Artificial neural network-based repair and maintenance cost estimation model for rice combine harvesters

Apsornrat Numsong[1], Jetsada Posom[1,2], Somchai Chuan-Udom[1,2*]

(1. *Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Thailand*;
2. *Applied Engineering for Important Crops of the North East Research Group, Khon Kaen University, Thailand*)

Abstract: This research proposes an artificial neural network (ANN)-based repair and maintenance (R&M) cost estimation model for agricultural machinery. The proposed ANN model can achieve high estimation accuracy with small data requirement. In the study, the proposed ANN model is implemented to estimate the R&M costs using a sample of locally-made rice combine harvesters. The model inputs are geographical regions, harvest area, and curve fitting coefficients related to historical cost data; and the ANN output is the estimated R&M cost. Multilayer feed-forward is adopted as the processing algorithm and Levenberg-Marquardt backpropagation learning as the training algorithm. The R&M costs are estimated using the ANN-based model, and results are compared with those of conventional mathematical estimation model. The results reveal that the percentage error between the conventional and ANN-based estimation models is below 1%, indicating the proposed ANN model's high predictive accuracy. The proposed ANN-based model is useful for setting the service rates of agricultural machinery, given the significance of R&M cost in profitability. The novelty of this research lies in the use of curve-fitting coefficients in the ANN-based estimation model to improve estimation accuracy. Besides, the proposed ANN model could be further developed into web-based applications using a programming language to enable ease of use and greater user accessibility. Moreover, with minor modifications, the ANN estimation model is also applicable to other geographical areas and tractors or combine harvesters of different countries of origin.

Keywords: repair and maintenance cost, estimation model, artificial neural network, curve fitting coefficients, combine harvesters

DOI: 10.25165/j.ijabe.20231602.5931

Citation: Numsong A, Posom J, ChuanUdom S. Artificial neural network-based repair and maintenance cost estimation model for rice combine harvesters. Int J Agric & Biol Eng, 2023; 16(2): 38–47.

1 Introduction

In agricultural machinery operation, fuel costs account for the largest proportion of the variable cost, varying by the extent of operation[1]. Another significant variable-cost item is repair and maintenance (R&M) cost, which is a function of annual usage and machinery service life[2]. The ownership and operating costs constitute the key cost components of agricultural machinery investment. The ownership (fixed) cost, including depreciation, interest on investment, tax and insurance, is straightforward. On the other hand, the operating costs (e.g., fuel, lubricant, labor, R&M) vary in amount subject to the extent of operation.

As a result, a model to estimate the operating costs of agricultural machines, particularly the R&M outlay, is vital to the effective cost management and maximum return on investment[4-6]. The R&M cost of agricultural machinery is nonlinear, subject to a variety of factors including machinery age, extent of operation, harvest area, soil condition, crop type, and operators' skills and experience[7].

Existing R&M cost estimation techniques are categorized into two groups: mathematical and ANN models. The first group consists of: (a) the American Society of Agricultural Engineers (ASAE) standards[2,4,5,8-14], and (b) curve fitting methods[15-21]. The mathematical models require a substantial amount of data for accurate estimation. In addition, both ASAE and curve fitting methods suffer from area-specificity issues. Given the variable nature of R&M cost, definitive mathematical relationships between R&M cost and the determining factors (machinery age, extent of operation, harvest area, soil condition, crop type) are less straightforward and difficult to establish.

As an alternative to the conventional mathematical methods, artificial neural network (ANN) is adopted to estimate the R&M cost of agricultural machinery. Ranjbar et al.[22] comparatively estimated the R&M costs of tractors using two neural network structures (between single network and separate networks), the result found that a single network gave a better result than using separate networks for estimation of each cost component. They summarized that neural network could be improved the economic decision making capabilities of machinery managers. Rohani et al.[23] estimated the R&M costs of two-wheel-drive tractors using ANN and conventional mathematical models; and reported that the ANN model provided the accuracy with the coefficient of determination (R^2) and root mean square error (RMSE) of 0.99 and 0.3674, respectively. BDLRF with feed-forward back-propagation (FFBP) algorithms Azim et al.[24] predict the R&M cost of two-wheel-drive tractors using the multi-layer neural network with Feed Forward Backpropagation training algorithm (FFBP), the performance of Backpropagate Declining Learning Rate Factor algorithm (BDLRF) has been compared with Feed-Forward

Received date: 2020-05-26 **Accepted date:** 2022-10-09

Biographies: Apsornrat Numsong, PhD, research interest: agricultural mechanization, Email: apsornrat@gmail.com; **Jetsada Posom**, PhD, Associated Professor, research interests: biosensor, Email: jetspo@kku.ac.th.

