Energy 245 (2022) 123295

Contents lists available at ScienceDirect

Energy

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

Random forest-based modeling for insights on phosphorus content in hydrochar produced from hydrothermal carbonization of sewage sludge

Oraléou Sangué Djandja [a], Adekunlé Akim Salami [b], Zhi-Cong Wang [a], Jia Duo [c, d, e, **], Lin-Xin Yin [a], Pei-Gao Duan [a, c, d, e, *]

[a] Shaanxi Key Laboratory of Energy Chemical Process Intensification, School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, PR China
[b] Centre d'Excellence Régional pour la Maîtrise de l'Electricité (CERME), Université de Lomé, Lomé, BP 1515, Togo
[c] Xinjiang Key Laboratory of Environmental Pollution and Bioremediation, Xinjiang Institute of Ecology and Geography, Chinese Academy of Science, Urumqi, 830011, China
[d] National Engineering Technology Research Center for Desert-Oasis Ecological Construction, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 818 South Beijing Road, Urumqi, 830011, Xinjiang, China
[e] University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Article history:
Received 8 August 2021
Received in revised form
29 December 2021
Accepted 21 January 2022
Available online 22 January 2022

Keywords:
Sewage sludge
Hydrothermal carbonization
Hydrochar
Phosphorous
Machine learning
Random forest

ABSTRACT

The hydrochar produced from hydrothermal carbonization(HTC) of sewage sludge (SS) usually has a high phosphorous (P) content, and that would result in fouling and energy efficiency reduction. Therefore, it is important to monitor the P content during the hydrochar production process. This work suggests a data-driven Random Forest-based model to predict the total P content in the hydrochar (TP-hc) from the HTC of SS. Various configurations of inputs features were examined, including the data of proximate analysis, ultimate analysis, ultimate and proximate analyses, and for each configuration, either if the total P in the SS (TP-ss) was known or not. Overall, the models including TP-ss as input have accurately predicted the TP-hc with an R^2 located in [92–95%]. Features' importance approach and partial dependence analysis pointed out that the TP-ss, ash content, reaction temperature (T), reaction time (t), and initial pH of feedwater exhibit positive effect on the TP-hc. In contrast, contribution of the volatile matter (VM) of SS was mostly negative. Dry matter loading exhibits no obvious monotonicity with TP-hc. This work could guide the production of SS-hydrochar with the desired P content, and thus avoid time and resources consuming for many trials.

1. Introduction

Phosphorous (P) is one of the most common macronutrients in commercial fertilizers that are needed for sustaining plant growth and maximum crop production. Every year, approximately 15 million tons of fertilizers containing P are used worldwide on farmland [1]. Generally, P is an essential nutrient for animals and humans as it is used in agriculture, pesticides, laundry detergents, matchstick, steel production, and the smooth functioning of the human body. In China, about 50 kg/ha/year was reported as input to the cropland in 2018 [2]. P is also the staff of life as it enables the synthesis of adenosine triphosphate (ATP), ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). However, P is a finite nutrient obtained mainly from non-renewable rocks (phosphate rock is the primary source of P) concentrated in a few countries. As modern agriculture is more dependent on the availability of P based chemical fertilizer, the distorted distribution of P resources influences the trend of agricultural and food development through the fluctuating and unstable price of phosphate rock [3]. Also, the running out of P natural resources (estimated to be depleted in 50–100 years [4,5]) makes the availability of *P*-one of the most

* Corresponding author. Shaanxi Key Laboratory of Energy Chemical Process Intensification, School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, PR China.
** Corresponding author. Xinjiang Key Laboratory of Environmental Pollution and Bioremediation, Xinjiang Institute of Ecology and Geography, Chinese Academy of Science, Urumqi, 830011, China.
E-mail address: pgduan@xjtu.edu.cn (P.-G. Duan).

https://doi.org/10.1016/j.energy.2022.123295
0360-5442/

List of acronyms and abbreviations

SS	Sewage sludge
HTC	Hydrothermal carbonization
TP-ss	Total phosphorous content in the sewage sludge
TP-hc	Total phosphorous content in the hydrochar
IP	Inorganic phosphorous
OP	Organic phosphorous
OM	Organic matter content (wt.%, db) of the sewage sludge
VM	Volatile matter content (wt.%, db) of the sewage sludge
FC	Fixed carbon content (wt.%, db) of the sewage sludge
Ash	Ash content (wt.%, db) of the sewage sludge
C	Carbon content (wt.%, db) of the sewage sludge
H	Hydrogen content (wt.%, db) of the sewage sludge
N	Nitrogen content (wt.%, db) of the sewage sludge
O	Oxygen content (wt.%, db) of the sewage sludge
t	Reaction time (min)
T	Reaction temperature (°C)
DM	Dry matter loading (wt.%)
PCC	Pearson correlation coefficient
ML	Machine learning
RF	Random Forest
RMSE	Root mean square error
R^2	Square of the correlation coefficient
SHAP	Shapley additive explanation

severe challenges that confront the world now and in the future. To overcome these challenges, growing interest in achieving more sustainable P has led to a focus on reclaiming P from biowastes such as sewage sludge (SS).

In fact, through wastewater treatment plants, a high proportion of P consumed by human activities is subsequently retained in SS. As the amount of SS is increasing over the years with population and urbanization growing, SS is gaining attention as an alternative renewable source of P. The amount of *P*-rich SS has risen considerably in recent years, and a further increase is expected [5]. The current wastewater treatment technologies provide the opportunity to retain up to 96% of the influent P in the SS [6]. The total content of P in SS also depends on the location and the type of SS. Among various organic wastes, SS exhibits the second-largest amount of P, after the bone meal (this one is produced in lower amount, compared to SS) [7]. Generally, compared to lignocellulosic biomass, SS has a high level of P [8]. SS is therefore seen as a promising renewable source of P.

The elemental P contained in SS grants it a fertilising ability, while its organic matter can also act as a suitable soil conditioner. However, it is urged that a proper pretreatment of SS is carried out before its disposal in farm lands. Besides its ability of P and other nutrients supply, SS contains other compounds that represent risks for the environment and human beings. Also, SS has quick P liberation rates that may surpass the needs of crops and/or the retention ability of soils, resulting in P losses and eutrophication of water bodies [9]. Many studies and a number of full-scale installations implemented in some countries show that the reclaiming of P from SS is technically achievable [10,11], but can face some challenges. Recently, many studies have been directed towards hydrothermal carbonization (HTC) as a benign process for both P recovery and alleviation of energy crisis.

HTC is a process with low environmental risks that not only can remove and destroy potentially hazardous substances in SS, but also has the ability to enable P recycling even in the case of energetic use of hydrochar. This process takes place in hot compressed water at temperatures usually ranging in (180–260 °C) and autogenous pressures (2–10 MPa), and accept high moisture SS without pre-drying. Compared to SS, the resulting hydrochar exhibits better storage properties, given that it is more hydrophobic and more biologically inert. Compared to SS, P content in the hydrochar increases with HTC reaction severity as precipitated phosphate salts [12]. Both high content and lower content of P in hydrochar can be preferred according to the application of hydrochar. On the one hand, the high content of P in the hydrochar is advantageous both for its use as a soil conditioner as well as for its combustion for energy production (as P could be recovered from the ash) [12,13]. The relatively lower solubility of P in hydrochar could be a new and promising strategy for addressing the SS-related P loss problem [14]. Hydrochar's microporous structure, surface functional groups, and intrinsic minerals may improve the ability of the soil to absorb and retain nutrients, potentially preventing P loss and thus improving the plant's P use efficiency [15]. On the other hand, the removal of P from hydrochar reduces the fouling of the hydrochar, therefore, increases its energy efficiency.

