CHEMICAL ENGINEERING TRANSACTIONS
VOL. 106, 2023
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Jeng Shiun Lim, Nor Alafiza Yunus, Peck Loo Kiew, Hon Huin Chin

ISBN 979-12-81206-05-2; **ISSN** 2283-9216
DOI: 10.3303/CET23106136

Modelling Soil Respiration of Food Waste Compost Amended Soil

Siti Aisyah Mohd Dolit[a], Nur Raudhah Azman[a], Umi Aisah Asli[a], Aidee Kamal Khamis[b], Norfhairna Baharulrazi[a], Nor Alafiza Yunus[a,*]

[a]Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, MY
[b]Ibnu Sina Institute for Scientific and Industrial Research (ISI-SIR), Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, MY
alafiza@utm.my

Soil respiration plays a crucial role in the worldwide carbon cycle and displays great sensitivity to shifts in soil temperature and moisture levels. Accurate prediction of soil respiration under different ratios of food waste compost (FWC) amended soil in various variables requires a clear understanding of the processes involved. This research introduces an appropriate model aimed at estimating soil respiration. This paper employed three distinct regression models: multiple linear (Model 1), first-order polynomial (Model 2), and second-order polynomial (Model 3). These models were employed to predict soil respiration by assessing its relationship with various factors. The study examined several factors, including FWC amended ratio (A), pH (B), electrical conductivity (EC) (C), organic matter (OM) (D), carbon-to-nitrogen ratio (C/N) (E), moisture content (F), porosity (G), and microbial count (H). These factors were considered potential influencers of the CO_2 efflux response. It was observed that A,B,C and E exhibited p-values below 0.05 signifying their significance in the context of the study. Among the regression models, Model 3 demonstrated the lowest mean squared error of prediction (MSEP) and root mean square error (RMSE), 1.142 % and 0.153, respectively. The suitability of Model 3 for predicting soil respiration was attributed to its capacity to account for interaction effects among independent variables. Conversely, the results indicated that a non-linear model provide a better understanding of soil respiration under different ratios of FWC amended soil due to the smallest MSEP and RMSE, suggesting that the predictive model for CO_2 efflux aligned more with second-order behaviour.

1. Introduction

Soil respiration approximately provides ~75 × 10^{15} g of carbon projected annually into the global carbon budget, making it the second-largest contributor to gross carbon dioxide (CO_2) release into the atmosphere after oceans (Sharma et al., 2020). Slight adjustments in the rate of soil respiration can have notable effects on the yearly carbon absorption of terrestrial ecosystems because this process plays a significant role in transferring carbon between the Earth's biosphere and its atmosphere (Su et al., 2019). For instance, if the CO_2 released from the soil surpasses the amount produced by plants, it can considerably impact atmospheric CO_2 levels. Recent literature has been particularly focused on soil respiration due to its dual role in shaping net ecosystem carbon budgets and its relevance in the context of global changes (Chia et al., 2021). Accurate modeling of soil respiration is imperative for comprehending alterations in carbon storage within ecosystems and changes in carbon movement to the atmosphere caused by shifts in climate. While there's been recognition through workshops and synthesis efforts about the necessity of creating and validating models for soil respiration in conjunction with empirical findings (Ryan and Law, 2005), this approach has rarely been done. Presently, attempts have been made to experimentally separate soil respiration into its different components, yet the outcomes have displayed variations across different methodologies. The process of partitioning during experimental treatments becomes even more complicated due to distinct responses of these components to environmental shifts connected to climate change (Ryan and Law, 2005), and strong covariation among factors (Chen et al., 2011). Thus, achieving precise separation of soil respiration components remains an ongoing

Paper Received: 17 August 2023; Revised: 17 September 2023; Accepted: 10 October 2023
Please cite this article as: Dolit S.A.M., Azman N.R., Asli U.A., Khamis A.K., Baharulrazi N., Yunus N.A., 2023, Modelling Soil Respiration of Food Waste Compost Amended Soil, Chemical Engineering Transactions, 106, 811-816 DOI:10.3303/CET23106136

challenge. Quantifying or modelling variations in soil respiration under different ratios of FWC amended soil is crucial for further investigating the processes behind changes in soil respiration caused by various factors. Because soil ecosystems are complicated, most research to date have relied on empirical models (Subke et al., 2006), which are based on significant relationships between temperature (part of which includes soil moisture) and soil respiration (Phillips et al., 2016). However, the correlative model of CO_2 efflux is required to increase quantitative knowledge of soil respiration by accounting for other parameters such as C/N ratio and pH value (Hibbard et al., 2005).
This study initially involved measuring CO_2 efflux and applying Z-score limit for pre-processing to eliminate outliers. Once optimal outcomes were achieved, a Type III Sum of Squares analysis was employed to identify three highly influential factors affecting CO_2 efflux. Subsequently, both multiple linear and non-linear regression were utilized to determine the most suitable correlative model for CO_2 efflux. Finally, a comparison of outcomes among different models was conducted, alongside an assessment of the error percentage between predicted and observed CO_2 efflux values.