***Corresponding author: Somchai Chuan-Udom**, PhD, Associated Professor, research interests: agricultural mechanization. Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, 123 Mittraphap Road, Khon Kaen, 40002, Thailand. Tel: +668-9712-1247, Email: somchai.chuan@gmail.com.

Backpropagation algorithm (FFBP), the result shows that training Feed Forward Backpropagation algorithm (FFBP) surpasses the (BDLRF) algorithm in predicting tractor R&M costs by using separate networks rather than a single network.

Despite significantly less data requirement, the existing ANN models suffer from limited estimation accuracy compared to the mathematical models[25]. In light of large data requirements of mathematical models and limited predictive accuracy of existing ANN models, this research proposes an ANN-based R&M outlay estimation model for agricultural machinery. The proposed ANN-based estimation model can achieve high predictive accuracy with small data requirement. In this study, the proposed ANN-based model is implemented to estimate the R&M costs using a sample of locally-made rice combine harvesters in the rice-growing regions of Thailand.

The inputs of the ANN estimation model are geographical regions (Thailand's northern, northeastern, and central regions), size of harvest area, and curve fitting coefficients related to historical cost data; and the ANN output is estimated R&M cost.

The neural network convergence algorithm of Levenberg-Marquardt backpropagation has many advantages compared to the traditional backpropagation, Levenberg Marquardt (LM) based back propagation (BP) has better performance (in term of convergence speed and rate) than other algorithms such as Artificial Bee Colony-Levenberg Marquardt (ABC-LM), Artificial Bee Colony- back propagation (ABC-BP) and back propagation neural network (BPNN) algorithms[26]. Sapna et al.[27] concluded that Levernberg-Marquardt algorithm gives the best performance in the prediction of diabetes compared to any other backpropogation algorithm. Multilayer feed-forward and Levenberg-Marquardt backpropagation learning algorithms are used for R&M cost estimation. Unlike previous ANN models, the proposed ANN estimation model incorporates curve-fitting coefficients, which is part of the mathematical technique, into the model to improve the predictive accuracy. To validate, the ANN-based R&M cost estimations are calculated, and results are compared with those of conventional estimation model. The proposed ANN-based model is useful for setting the rental rate or service charge of agricultural machinery in an efficient and reasonable manner, given the significance of R&M cost in profitability.

2 Research methodology

The research methodology consists of three stages: data collection, evolution of the algorithmic scheme and training, and validation. In the data collection stage, field survey is carried out to garner data on purchase prices, harvest areas, machine ages, and historical annual R&M costs. The curve-fitting coefficients of mathematical functions are then determined.

In the ANN algorithmic scheme evolution stage, multilayer feed-forward is adopted as the processing algorithm and Levenberg-Marquardt (LM) backpropagation learning as the training algorithm. The post-training ANN-based algorithmic scheme is subsequently established to estimate the R&M cost. In the validation stage, the ANN-based R&M cost estimations are calculated, and results are compared with those of conventional estimation model. The validation stage is detailed in the Results and Discussion section.

2.1 First stage: data collection

2.1.1 Study area and research data

In this research, a field survey was undertaken with a random sample of 100 owners of locally-made rice combine harvesters in 30 rice-growing provinces (excluding pre-owned vehicles). The owners have maintained detailed records of R&M costs since the first year of machine acquisition. Since the survey participants are required to have a complete record of repairs and maintenance, the sample size is therefore limited to 100 combine harvesters. Besides, previous research works on the relationship between R&M cost and usage relied on a minimum sample size of 30[4,7]. Since there exists no official and systematic record keeping of R&M costs in Thailand, this research conducts a field survey by face-to-face interviews using a semi-structured questionnaire to collect the data. This method is straightforward and efficient to collect data from participants[27].

The 30 rice-growing provinces consist of 7 provinces in the North (20 combine harvesters), 11 provinces in the Northeast (40), and 12 provinces in the Central Plains (40). In Thailand, rice cultivation is densely concentrated in the central region due to fertile lands and efficient irrigation. The soil condition of the northern and northeastern regions are of saline soil and gravelly, while that of the central region is of clay loam. Since topographical features vary from region to region which influence operation and R&M costs of rice combine harvesters, this research thus uses the data of different geographical regions. The average age of the rice combine harvesters is six years. The field survey data include initial acquisition costs (purchase prices), years in service, annual harvest areas, and annual R&M outlays, including lubricants, oil filter, spare parts, and labor.

Table 1 lists the specifics of the surveyed rice combine harvesters. Rice combine harvesters are categorized by cutting widths into small and large combine harvesters. Due to different farm scales and crop types, the large cutting widths (5-6 m) are normally used in European countries and the U.S., while the small widths (1-4 m) are ubiquitous in Asian countries[28].