Although HTC had recently become a hot pot technology for SS management and P reclamation, reaction mechanisms and chemical states of P during the process still need further understanding. In fact, the P content in hydrochar is impacted by many features, including the chemical composition of SS and the HTC reaction conditions (temperature, time, pH of the process water, and dry matter concentration). The optimized point of the process is always obtained after many experiment trials, which leads to time consumption and resource wastage. The values of parameters investigated and the composition of the SS vary from one research to another, making it difficult to have a global understanding of the influence of each parameter. Therefore, in this new era where the term Industry 4.0 is attracting more attention, a Machine Learning-based model to facilitate a global understanding of the impact of each of the above-mentioned parameters and predicting the content of P in the hydrochar for optimized conditions prior to the experiment would help to make significant progress in designing HTC reactions. This study aims to suggest a data-driven Machine Learning model based on Random Forest (RF) algorithm for predicting the total content of P in hydrochar from SS. This algorithm has been successfully implemented in many cases, such as the prediction of the carbon content and the higher heating value of hydrochar [16] and energy recovery from hydrochar [17] derived from wet organic wastes, the prediction of the bio-oil yield and its hydrogen content [18] and carbon content and yield of biochar [19] from biomass pyrolysis, and the uncertainty and sensitivity analyses of co-combustion/pyrolysis processing of biowastes [20]. However, no particular attention has been given to the P migration during hydrothermal carbonization of SS. In this work, input features including proximate analysis results (fixed carbon, ash and volatile matter contents), elemental composition (C, H, N, O), and the total P content in SS (TP-ss), and HTC reaction conditions (reaction temperature, reaction time, initial pH of the feedwater, and the dry matter content in the reactor) collected from a systematic literature review, were employed to build RF based models for the prediction of the total content of P in the hydrochar (TP-hc). The features' importance was analyzed and discussed. This work could guide the computational selection of reaction conditions associated with the properties of a given SS to reach a desired amount of P in

hydrochar quickly.

2. Methodology

2.1. Data collection and pre-processing

The dataset was built from a systematic review of the literature. Published papers were extensively researched from databases (Google Scholar, Web of Sciences, ScienceDirect and Scopus) using keywords sewage sludge or wastewater combined with hydrochar, hydrothermal carbonization, hydrothermal treatment, hydrothermal conversion, hydrothermal decomposition, subcritical water conversion or thermal hydrolysis. The target papers were downloaded from publishers' websites. Only experiments conducted in subcritical conditions were included in the dataset. A total of 23 papers reporting 26 various sewage sludge with 185 data points for hydrochar was retained for the characterization of P contents in sewage sludge and hydrochar. In this dataset, 109 data points that provide complete information on the proximate analysis (volatile matter (VM (wt.%)), fixed carbon (FC (wt.%)), and ash (wt.%))), elemental analysis (C(wt.%), H(wt.%), N(wt.%), O(wt.%)) and the TP-ss (mg/g)) of the SS, the HTC reaction conditions (including temperature (T (°C)), time (t (min)), initial pH of the feedwater and the dry matter content (DM (wt.%))), and the TP-hc (mg/g)), were considered for designing the model. The details of the whole dataset are provided in the supporting information S1. The DM represent the solid concentration relative to the total feed in the reactor. The properties of the SS and the hydrochar are on a dry weight basis.

2.2. Predictive modelling and evaluation

2.2.1. Random forest approach

The RF method can be used for problems involving classification or regression analysis. The RF for regression is a bagging ensemble learning method that builds a multitude of regression trees and then aggregates them to output a final prediction [21,22]. It is an improvement of the decision tree method [23]. During training, sub-samples are randomly gathered from the predictors' training samples through bootstrap sampling, and the forest grows several trees to match the gathered sub-samples [24]. This reduces the risk that the same strong predictor variables are selected when a split is to be performed, thus preventing the regression trees from being too correlated [25]. The final prediction is obtained from the average of the outputs of all individual decision trees, with minimal variance and better generalized predictive ability [22]. The basics of these methods are depicted in Fig. 1. The RF algorithm is described in detail in Ref. [23].

In this work, the RF algorithm was implemented using the scikit-learn library (https://scikit-learn.org/stable/#), which is based on Anaconda (https://www.anaconda.com/) and using python programming language (Python 3.8.5). In the RF model, the number of estimators (number of trees in the forest) and the Max_depth (the maximum depth of the tree) are the main parameters [24]. In practice, it is recommended to set a large number of decision trees, allowing the convergence of the prediction error to a stable minimum [25]. Therefore, the number of decision trees was set in the range of 50−500, and the Max-depth was set in 2−30. The other parameters were set to default values (as presented in the supporting information S2). The optimization was conducted using two methods, including, on one hand, the grid searching and 5-fold cross-validation method, and on the other hand, the combination of the values of these hyperparameters two-by-two.

Six (06) various configurations of the input features were assessed. These configurations include the data of proximate analysis alone as inputs, ultimate analysis alone, both ultimate and proximate analyses, and for each of these three configurations, whether the TP-ss is known or not. The dataset was randomly split into two subsets for each running, including 70% used for training and the remaining 30% used for the model validation. For each configuration and a given couple (number of trees, Max-depth), the model was run 100 times to ensure the convergence of the prediction error. The accuracy of the model was evaluated using the square of the correlation coefficient (R^2), and the Root Means Square Error (RMSE).

2.2.2. Partial importance and dependence

Once the accuracy of the model is confirmed, partial importance was conducted to measure each predictor variable's effect on the response variable (TP-hc). For this purpose, an extension of the Shapely values from game theory called SHAP (Shapley Additive exPlanations) was used. This method with the scikit-learn toolbox is a helpful tool to have insight into the results of a Machine Learning prediction model [22]. According to this method, each input variable value is a player in a game where the prediction is the profit. Then the Shapley values tell us how to distribute the profit among the predictors proportionally. The partial dependence plots were also plotted from the prediction model (these plots are not the scatter plots).

3. Results and discussion

3.1. Characterization of P in sewage sludge

The TP-ss is subdivided into inorganic P (IP) and organic P (OP). The characterization of TP, IP, and OP fractions of SS and hydrochar are presented in Fig. 2a. As depicted in Fig. 2a, IP is the main fraction of the TP-ss. This claims that the P in SS is mainly chemically complexed with inorganic metals contained in the ash of the SS. These metals mainly include Al, Ca, Mg, Fe and Mn. The OP content is still very lower. This is supported when observing from the violin plot of Fig. 2b that at high organic matter (OM) content in SS, lower TP-ss is observed, while the inverse is observed for the ash content. From Fig. 2a, one can also notice that IP is also the major fraction of the total P in hydrochar (TP-hc). This supports that the interactions between inorganic metals and TP-ss are of great importance in controlling P migration to the hydrochar under various HTC reaction conditions.

