**SD**	**Variance**
SVM model	0.82	0.90	−0.03	3.07	49.28	44.12	0.93	−3.04	14.32	0.91	0.82
RFR model	5.99	2.43	−0.01	8.32	359.53	103.30	0.79	0.94	2.08	2.45	5.99

Table 3 shows the statistical test, "Goodness of fit", which shows the performance of the developed SVM and RFR models. The calculated values of x^2 (0.82), RMSE (0.90), MBE (−0.03), SSE (49.28), AARD (44.12) and R^2 (0.93), and the residual analysis skewness (−3.04), kurtosis (14.32), SD (0.91) and Var (0.82) show the low level of error of the SVM model. The values of x^2 (5.99), RMSE (2.43), MBE (−0.01), MPE (8.32), SSE (359.53), AARD (103.30), and R^2 (0.79) were calculated for the RFR model in the table. The skewness (0.94), kurtosis (2.08), SD (2.45) and variance (5.99) parameters were determined by the residual analysis. Both developed models showed satisfactory performance in modeling the *HHV* values. Considering that, R^2 is used as the main indicator of the model's ability for estimation.

The scatterplot visualization technique was used to analyze correlations of variables on the x and y axes and to detect associations and anomalies in a multidimensional dataset [38]. Figure 4 shows the parity plot of the predicted and targeted *HHV* of the developed SVM

and RFR models for training and test data. Both models show an extensive overlap of the data, with a coefficient of determination of R^2 = 0.79 for the RFR model, while the overlap values in the SVM model have a higher coefficient of determination (R^2 = 0.93). In the study conducted by Xing et al. [38], machine learning models were built to estimate the *HHV* value of biomass based on the input parameters of ultimate and proximate analysis. The RFR model in the study shows a great fit for prediction (R^2 > 0.94), while the SVM model also shows good performance (R^2~0.90). The RFR models give good results in terms of performance. The parameters R^2, MAPE and RMSE are calculated at 0.94, 0.57 and 2.56, respectively, while for the SVM model the parameters R^2, MAPE and RMSE are 0.90, 0.76 and 3.53, respectively. The study also showed the relative importance of the input parameters of ultimate analysis in the models for C (61.6%), H (20%), O (9.6%) and N (8.8%). Considering everything, it can be concluded that the developed models are suitable for estimating the *HHV* based on the input parameters of the ultimate analysis. Using the performed statistical test "Goodness of Fit", the parameters showed a low level of error in estimating the *HHV*, while the SVM model shows a higher level of performance in modeling.