Table 1 Specifics of locally-made rice combine harvesters

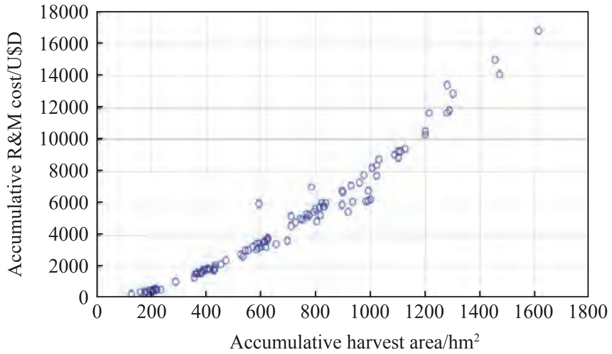
	Northern	Northeastern	Central plains
Acquisition cost/USD	63 500-77 500	60 500-74 500	64 000-77 500
Engine power/hp	195-270	195-270	195-270
Cutting width/m	2.8-3.25	2.85-3.25	2.85-3.75
Gain capacity/t	2.5-2.7	2.5-2.7	2.5-2.7
Average age/a	5.7	6.2	7
Average harvest area per vehicle/$hm^2 \cdot a^{-1}$	179.52	189	202
Average R&M costs per vehicle/$USD \cdot a^{-1}$	507	643	527.3

Source: author's survey.

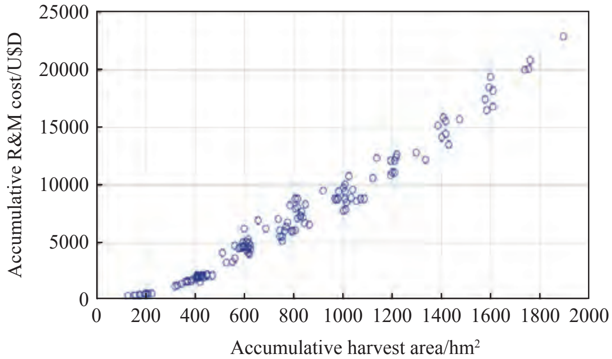
2.1.2 Mathematical curve-fitting coefficients

Figures 1a-1c respectively illustrate the accumulative R&M cost, using MATLAB/Simulink, of rice combine harvesters (USD) relative to harvest area (hm^2) of Thailand's northern, northeastern, and central regions. The relationships between R&M cost and harvest area are nonlinear. Conventionally, the accumulative R&M cost as a percentage of initial purchase price is a function of accumulative hours of machinery use, and the machine is replaced upon reaching a predetermined maximum hour-usage threshold. However, this practice is impractical in the Thai setting due to a lack of hour-based R&M expenditures. In Thailand, the R&M outlay is in lump sum amount per total harvest area annually.

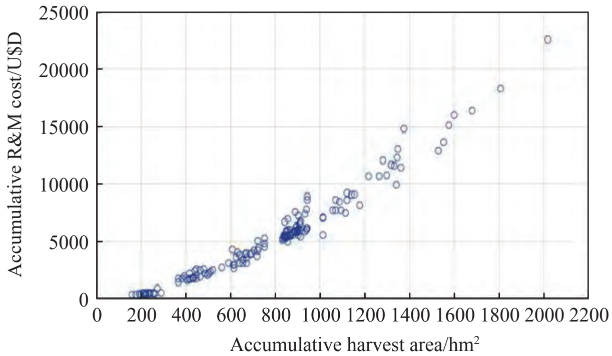
With the accumulative R&M cost relative to harvest area, the curve fitting models, based on coefficients of power, polynomial, exponential, and linear functions, are subsequently established. The accumulative R&M cost (y) based on power (Equation (1)), polynomial (Equation (2)), exponential (Equation (3)), and linear