3.2. Modelling dataset description

To elaborate the model, 109 datapoints providing complete information on the proximate analysis, elemental analysis and the TP-ss contents of the SS, the HTC reaction conditions, and the TP-hc are considered. The choice of these features is based on the fact that they have been found to impact the TP-hc. The statistical indicators of these factors are presented in Table 1, and their distribution is depicted by violin and swarm plots in Fig. 2c and d. As shown in Table 1, for most of the input features, the mean, the median and the mode are very close. This reveals that these data follow trends of symmetrical distribution and can therefore be fitted by a Gaussian function. The TP-ss, TP-hc and FC present skewness properties. For e.g., for FC, the mean is greater than the median, and the median is greater than the mode, pointing out a positive skewness of its distribution.

The correlation matrix between variables is depicted in Fig. 3. A higher Pearson Correlation Coefficients (PCC) of 0.88 is observed between TP-ss and TP-hc. This claims the great dependence of TP-hc on TP-ss. However, the squared PCC (0.79) shows that this relationship is insufficient for estimating the TP-hc from TP-ss. VM

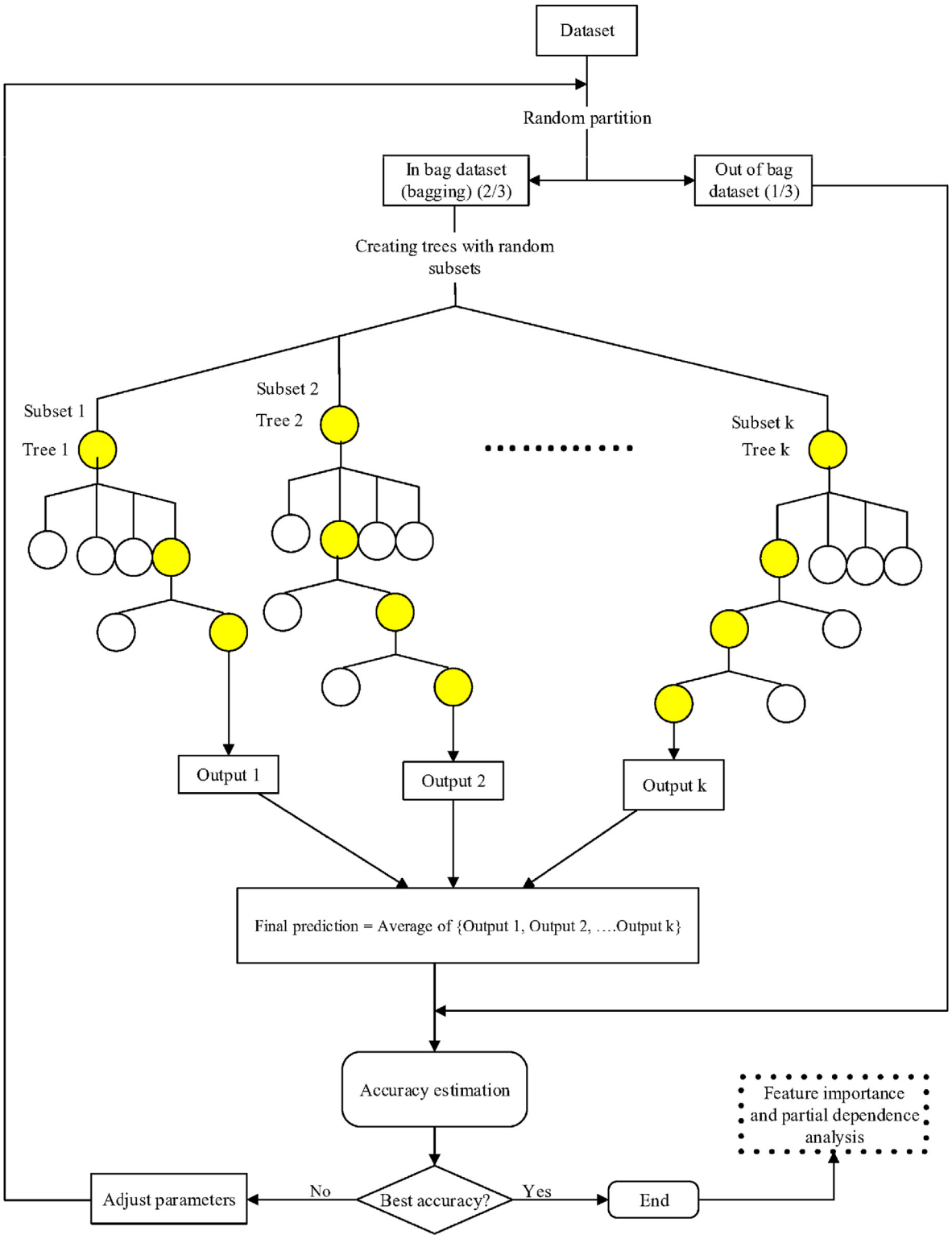

Fig. 1. Random Forest method implementation.

exhibits a significant relationship with both FC and Ash contents. This is attributed to the fact that the data are on a dry basis, and VM + FC + Ash = 100 for each observation. Similarly, Ash, C, H, N, O exhibit significant relationships between each other due to the relationship ash + C + H + N + O = 100.

3.3. Model development and accuracy

For the model development, the optimization was conducted using two methods. First, the grid searching and 5-fold cross-validation method were employed. Second, the combination of hyperparameters two-by-two was examined. Although the first