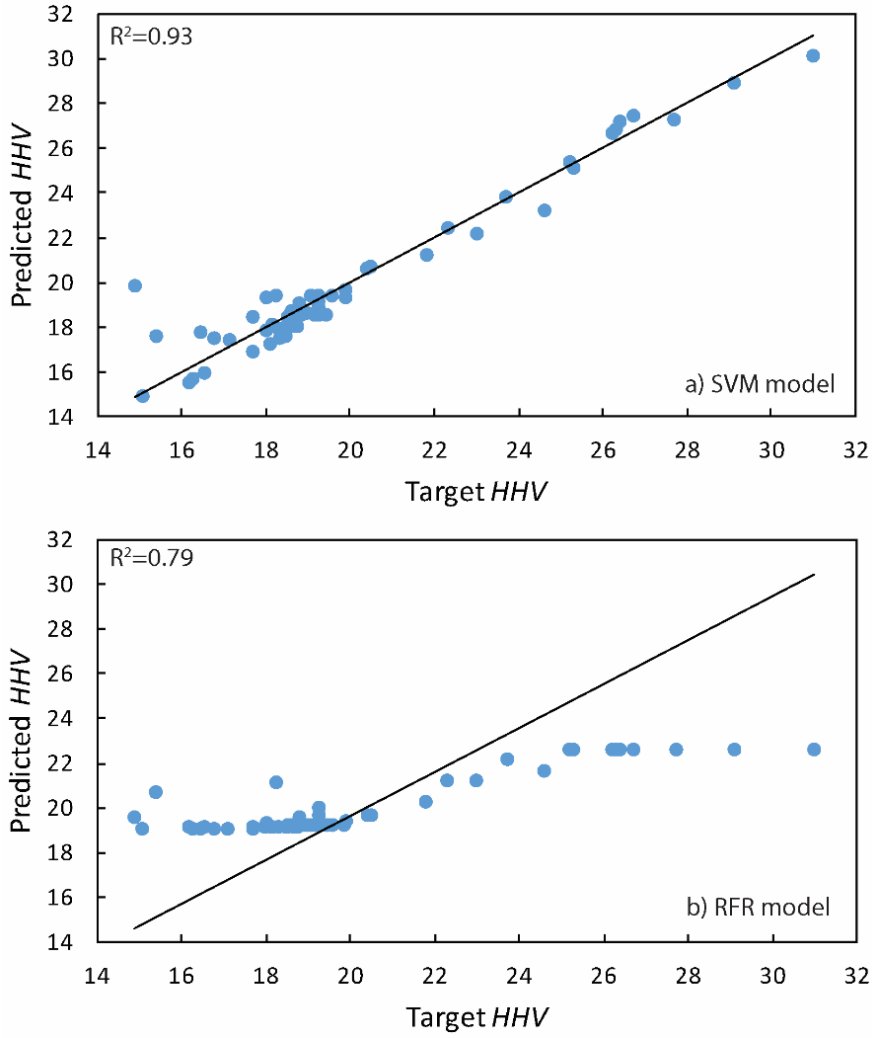


Figure 4. Parity plot of predicted *HHV* vs. Target *HHV* for developed (**a**) SVM and (**b**) RFR models.

The developed SVM and RFR models show good ability in estimating the *HHV* of corn, soybean and sunflower biomass. In developing the model, the input parameters of the ultimate analysis (percent concentration of C, H, N, S and O) were used to estimate *HHV*. In the case of the RFR model, the output value is most influenced by the variables in the following order: O (1.00), C (0.99), N (0.94), S (0.78) and H (0.57). The SVM model

with a number of five independent input parameters, as a regression model, shows the best performance with nine support vectors using RDF (Radial Basis Function) as a Kelner type. Data from 51 different biomass samples were used for the study (27 for corn biomass, 15 for soybean biomass and 19 for sunflower biomass). The SVM model showed better performance than the RFR model because the mentioned model generalizes and covers the data better for a small and medium data set, while the RFR model shows better estimation performance for a medium and large data set [39,40].

5. Conclusions

Agricultural biomass generated from corn, sunflower and soybean production has a great potential as a feedstock for energy production. Large amounts of biomass are produced as a byproduct during agricultural production and are often unused. In order to assess the possibility of using biomass as a fuel, its energy properties and composition must be determined. Nonlinear models offer the possibility to estimate the *HHV* of biomass with high accuracy. The developed nonlinear models in the form of Random Forest Regression (RFR) and Support Vector Machine (SVM) were determined as successful tools in estimating biomass *HHV*. Thus, this work shows the satisfactory ability of the SVM (R^2 = 0.93) and RFR models (R^2 = 0.79) in estimating *HHV* based on the input parameters of the ultimate analysis of observed agricultural biomass.

Author Contributions: Conceptualization, I.B., L.P., A.M. and A.A.; methodology, I.B., M.K. and B.M.; software, I.B., L.P. and K.Š.; validation, L.P., T.K., V.J. and A.M.; formal analysis, I.B. and L.P.; investigation, J.K., K.Š., M.G. and M.K.; writing—original draft preparation, I.B.; writing—review and editing, L.P., T.K., A.M. and A.A.; visualization, I.B. and L.P.; supervision, A.M.; project administration, A.A.; funding acquisition, A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund, under the Operational program competitiveness and cohesion 2014–2022, project no. KK 01.2.1.02.0286, "*Development of innovative pellets from forest and/or agricultural biomass—INOPELET*".