2. Experimental Procedure

2.1 Data collection and data pre-processing

There are 8 input variables in this study including food waste compost (FWC) amended ratio (A), pH (B), electrical conductivity (EC) (C), organic matter (OM) (D), carbon-to-nitrogen ratio (C/N) (E), moisture content (F), porosity (G), and microbial count (H). A set of CO_2 efflux data was measured from all FWC amended soil treatments by using IRGA analyzer. The data observations were collected based on previous research (Dolit et al., 2022) as shown in Table 1. Before establishing the relationship between input and output variables, the data pre-processing step was conducted to remove outliers. The default significance level is 5 % and the p-value is obtained with a Monte Carlo simulation approach.

Table 1: Effect of different ratio of FWC amended soil on the physical properties of sandy soil (P<0.05)

Soil: Compost ratio	Porosity (%)	Organic matter (%)	pH	C/N ratio	Electrical conductivity (dS/m)
100:0	91.97±0.11	6.62 ± 0.88	4.77±0.00	5.9	1.35 ± 1.06
95:5	92.34±0.11	8.59 ± 4.60	6.86±0.05	7.2	5.604 ± 0.16
85:15	94.26±0.11	15.67 ± 3.30	6.88±0.04	8.8	9.398 ± 0.006
75:25	94.42±0.19	18.67 ± 2.83	6.98±0.03	11.9	12.988 ± 0.24
65:35	95.00±0.20	21.67 ± 0.47	7.03±0.01	16.3	15.281 ± 0.02

2.2 Multiple linear regression

A multiple linear regression model referred to as Model 1 was introduced and presented as Eq(1).

$$CO_{2_{efflux}} = \sum_{i=1}^{n} \beta_i X_i \tag{1}$$

where X_i and $CO_{2_{efflux}}$ are the input and output variables and βi is the unknown predictor coefficients. To perform the regression analysis of Model 1, XLSTAT in Microsoft Excel was employed. From the Eq(1), the significance of each input variable was determined using a p-value threshold of less than 0.05. In statistical terms, a p-value below 0.05 is typically considered statistically significant, leading to the rejection of the null hypothesis (H_0). The null hypothesis assumes no relationship between the variables under study, implying that one variable does not impact the other. Three of the most significant variables were further subjected to multiple linear and non-linear regression analyses to ascertain the optimal CO_2 efflux.

2.3 Multiple non-linear regression

In this study, both first and second-order non-linear models were developed and compared. Second-order modeling is anticipated to provide more accurate predictions when the process exhibits non-linear behavior and contains a significant amount of nonlinearities, in contrast to the first-order linear model. First-order (Model 2) and second-order predictive modelling (Model 3) are shown in Eq(2) and Eq(3) respectively.

$$Y = \sum_{i=1}^{n} \beta_i X_i + \sum_{i,i=1,i\neq j}^{n} \beta_{ij} X_i X_j \tag{2}$$

$$Y = \sum_{i=1}^{n} \beta_i X_i + \sum_{i,i=1,i\neq j}^{n} \beta_{ij} X_i X_j + \sum_{i=1}^{n} \beta_{ii} X_{ii}^2 \quad (3)$$

where Y represents the output variable, Xi corresponds to the input variables (first-order terms), while Xii^2 and X_iX_j denote the second order terms derived from the input variables, βii, βij and βi are the unknown predictor coefficients. In a second-order analysis, the independent variables terms change while the dependent variable remains in first-order form. There is no y-intercept or zero-order terms in the model because all the data are mean-centered and standardized beforehand. These models were regressed and interpreted the goodness of fit statistic by using XLSTAT in Microsoft Excel. The sum square of error (SSE), mean square of error (MSE), and root mean square of error (RMSE) were assessed to evaluate the explanation of the response variable in both model predictions.