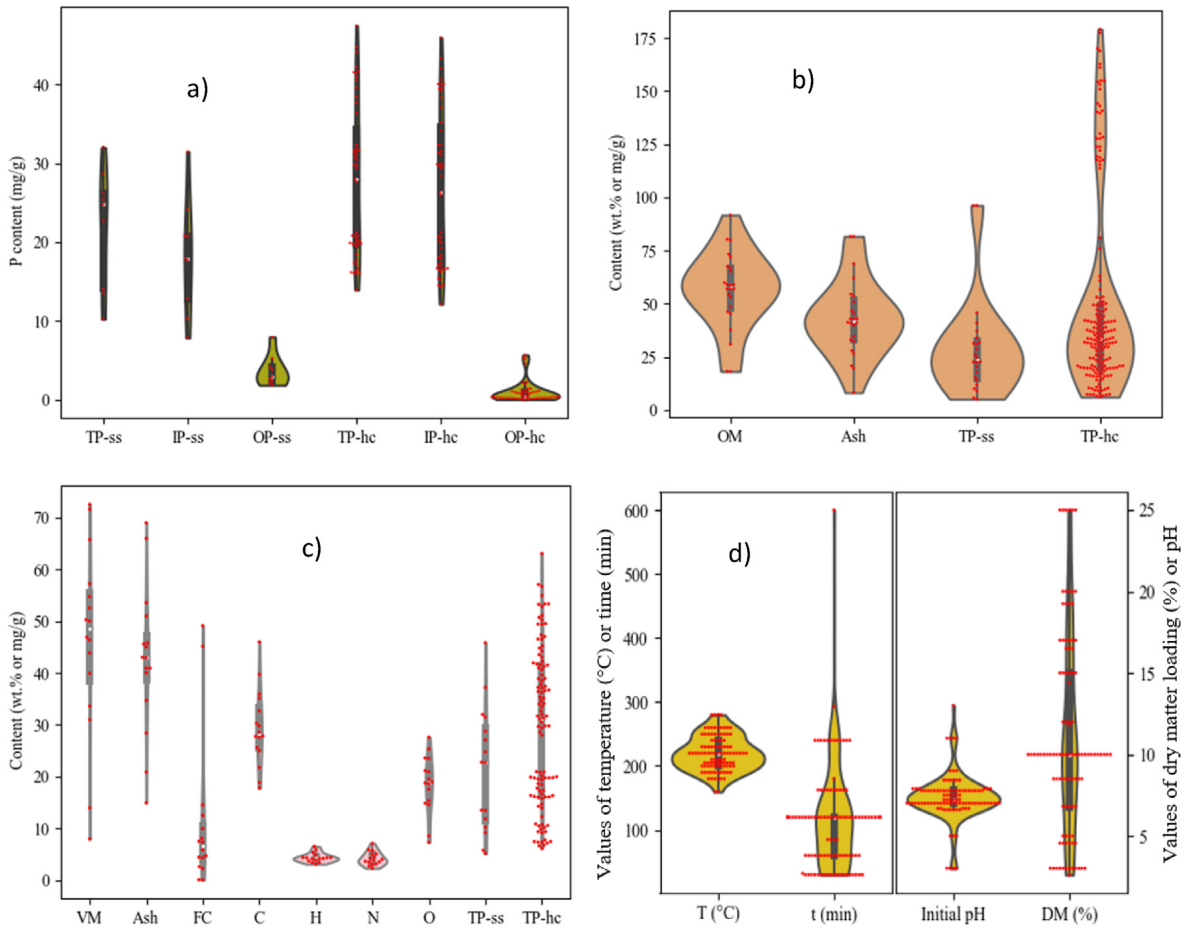

Fig. 2. Characterization of P in SS and hydrochar (a) 09 various SS and corresponding 51 datapoints for hydrochar; b) 24 various SS and corresponding 176 datapoints for hydrochar); c) and d) Distribution of parameters of the dataset considered for modelling.

Table 1
Statistical indicators of the data employed for modelling.

	Min	Max	Mean	Std	Mode	Median
VM (wt.%)	8.00	72.48	46.11	18.09	47	48.42
Ash (wt.%)	14.96	68.91	42.64	14.05	41	42.94
FC (wt.%)	0.10	49.10	11.20	14.60	5	6.655
C (wt.%)	17.83	45.96	29.40	7.36	28	28.215
H (wt.%)	3.16	6.57	4.43	0.91	5	4.3
N (wt.%)	2.34	7.15	4.24	1.29	4	3.935
O (wt.%)	7.36	27.61	18.54	5.52	19	18.89
TP-ss (mg/g)	5.22	45.80	21.38	11.98	6	22.738
T (°C)	160.00	280.00	221.93	29.23	220	220
t (min)	30.00	600.00	115.14	80.02	120	120
Initial pH	3.00	13.00	7.37	1.59	7	7.01
DM (%)	2.61	25.00	10.89	5.90	10	10
TP-hc (mg/g)	6.20	63.00	29.44	15.05	20	31.1

method provided a lower number of trees and max-depth for most of the models, the variation of the error increased after many trials by using the selected hyperparameters. Therefore, the second method was employed, and for each couple of (number of trees, Max-depth), each model was run 100 times, and the minimum and maximum of the R^2 and RMSE were reported (see the supporting information S1). The best couple of hyperparameters and the performances for each configuration, based on both training and validation, are reported in Table 2. The results revealed the remarkable ability of the Random Forest algorithm in predicting TP-hc. Although all models show great performances for training, validation performance becomes worse when the TP-ss is not included as input feature. This points out once again the main contribution of TP-ss for TP-hc. The models including TP-ss as input (model 1, model 3 and model 5 from Table 2) can predict the TP-hc with $R^2 > 0.92$, with model 5 providing the best prediction (the R^2 and RMSE were primarily stable over 100 runs). When the TP-ss is not included, the prediction performances decreased in the order of model 6 (knowing both elemental and proximate analysis), model 2 (knowing only proximate analysis) and model 4 (knowing only elemental analysis). The superiority of model 2 over model 4 may be attributed to the ash content. The correlation plots between the predicted and experimental TP-hc for the models are depicted in Fig. 4.

3.4. Features' importance

To analyze the importance of each input feature on the output (TP-hc), SHAP method was used. Only the superior models obtained when TP-ss is either included or not (model 5 and model 6) were considered. The impacts of each input feature on the TP-hc are depicted in Fig. 5. For Fig. 5a and b, the horizontal position indicates the degree and the direction of impact (lower or higher, negative or positive) of a given predictor variable. The red color indicates a high value of the variable, while the blue color indicates a low value. A positive SHAP value indicates that the predictor positively influences the output. Fig. 5c and d quantify the impacts of each input. One can notice that when TP-ss is considered, the

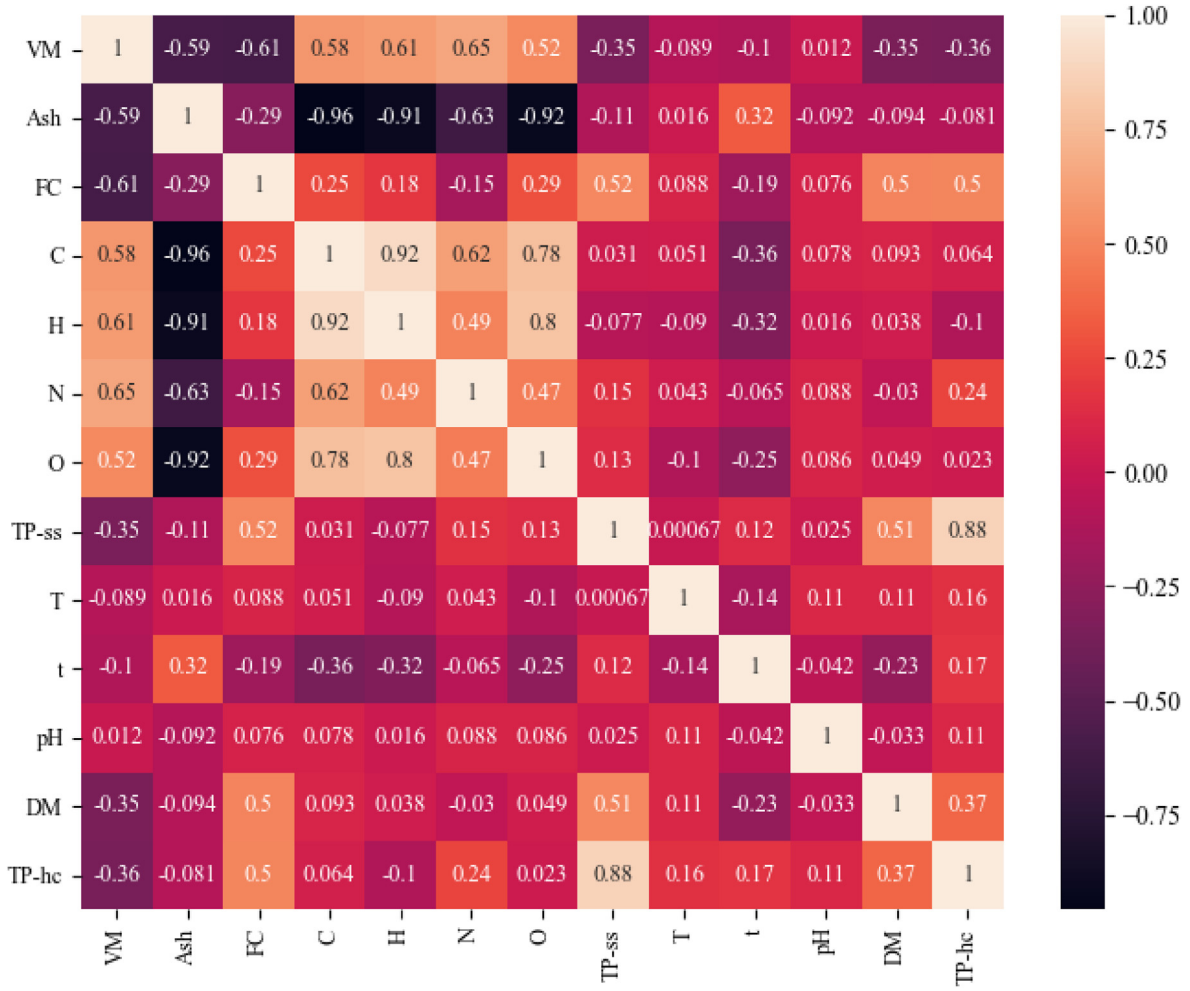

Fig. 3. Pearson correlation coefficients (PCCs) of any two variables.