Data Availability Statement: Not applicable.

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

References

1. Antar, M.; Lyu, D.; Nazari, M.; Shah, A.; Zhou, X.; Smith, D.L. Biomass for a sustainable bioeconomy: An overview of world biomass production and utilization. *Renew. Sustain. Energy Rev.* **2021**, *139*, 110691. [CrossRef]
2. Biswas, B.; Pandey, N.; Bisht, Y.; Singh, R.; Kumar, J.; Bhaskar, T. Pyrolysis of agricultural biomass residues: Comparative study of corn cob, wheat straw, rice straw and rice husk. *Bioresour. Technol.* **2017**, *237*, 57–63. [CrossRef] [PubMed]
3. Cai, J.; Wu, W.; Liu, R. An overview of distributed activation energy model and its application in the pyrolysis of lignocellulosic biomass. *Renew. Sustain. Energy Rev.* **2014**, *36*, 236–246. [CrossRef]
4. Erol, M.; Haykiri-Acma, H.; Küçükbayrak, S. Calorific value estimation of biomass from their proximate analyses data. *Renew. Energy* **2010**, *35*, 170–173. [CrossRef]
5. Khunphakdee, P.; Korkerd, K.; Soanuch, C.; Chalermsinsuwan, B. Data-driven correlations of higher heating value for biomass, waste and their combination based on their elemental compositions. *Energy Rep.* **2014**, *8*, 36–42. [CrossRef]
6. Asonja, A.; Desnica, E.; Radovanović, L.Z. Energy efficiency analysis of corn cob used as a fuel. *Energy Sources Part B Econ. Plan. Policy* **2017**, *12*, 1–7. [CrossRef]
7. Krisnawati, A.; Adie, M.M. Variability of Biomass and Harvest Index from Several Soybean Genotypes as Renewable Energy Source. *Energy Procedia* **2015**, *65*, 14–21. [CrossRef]
8. Muchlish Adie, M.; Krisnawati, A. Soybean Opportunity as Source of New Energy in Indonesia. *Int. J. Renew. Energy Dev.* **2014**, *3*, 37–43. [CrossRef]
9. Antonopoulou, G.; Dimitrellos, G.; Beobide, A.S.; Vayenas, D.; Lyberatos, G. Chemical Pretreatment of Sunflower Straw Biomass: The Effect on Chemical Composition and Structural Changes. *Waste Biomass Valorization* **2015**, *6*, 733–746. [CrossRef]
10. Qian, H.; Zhu, W.; Fan, S.; Liu, C.; Lu, X.; Wang, Z.; Huang, D.; Chen, W. Prediction models for chemical exergy of biomass on dry basis from ultimate analysis using available electron concepts. *Energy* **2017**, *131*, 251–258. [CrossRef]
11. Noushabadi, A.S.; Dashti, A.; Ahmadijokani, F.; Hu, J.; Mohammadi, A.H. Estimation of higher heating values (HHVs) of biomass fuels based on ultimate analysis using machine learning techniques and improved equation. *Renew. Energy* **2021**, *179*, 550–562. [CrossRef]