2.4 Model validation

Five test data set were used to evaluate the performance of the models. The prediction analysis relied on the coefficients (β) obtained from the regression models, which were then used to predict the CO_2 efflux. The process of model validation was conducted using XLSTAT statistical software.
To compare the results and determine the best model, the RMSE was calculated for each model. The RMSE serves as a measure of the average difference between the actual CO_2 efflux values and the predicted values generated by the model. Lower RMSE values indicate better model performance, as they signify a smaller discrepancy between the predicted and observed values. Therefore, by comparing the RMSE values of different models, the one with the lowest RMSE can be considered the best-performing model in terms of accuracy.

3. Results and Discussion

3.1 Data collection and pre-processing

638 CO_2 efflux observations for all variables of FWC amended soil treatments were obtained. These data were pre-processed to obtain preliminary information such as the outliers. There were 597 observations left after outlier removal.

3.2 Multiple linear regression of predictive model

There are a total of three predictive models as shown in Eq(1), Eq(2) and Eq(3) for the input soil respiration variables. The input and output sets X and Y are then regressed by using multiple least square regression (MLSR) algorithm to obtain the value of β. Each model gives different β values which leads to different prediction results for each method. Firstly, a set of training data was regressed with a multiple linear model.
Table 2 showed that there are 4 variables having p-values less than 0.05 including A,B,C and E which denoted that they are significant variables. C/N is the most influential variable. However, variables D,F,G and H are not bring significant information to explain the variability of the CO_2 efflux and can be removed from the model. Based on the table, the determination R^2 coefficient was 0.887, nearly close to 1, indicating that the prediction and actual values were nearly fit. Besides, the F value (8.787) and the p-value (<0.002, which is less than 0.05) of this model, both implied that this is a significant model. However, the sum square error (SSE) and mean square error (MSE) is not too close to 0. The mathematical equation for these model terms was presented in Eq(4).

$$CO_2\ Efflux = 69.079 + 0.075A - 0.767B + 0.481C - 0.035D - 0.407E + 0.005F - 0.688G - 0.008H \quad (4)$$

According to the Table 2, three of the most significant factors were selected and regressed with multiple linear regression again to determine the optimal CO_2 efflux. The chosen variables were FWC Amended Ratio (A), pH (B) and C/N ratio (C) while the CO_2 efflux was the response. The linear regression model (Model 1) was estimated by using Eq(1) in the TXLSTAT for the CO_2 efflux response. The goodness of fit statistics was used to evaluate the results.
Based on Table 3, the determination R^2 coefficient was 0.317, far away from unity, indicating that the prediction and actual values were not fit. Besides, the F value (2.167) and the p-value of the F statistic (0.138, which is more than 0.05) of this model, both implied that this is not a significant model. The linear equation for these model terms was presented in Eq(5).

Table 2: Type III Sum of Squares analysis (CO_2 Efflux)

Source	Sum of Square	DF	Mean Square	F Value	Pr > F	p-values signification codes
Model	3.407	8	0.426	8.787	0.002	**
A	0.364	1	0.364	7.509	0.023	°
B	0.320	1	0.320	6.606	0.030	*
C	0.231	1	0.231	4.771	0.047	***
D	0.077	1	0.077	1.597	0.238	°
E	0.374	1	0.374	7.721	0.021	*
F	0.010	1	0.010	0.212	0.656	°
G	0.137	1	0.137	2.820	0.127	*
H	0.000	1	0.000	0.003	0.956	°
Pure Error	0.436	9	0.048			
Corrected Error	3.843	17				
R-squared	0.887					

*Signification codes: 0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < ° < 1*

Table 3: Analysis of variance (CO_2 Efflux)

Source	Sum of Square	DF	Mean Square	F Value	Pr > F	p-values signification codes
Model	1.219	3	0.406	2.167	0.138	°
Pure Error	2.625	14	0.187			
Corrected Error	3.843	17				
R-squared	0.317					

Computed against model Y=Mean(Y)

$$CO_2\,Efflux = -1.290 - 0.034A + 0.108B + 0.172C \tag{5}$$

As conclusion, all the eight input variables are important. However, the multiple linear model with three significant variables (Model 1) is not compatible and less accurate to predict the CO_2 efflux. Therefore, a set of training data was proceeded to regress with a multiple non-linear model.

3.3 Multiple non-linear regression of predictive model

For the multiple non-linear models, the regressions were extended by using first-order predictive modelling (Model 2) and second-order predictive modelling (Model 3). The polynomial regression models were estimated using XLSTAT for the CO_2 efflux. The goodness of fit statistics was used to evaluate the results.
Table 4 shows the SSE, MSEP and RMSE towards all model equations. After obtaining the forecasted output, MSEP was calculated between actual output data with the predicted value to investigate the prediction model efficiency. The prediction graph was shown in Figure 1 for each model. The MSEP closest to zero is the most accurate.