Table 2
Optimum parameters and corresponding performance obtained from hyperparameters' turning[a].

Input features	Estimators	Max_depth	Training				Validation			
			R^2 Min	R^2 Max	RMSE Min	RMSE Max	R^2 Min	R^2 Max	RMSE Min	RMSE Max
Proximate with TP-ss (model 1)	100.00	10.00	95.54	96.60	2.77	3.17	92.11	94.73	3.44	4.44
Proximate without TP-ss (model 2)	500.00	20.00	94.34	95.19	3.30	3.57	80.36	83.09	6.12	6.60
Ultimate with TP-ss (model 3)	100.00	10.00	96.17	96.86	2.66	2.94	92.37	94.54	3.48	4.11
Ultimate without TP-ss (model 4)	100.00	15.00	95.26	96.20	2.93	3.27	78.52	82.99	6.14	6.90
All data with TP-ss (model 5)	300.00	5.00	94.88	95.33	3.25	3.40	93.22	94.46	3.50	3.88
All data without TP-ss (model 6)	200.00	25.00	94.99	96.10	2.97	3.36	80.94	85.31	5.71	6.50

[a] The minimum and maximum are obtained after 100 runs of each model, R^2 is in %, RMSE in mg/g.

temperature followed by the DM are the most influential HTC reaction conditions. However, with the absence of TP-ss, the DM become more influential than the temperature. This can be explained by the fact DM has a great relationship with the amount of TP-ss loaded in the reactor. A similar explanation can be given to the improvement in the impact of ash observed when TP-ss was not considered. When TP-ss was included, the contribution of the proximate analysis was lower than that of the HTC reaction conditions, while the inverse was noticed without TP-ss as an input.

3.5. Partial dependence analysis and interpretation

To provide more insight into the features' impacts, partial dependence analysis was conducted for model 5 (the best model). Note that the partial dependence analysis consists in adjusting one predictor or two predictors for predicting the output, while the other inputs are constrained to the mean values. For the most influential features, the one-way partial dependence plots obtained from the prediction with model 5 are depicted in Fig. 6, and the

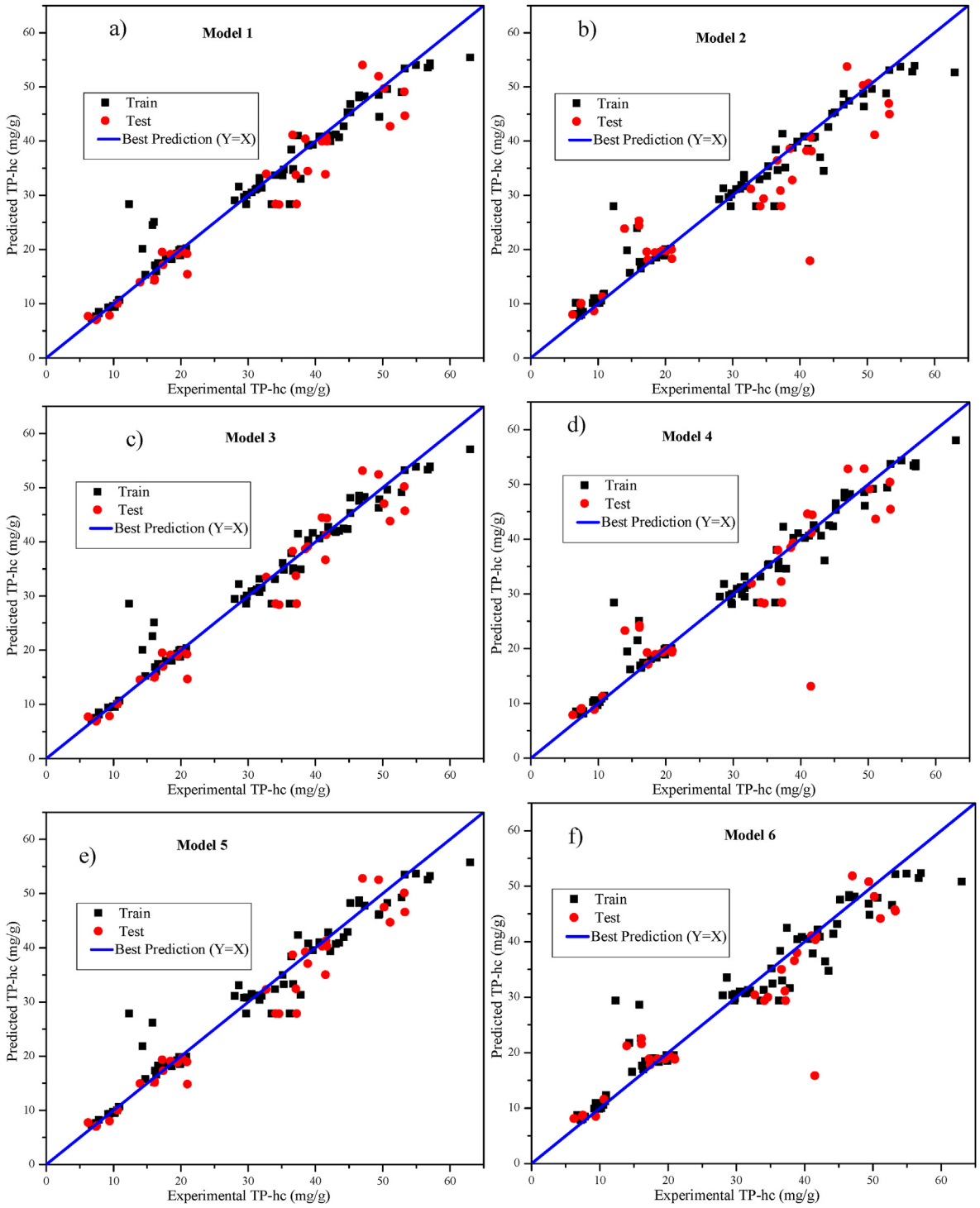

Fig. 4. Comparison of the experimental and predicted (by the various RF models) values of TP-hc: (a) Model 1; (b) Model 2; (c) Model 3; (d) Model 4; (e) Model 5; (f) Model 6.

two-ways partial dependence plots are reported in Fig. S1 of the supporting information S2.