12. Dai, Z.; Chen, Z.; Selmi, A.; Jermsittiparsert, K.; Denić, N.M.; Nešić, Z. Machine learning prediction of higher heating value of biomass. *Biomass-Convers. Biorefinery* **2021**, *11*, 1–9. [CrossRef]
13. Roman, K.; Barwicki, J.; Rzodkiewicz, W.; Dawidowski, M. Evaluation of Mechanical and Energetic Properties of the Forest Residues Shredded Chips during Briquetting Process. *Energies* **2021**, *14*, 3270. [CrossRef]
14. Ibikunle, R.A.; Lukman, A.F.; Titiladunayo, I.F.; Akeju, E.A.; Dahunsi, S.O. Modeling and robust prediction of high heating values of municipal solid waste based on ultimate analysis. *Energy Sources Part A Recover. Util. Environ. Eff.* **2020**, 1–18. [CrossRef]
15. Nieto, P.G.; Garcia-Gonzalo, E.; Lasheras, F.S.; Paredes–Sánchez, J.; Fernández, P.R. Forecast of the higher heating value in biomass torrefaction by means of machine learning techniques. *J. Comput. Appl. Math.* **2019**, *357*, 284–301. [CrossRef]
16. Bychkov, A.L.; Denkin, A.I.; Tikhova, V.D.; Lomovsky, O.I. Prediction of higher heating values of plant biomass from ultimate analysis data. *J. Therm. Anal. Calorim.* **2017**, *130*, 1399–1405. [CrossRef]
17. Grubor, M.; Krička, T.; Bilandžija, N.; Jurišić, V.; Voća, N.; Antonović, A.; Matin, A. Maize straw as an energy source for solid fuel production. In *53. Hrvatski i 13. Međunarodni Simpozij Agronoma*; Rozman, V., Antunović, Z., Eds.; Sveučilišta Josipa Jurja Strossmayera u Osijeku: Osijek, Croatia, 2018.
18. Matin, A.; Krička, T.; Bilandžija, N.; Jurišić, V.; Grubor, M.; Antonović, A.; Lakić, J.; Voća, N. Varieties influence on soybean straw biomass combustible and noncombustible properties. In *53. Hrvatski i 13. Međunarodni Simpozij Agronoma*; Rozman, V., Antunović, Z., Eds.; Sveučilišta Josipa Jurja Strossmayera u Osijeku: Osijek, Croatia, 2018.
19. Demirbaş, A. Fuel Characteristics of Olive Husk and Walnut, Hazelnut, Sunflower, and Almond Shells. *Energy Sources* **2002**, *24*, 215–221. [CrossRef]
20. Krička, T.; Matin, A.; Voća, N.; Pospišil, A.; Grubor, M.; Šaronja, I.; Jurišić, V. Changes in nutritional and energy properties of soybean seed and hull after roasting. *Res. Agric. Eng.* **2018**, *64*, 96–103. [CrossRef]
21. Tumuluru, J.S.; Kremer, T.; Wright, C.T.; Boardman, R.D. Proximate and Ultimate Compositional Changes in Corn Stover during Torrrefaction using Thermogravimetric Analyzer and Microwaves. In Proceedings of the 2012 ASABE Annual International Meeting, Dallas, TX, USA, 29 July–1 August 2012; pp. 1348–1365. [CrossRef]
22. Tahir, M.H.; Irfan, R.M.; Hussain, M.B.; Alhumade, H.; Al-Turki, Y.; Cheng, X.; Karim, A.; Ibrahim, M.; Rathore, H.A. Catalytic Fast Pyrolysis of Soybean Straw Biomass for Glycolaldehyde-Rich Bio-oil Production and Subsequent Extraction. *ACS Omega* **2021**, *6*, 33694–33700. [CrossRef]
23. Ahmad, S.; Zhu, X.; Wei, X.; Zhang, S. Influence of process parameters on hydrothermal modification of soybean residue: Insight into the nutrient, solid biofuel, and thermal properties of hydrochars. *J. Environ. Manag.* **2021**, *283*, 111981. [CrossRef]
24. Riva, G.; Foppa Pedretti, E.; Toscano, G.; Duca, D.; Rossini, G.; Mengarelli, C. Torrefaction of residues and by-products from sunflower chain. *J. Agric. Eng.* **2013**, *44*, 601–606. [CrossRef]
25. Liu, X.; Hua, W.; Wu, S. Characterization of Thermo-Chemical Degradation and Pyrolysis Properties for Three Kinds of Biomass Residues. *BioResources* **2016**, *11*, 8806–8819. [CrossRef]
26. Tortosa Masiá, A.A.; Buhre, B.J.P.; Gupta, R.P.; Wall, T.F. Characterising ash of biomass and waste. *Fuel Process. Technol.* **2007**, *88*, 1071–1081. [CrossRef]
27. Turzyński, T.; Kluska, J.; Ochnio, M.; Kardaś, D. Comparative Analysis of Pelletized and Unpelletized Sunflower Husks Combustion Process in a Batch-Type Reactor. *Materials* **2021**, *14*, 2484. [CrossRef] [PubMed]
28. Özyuğuran, A.; Yaman, S.; Küçükbayrak, S. Prediction of calorific value of biomass based on elemental analysis. *Int. Adv. Res. Eng. J.* **2018**, *2*, 254–260. Available online: https://www.dergipark.gov.tr (accessed on 2 September 2022).
29. Schonlau, M.; Zou, R.Y. The random forest algorithm for statistical learning. *Stata J.* **2020**, *20*, 3–29. [CrossRef]
30. Ding, S.; Zhu, Z.; Zhang, X. An overview on semi-supervised support vector machine. *Neural Comput. Appl.* **2017**, *28*, 969–978. [CrossRef]
31. Zendehboudi, A.; Baseer, M.; Saidur, R. Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* **2018**, *199*, 272–285. [CrossRef]
32. Arsenović, M.; Pezo, L.; Stanković, S.; Radojević, Z. Factor space differentiation of brick clays according to mineral content:Prediction of final brick product quality. *Appl. Clay Sci.* **2015**, *115*, 108–114. [CrossRef]
33. Bilandžija, N.; Voća, N.; Leto, J.; Jurišić, V.; Grubor, M.; Matin, A.; Krička, T. Yield and Biomass Composition of Miscanthus x Giganteus in the Mountain Area of Croatia. In *Transactions of FAMENA*; University of Zagreb: Zagreb, Croatia, 2018; Volume 42, pp. 51–60.
34. Rutledge, D.N. Comparison of Principal Components Analysis, Independent Components Analysis and Common Components Analysis. *J. Anal. Test.* **2018**, *2*, 235–248. [CrossRef]
35. Beattie, J.R.; Esmonde-White, F.W.L. Exploration of Principal Component Analysis: Deriving Principal Component Analysis Visually Using Spectra. *Appl. Spectrosc.* **2021**, *75*, 361–375. [CrossRef] [PubMed]
36. Wang, X.; Huang, F.; Cheng, Y. Computational performance optimization of support vector machine based on support vectors. *Neurocomputing* **2016**, *211*, 66–71. [CrossRef]
37. Han, S.; Kim, H. Optimal Feature Set Size in Random Forest Regression. *Appl. Sci.* **2021**, *11*, 3428. [CrossRef]
38. Nguyen, Q.V.; Lin Huang, M.; Simoff, S. Enhancing Scatter-plots with Start-plots for Visualising Multi-dimensional Data. In Proceedings of the 2020 24th International Conference Information Visualisation (IV), Melbourne, Australia, 7–11 September 2020; pp. 80–85. [CrossRef]

39. Xing, J.; Luo, K.; Wang, H.; Gao, Z.; Fan, J. A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy* **2019**, *188*, 116077. [CrossRef]
40. Kramer, K.A.; Hall, L.O.; Goldgof, D.B.; Remsen, A.; Luo, T. Fast Support Vector Machines for Continuous Data. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2009**, *39*, 989–1001. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.