Table 4: Goodness of Fit Statistics

Predictive Modelling	Model	SSE	MSEP	RMSE
Multiple Linear	1	2.625	0.187	0.433
Polynomial First Order	2	0.597	0.054	0.233
Polynomial Second Order	3	0.188	0.023	0.153

The MSEP values of Model 3 is significantly smallest compared to others. CO_2 efflux prediction fit second order well due to the interaction variable effect which the independent variables are multiplied with another independent variable. On the other hand, the response is well modelled by a non-linear function shows CO_2 efflux predictive modelling reflects second order behaviour more than a linear behaviour. Therefore, second order modelling is the most suitable for soil respiration prediction. The mathematical equation for Model 2 and Model 3 terms were presented in Eq(6) and Eq(7), respectively.

$$CO_2\,efflux = 281.925 - 9.4006A + 39.983B + 48.179C + 1.357AB - 0.005AC - 6.853BC \tag{6}$$

$$CO_2\,efflux = 152.074 + 6.639A - 10.884B - 35.018C - 1.024AB + 0.037AC + 5.654BC - 0.0002A^2 - 1.977B^2 - 0.208C^2 \quad (7)$$

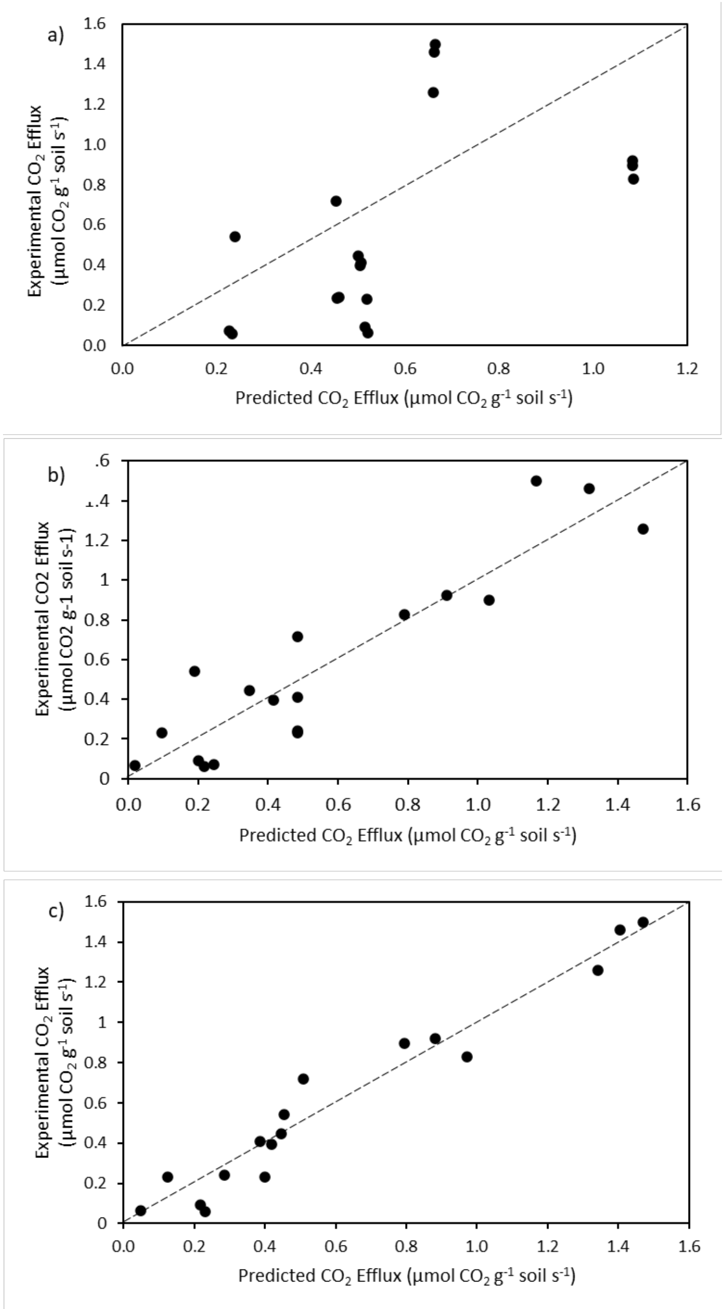

Figure 1: Comparison of predicted vs experimental CO2 efflux for (a) Model 1, (b) Model 2 and (c) Model 3

3.4 Predictive modelling validation for soil respiration

Assessing the model's quality requires quantifying and reporting the predictive accuracy of the generated models. The model is tested on an independent dataset that was not used to develop the model. A broadly used method to evaluate predictive validity for continuous outcomes is the MSE. The MSE is calculated as the average of the squared differences between the observed and predicted values. A smaller value for MSE indicates that the predicted values are closer to the observed data and therefore a better prediction. The results from the predicted and experimental CO_2 efflux were shown in Table 5.