Obviously, the TP-hc increased with increasing TP-ss (Fig. 6a). From the side of the proximate analysis components (Fig. 6b-d), TP-hc has an increasing trend with increasing ash content, while it exhibits a decreasing trend with the increasing of VM and FC. The positive contribution of ash would mainly be linked to its contents of some metals such as Ca and Mg, which significantly promote the precipitation of P as PO_4^{3-} on the hydrochar surface [26–28], with increasing temperature. The decreasing of TP-hc with increasing VM and FC is understandable, given that IP is the main fraction of TP-ss, and a high organic matter (VM and FC) lowers the ash

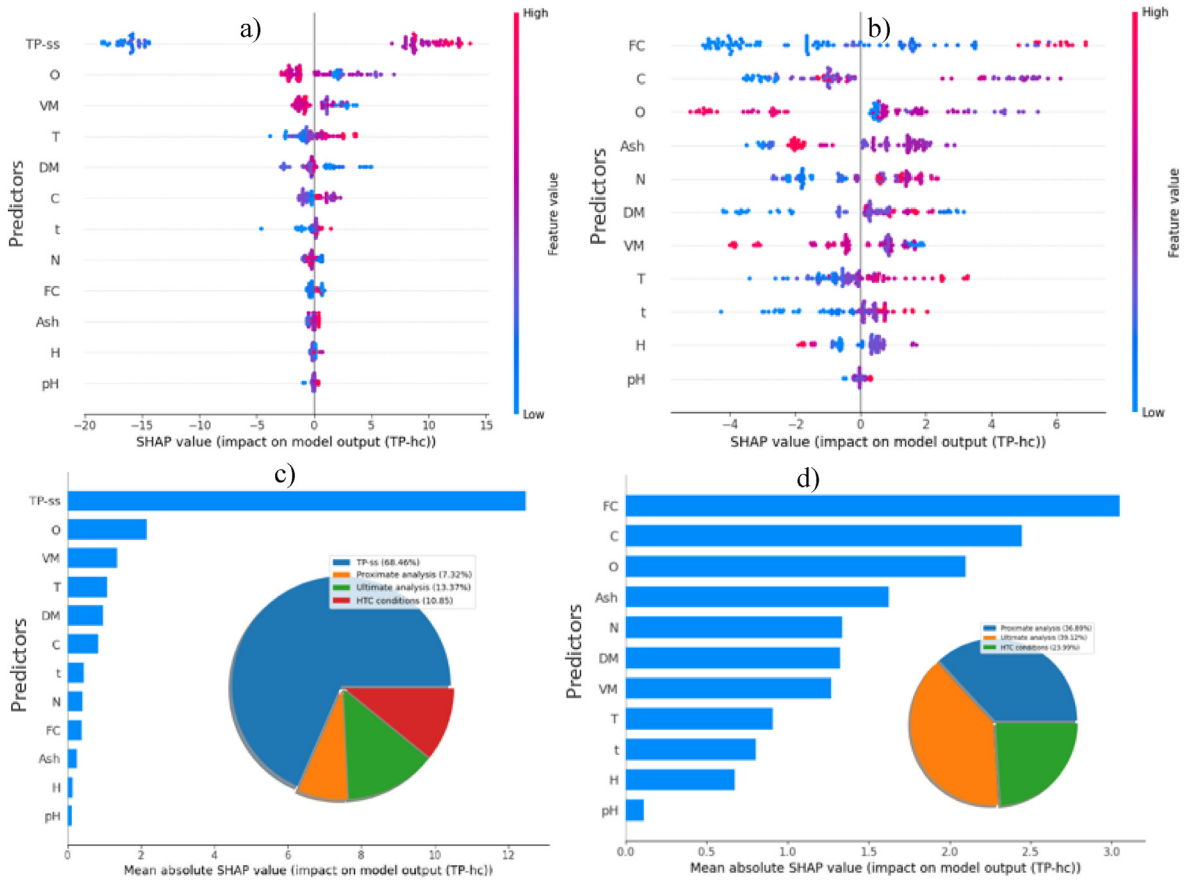

Fig. 5. Impact of predictors on the model output (TP-hc) according to the RF model, based on SHAP method: (a) SHAP value distributions for model 5; (b) SHAP value distributions for model 6; (c) Mean absolute SHAP values for model 5; (d) Mean absolute SHAP values for model 6.

content followed by a reduced TP-ss. Also, a high carbohydrates content in SS could release a high amount of acids during HTC, which may drop the pH of the process water, and therefore inhibiting the P transfer to the hydrochar. The opposite effect could be observed for high protein SS. Some positive values observed for FC can be explained by the fact that during HTC, the microbial cells and salts contained in SS could be thoroughly mixed, promoting interaction between dissolved cations and the hydrolyzed intracellular polyphosphate [29]. This let us also believe that OP would be mainly associated with FC in some SS samples. This is supported by the positive contribution observed for the C content (Fig. 6e).

On the side of HTC reaction conditions (Fig. 6f-i), the TP-hc increased with increasing temperature, reaction time and initial pH of the feedwater, while the DM does not exhibit obvious monotonicity with TP-hc. With increasing temperature, the precipitation of P with inorganic metals contained in SS is promoted, leading to an increased TP-hc. Also, the increasing temperature may promote free-radical mechanisms which enhance the decomposition of OP compounds into inorganic compounds [5,27]. However, at too high temperature, a part of P in SS could be released in the gaseous phase as phosphorus oxide [30], disfavoring TP-hc. Prolonged reaction time not only promotes the turning of the non-apatite IP into the apatite P [31], but also enhances the turning of Pyro-P into Ortho-P [27]. Thus, it favors the immobilization of P in the hydrochar. The contribution of the initial pH is attributed to the fact that many reactions related to P transformation, such as precipitation-solubilization and sorption-adsorption, are greatly affected by pH [27,32]. Decreasing pH may promote the dissolution of Ca-bound P that could decrease TP-hc, while basic condition promotes the dissolution of Al-bound P, which is subsequently transformed into Ca-bound P that becomes increasingly insoluble under alkaline conditions [33]. In a basic environment, even a lower amount of Ca can be bounded with P, leading to an increase in the apatite phosphorus [32]. The association of ortho-P with metals is also promoted in basic conditions [34]. However, this initial pH could be changed during the HTC reaction given the deamination of the proteins (releasing NH_3) [30,33,34], the solubilization of alkaline salts [28], the decomposition of carbohydrates forming organic acids [33–35], and the volatilization of acidic compounds that occur during HTC [30]. Thus, the effect of initial pH on TP-hc would be linked to the type of organic matter in SS. The DM has no obvious monotonicity with TP-hc variation. This irregular trend could be attributed to the chemical composition of SS. With high TP-ss and high content of inorganic metals such as Ca and Mg (that have high precipitation ability), TP-hc may increase with increasing DM. Interactions between features are presented through two-way partial dependence plots depicted in Fig. S1 of the supporting information S2. Overall, the results suggest that high ash and TP-ss contents, increasing temperature, higher reaction time and higher initial pH could promote the TP-hc.