Table 5: Predicted vs experimental data of CO_2 efflux

Data	Experimental CO_2 Efflux	Predicted CO_2 Efflux			Mean Square Error (MSE) (%)		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
1	0.225	0.234	0.218	0.225	3.810	2.936	0.084
2	0.130	0.519	0.107	0.130	298.386	18.174	0.333
3	0.397	0.457	0.485	0.409	14.965	22.082	2.978
4	1.406	0.661	1.319	1.416	53.006	6.161	0.756
5	0.882	1.082	0.912	0.869	22.651	3.400	1.555
			Average MSE (%)		78.566	10.550	1.142

The MSE (%) value between the predicted and the actual value for Model 1, Model 2 and Model 3 were 78.566 %, 10.55 % and 1.142 % respectively. The results obtained through XLSTAT proved that Model 3 produces the smallest MSE between experimental and predicted value of CO_2 efflux and gives a clearer trend of data. On the other hand, Model 1 and Model 2 give a significant increment in MSE, it indicates that CO_2 efflux prediction does not reflect a linear behavior. Therefore, second-order modelling (Model 3) should be chosen as the model fits a non-linear response and improve the framework for CO_2 efflux prediction. The polynomial second order modelling is shown in Eq(8) for CO_2 efflux.

$$\begin{aligned} CO_2\, efflux = {} & 152.074 + 6.639A - 10.884B - 35.018C - 1.024AB + 0.037AC \\ & + 5.654BC - 0.0002A^2 - 1.977B^2 - 0.208C^2 \end{aligned} \tag{8}$$

4. Conclusions

As conclusion, the study's findings underscored the significance of employing a second-order polynomial model in enhancing the precision of CO_2 efflux prediction. This assertion was substantiated by the model's remarkably minimal average error percentage, measuring at 1.14 %. This study aids in the order of CO_2 efflux predictive modelling and non-linearity of input to output variables.

Acknowledgments

The author would like to thank for the financial supports from Universiti Teknologi Malaysia (UTM) and Ministry of Higher Education Malaysia for the funding of this research through FRGS Grant (FRGS/1/2021/TK0/UTM/02/38).

References

Chen X., Post W.M., Norby R.J., Classen A.T., 2011, Modeling soil respiration and variations in source components using a multi-factor global climate change experiment, Climatic Change, 107(3-4), 459-480.

Chia S.R., Chew K.W., Leong H.Y., Ho S.H., Munawaroh H.S.H., Show P.L., 2021, CO_2 mitigation and phycoremediation of industrial flue gas and wastewater via microalgae-bacteria consortium: Possibilities and challenges, Chemical Engineering Journal, 425, 131436.

Dolit S.A.M., Asli U.A., Khamis A.K., Baharulrazi N., Azman N.R., Yunus N.A., 2022, Improving the water retention characteristics of sandy soil using food waste compost amendment and indigenous microorganisms, Chemical Engineering Transactions, 97, 211-216.

Hibbard K.A., Law B.E., Reichstein M., Sulzman J., 2005, An analysis of soil respiration across northern hemisphere temperate ecosystems, Biogeochemistry, 73(1), 29-70.

Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown, K., Vargas, R., 2016, The value of soil respiration measurements for interpreting and modeling terrestrial carbon cycling, Plant and Soil, 413(1-2), 1-25.

Ryan M.G., Law B.E., 2005, Interpreting, measuring, and modeling soil respiration, Biogeochemistry, 73(1), 3-27.

Sharma, S., Singh, P., Kumar, S., 2020, Responses of soil carbon pools, enzymatic activity, and crop yields to nitrogen and straw incorporation in a rice-wheat cropping system in North-Western India, Frontiers in Sustainable Food Systems, 4.

Su X., Wang Y., He Q., Hu X., Chen Y., 2019, Biochar remediates denitrification process and N_2O emission in pesticide chlorothalonil-polluted soil: Role of electron transport chain, Chemical Engineering Journal, 370, 587-594.

Subke J.A., Inglima I., Cotrufo M.F., 2006, Trends and methodological impacts in soil CO_2 efflux partitioning: A metaanalytical review, Global Change Biology, 12, 921-943.