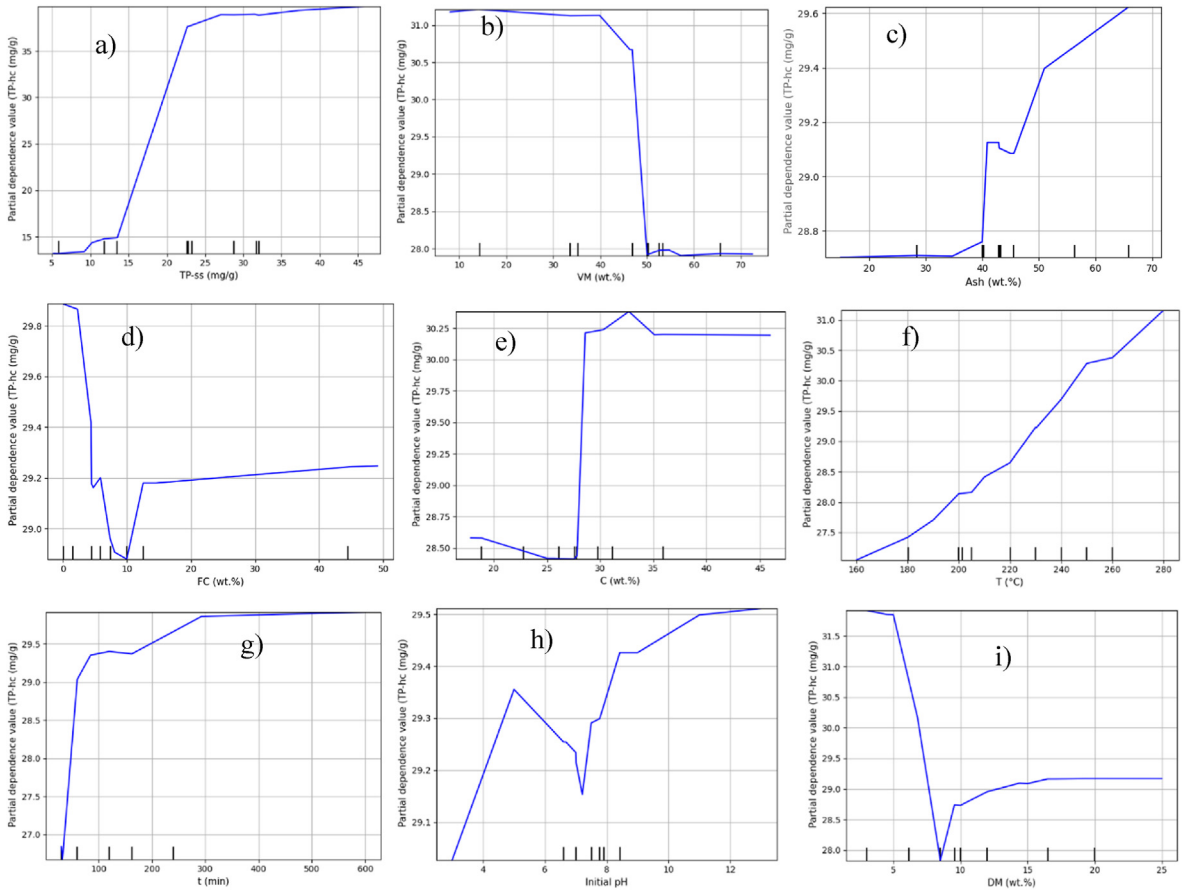

Fig. 6. One-way partial dependence plots for TP-hc (model 5): (a) Effect of TP-ss; (b) Effect of VM; (c) Effect of Ash; (d) Effect of FC; (e) Effect of C; (f) Effect of reaction temperature; (g) Effect of reaction time; (h) Effect of initial pH; (i) Effect of dry matter loading.

4. Conclusion

Various configurations of predictors, including proximate analysis, elemental composition and total P of various SS, and HTC reaction conditions, are examined for the prediction of the total P in hydrochar using Random Forest algorithm. Although all models show great performances for training, the performances of validation become worse when the TP-ss is not included as predictor. When TP-ss is included as a predictor, the TP-hc is predicted with $R^2 > 0.92$. The model's best prediction includes both elemental and proximate analysis data, TP-ss, and HTC reaction conditions. TP-ss is the most influential predictor. For the HTC reaction conditions, temperature hold the greatest positive impact, followed by the reaction time, and initial pH of the feedwater.

The results of this work revealed the great ability of the Random Forest algorithm in predicting the total content of P in hydrochar from sewage sludge. However, although the model has a great chance to always provide an accurate prediction when the feedstock is sewage sludge, for other kinds of biomass and wastes, the model may not be accurate, given that the biochemical properties of these feedstocks are different from that of sewage sludge. Hence, in the future, the dataset will be expanded and other *P*-containing wastes such as food waste and manure will be considered to develop a global model for all these wastes and compare with individual prediction for each type of waste.

Credit author statement

Oraléou Sangué Djandja and Adekunlé Akim Salami: Investigation; Conceptualization; Formal analysis; Methodology; Software; Data curation; Writing – original draft. Pei-Gao Duan and Jia Duo: Funding acquisition; Visualization; Investigation; Supervision; Validation; Writing – review & editing. Lin-Xin Lin: Conceptualization; Methodology; Software; Data curation. Zhi-Cong Wang: Conceptualization; Methodology; Software; Data curation.

Declaration of competing interest

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

Acknowledgements

We are grateful for the financial support of the National Natural Science Foundation of China (21776063; U1704127), the Scientific and Technological Innovation Team of the University of Henan Province (18IRTSTHN010).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.energy.2022.123295.

References

[1] Zhang H, Chen C, Gray EM, Boyd SE, Yang H, Zhang D. Roles of biochar in improving phosphorus availability in soils: a phosphate adsorbent and a source of available phosphorus. Geoderma 2016;276:1–6. https://doi.org/10.1016/j.geoderma.2016.04.020.

[2] Li B, Li P, Zeng XC, Yu W, Huang YF, Wang GQ, et al. Assessing the sustainability of phosphorus use in China: flow patterns from 1980 to 2015. Sci Total Environ 2020;704:135305. https://doi.org/10.1016/j.scitotenv.2019.135305.

[3] Heilmann SM, Molde JS, Timler JG, Wood BM, Mikula AL, Vozhdayev GV, et al. Phosphorus reclamation through hydrothermal carbonization of animal manures. Environ Sci Technol 2014;48:10323–9. https://doi.org/10.1021/es501872k.

[4] Xue T, Huang X. Releasing characteristics of phosphorus and other substances during thermal treatment of excess sludge. J Environ Sci 2007;19:1153–8. https://doi.org/10.1016/S1001-0742(07)60188-0.

[5] Gong W, Zou Z, Liu Y, Wang Q, Guo L. Hydrogen production and phosphorus recovery via supercritical water gasification of sewage sludge in a batch reactor. Waste Manag 2019;96:198–205. https://doi.org/10.1016/j.wasman.2019.07.023.

[6] Dagerskog L, Olsson O. Swedish sludge management at the crossroads. Stockholm Environment Institute; 2020.

[7] Cieślik B, Konieczka P. A review of phosphorus recovery methods at various steps of wastewater treatment and sewage sludge management. The concept of "no solid waste generation" and analytical methods. J Clean Prod 2017;142:1728–40. https://doi.org/10.1016/j.jclepro.2016.11.116.

[8] Zhang X, Zhang L, Li A. Hydrothermal co-carbonization of sewage sludge and pinewood sawdust for nutrient-rich hydrochar production : synergistic effects and products characterization. J Environ Manag 2017;201:52–62. https://doi.org/10.1016/j.jenvman.2017.06.018.

[9] Adhikari S, Gascó G, Méndez A, Surapaneni A, Jegatheesan V, Shah K, et al. Influence of pyrolysis parameters on phosphorus fractions of biosolids derived biochar. Sci Total Environ 2019;695:133846. https://doi.org/10.1016/j.scitotenv.2019.133846.

[10] Desmidt E, Ghyselbrecht K, Zhang Y, Pinoy L, Van Der Bruggen B, Verstraete W, et al. Global phosphorus scarcity and full-scale P-recovery techniques: a review. Crit Rev Environ Sci Technol 2015;45:336–84. https://doi.org/10.1080/10643389.2013.866531.

[11] Kwapinski W, Kolinovic I, Leahy JJ. Sewage sludge thermal treatment technologies with a focus on phosphorus recovery: a review. Waste Biomass Valor 2021;12:5837–52. https://doi.org/10.1007/s12649-020-01280-2.

[12] Wang L, Chang Y, Li A. Hydrothermal carbonization for energy-efficient processing of sewage sludge: a review. Renew Sustain Energy Rev 2019;108:423–40. https://doi.org/10.1016/j.rser.2019.04.011.

[13] Escala M, Zumbühl T, Koller C, Junge R, Krebs R. Hydrothermal carbonization as an energy-efficient alternative to established drying technologies for sewage sludge: a feasibility study on a laboratory scale. Energy Fuels 2013;27(1):454–60. https://doi.org/10.1021/ef3015266.

[14] Dai L, Tan F, Wu B, He M, Wang W, Tang X, et al. Immobilization of phosphorus in cow manure during hydrothermal carbonization. J Environ Manag 2015;157:49–53. https://doi.org/10.1016/j.jenvman.2015.04.009.

[15] Chu Q, Lyu T, Xue L, Yang L, Feng Y, Sha Z, et al. Hydrothermal carbonization of microalgae for phosphorus recycling from wastewater to crop-soil systems as slow-release fertilizers. J Clean Prod 2021;283:124627. https://doi.org/10.1016/j.jclepro.2020.124627.

[16] Li J, Zhu X, Li Y, Tong YW, Wang X. Multi-task prediction of fuel properties of hydrochar derived from wet municipal wastes with random forest, In: Applied Energy Symposium 2019: low carbon cities and urban energy systems. Paper ID: 0015.

[17] Li L, Flora JRV, Berge ND. Predictions of energy recovery from hydrochar generated from the hydrothermal carbonization of organic wastes. Renew Energy 2020;145:1883–9. https://doi.org/10.1016/j.renene.2019.07.103.

[18] Tang Q, Chen Y, Yang H, Liu M, Xiao H, Wu Z, et al. Prediction of bio-oil yield and hydrogen contents based on machine learning method: effect of biomass compositions and pyrolysis conditions. Energy Fuels 2020;34:11050–60. https://doi.org/10.1021/acs.energyfuels.0c01893.

[19] Zhu X, Li Y, Wang X. Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. Bioresour Technol 2019;288:121527. https://doi.org/10.1016/j.biortech.2019.121527.

[20] Wen S, Buyukada M, Evrendilek F, Liu J. Uncertainty and sensitivity analyses of co-combustion/pyrolysis of textile dyeing sludge and incense sticks : regression and machine- learning models. Renew Energy 2020;151:463–74. https://doi.org/10.1016/j.renene.2019.11.038.

[21] Pahlavan-Rad MR, Dahmardeh K, Hadizadeh M, Keykha G, Mohammadnia N, Gangali M, et al. Prediction of soil water infiltration using multiple linear regression and random forest in a dry flood plain, eastern Iran. Catena 2020;194:104715. https://doi.org/10.1016/j.catena.2020.104715.

[22] Li J, Pan L, Suvarna M, Tong YW, Wang X. Fuel properties of hydrochar and pyrochar : prediction and exploration with machine learning. Appl Energy 2020;269:115166. https://doi.org/10.1016/j.apenergy.2020.115166.

[23] Antoniadis A, Lambert-Lacroix S, Poggi JM. Random forests for global sensitivity analysis: a selective review. Reliab Eng Syst Saf 2021;206:107312. https://doi.org/10.1016/j.ress.2020.107312.

[24] Yang S, Zhang L, Fan J, Sun B. Experimental study on erosion behavior of fracturing pipeline involving tensile stress and erosion prediction using random forest regression. J Nat Gas Sci Eng 2021;87:103760. https://doi.org/10.1016/j.jngse.2020.103760.

[25] Fouedjio F. Exact conditioning of regression random forest for spatial prediction. Artif Intell Geosci 2020;1:11–23. https://doi.org/10.1016/j.aiig.2021.01.001.

[26] Han X, Wang F, Zhou B, Chen H, Yuan R, Liu S, et al. Phosphorus complexation of sewage sludge during thermal hydrolysis with different reaction temperature and reaction time by P K-edge XANES and 31P NMR. Sci Total Environ 2019;688:1–9. https://doi.org/10.1016/j.scitotenv.2019.06.017.

[27] Shi Y, Luo G, Rao Y, Chen H, Zhang S. Hydrothermal conversion of dewatered sewage sludge: focusing on the transformation mechanism and recovery of phosphorus. Chemosphere 2019;228:619–28. https://doi.org/10.1016/j.chemosphere.2019.04.109.

[28] Aragón-Briceño CI, Grasham O, Ross AB, Dupont V, Camargo-Valero MA. Hydrothermal carbonization of sewage digestate at wastewater treatment works : influence of solid loading on characteristics of hydrochar , process water and plant energetics. Renew Energy 2020;157:959–73. https://doi.org/10.1016/j.renene.2020.05.021.

[29] Huang R, Tang Y. Speciation dynamics of phosphorus during (hydro)thermal treatments of sewage sludge. Environ Sci Technol 2015;49:14466–74. https://doi.org/10.1021/acs.est.5b04140.

[30] Feng Y, Ma K, Yu T, Bai S, Pei D, Bai T, et al. Phosphorus transformation in hydrothermal pretreatment and steam gasification of sewage sludge. Energy Fuel 2018;32:8545–51. https://doi.org/10.1021/acs.energyfuels.8b01860.

[31] Xu Y, Yang F, Zhang L, Wang X, Sun Y, Liu Q, et al. Migration and transformation of phosphorus in municipal sludge by the hydrothermal treatment and its directional adjustment. Waste Manag 2018;81:196–201. https://doi.org/10.1016/j.wasman.2018.10.011.

[32] Wang T, Zhai Y, Zhu Y, Peng C, Wang T, Xu B, et al. Feedwater pH affects phosphorus transformation during hydrothermal carbonization of sewage sludge. Bioresour Technol 2017;245:182–7. https://doi.org/10.1016/j.biortech.2017.08.114.

[33] Ovsyannikova E, Arauzo PJ, Becker G, Kruse A. Experimental and thermodynamic studies of phosphate behavior during the hydrothermal carbonization of sewage sludge. Sci Total Environ 2019;692:147–56. https://doi.org/10.1016/j.scitotenv.2019.07.217.

[34] Zheng X, Zheng X, Jiang Z, Ying Z, Ying Z, Ye Y, et al. Migration and transformation of phosphorus during hydrothermal carbonization of sewage sludge: focusing on the role of pH and calcium additive and the transformation mechanism. ACS Sustainable Chem Eng 2020;8:7806–14. https://doi.org/10.1021/acssuschemeng.0c00031.

[35] Saetea P, Tippayawong N. Recovery of value-added products from hydrothermal carbonization of sewage sludge. 2013. https://doi.org/10.1155/2013/268947. Article ID 268947 |.