a. Northern

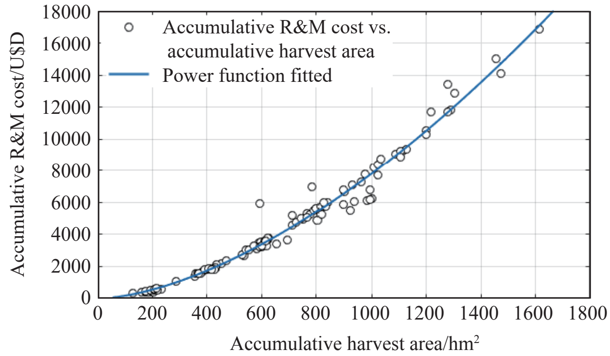


b. Northeastern

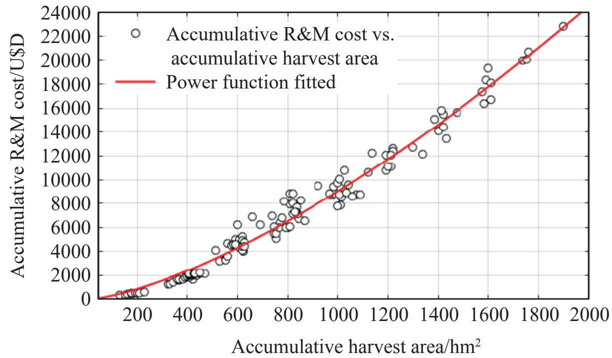


c. Central regions of thailand

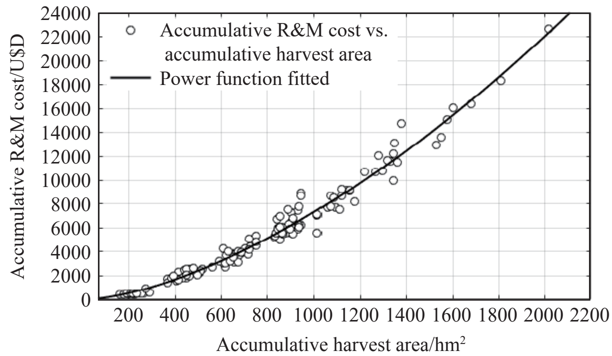
Figure 1 Accumulative R&M cost relative to harvest areas



a. Northern



b. Northeastern



c. Central regions

Figure 2 Power function-fitted accumulative R&M cost relative to accumulative harvest areas

functions (Equation (4)) are:

$$y = ax^b + c \tag{1}$$

$$y = ax^2 + bx + c \tag{2}$$

$$y = ae^{bx} + c \tag{3}$$

$$y = ax + b \tag{4}$$

where, y is the accumulative R&M cost; a, b, c are the curve fitting coefficients; and x is accumulative harvest area. The number e, also known as Euler's number, is a mathematical constant approximately equal to 2.718 28.

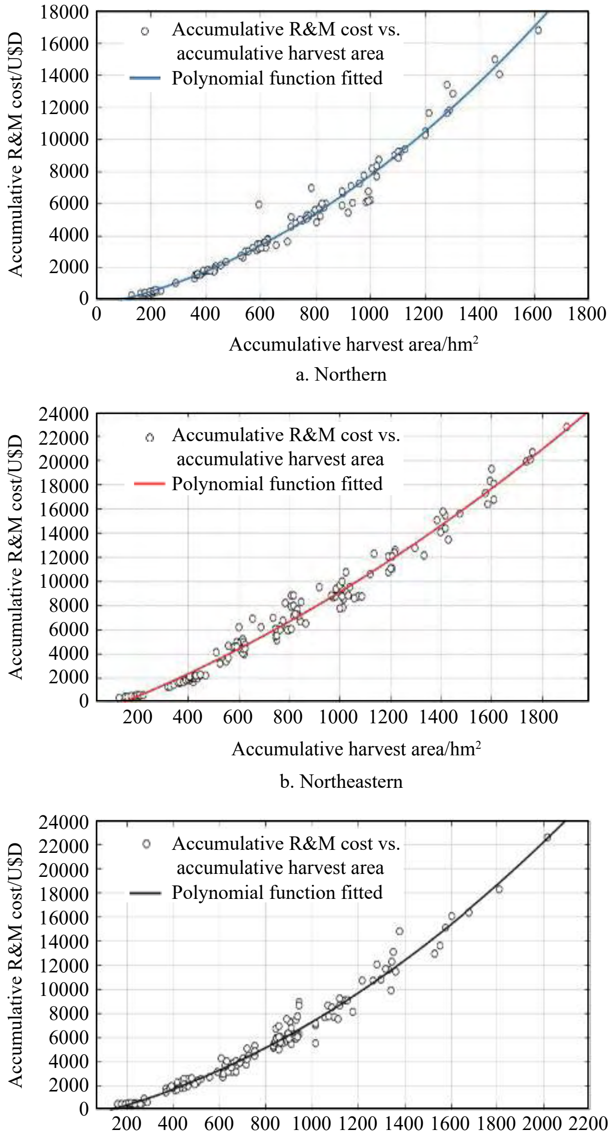
To obtain the curve fitting coefficients of the power function, the accumulative R&M cost and harvest area data are fitted into Equation (1) (i.e., the power function). The results are graphically depicted by geographical region in Figures 2a-2c.

To acquire the curve fitting coefficients of the polynomial function, the accumulative R&M cost and harvest area data are fitted into the polynomial function (Equation (2)). Figures 3a-3c illustrate the polynomial function-fitted accumulative R&M cost relative to accumulative harvest area of the country's North, Northeast, and Central Plains, respectively.

Figures 4a-4c show the exponential function-fitted accumulative R&M cost (Equation (3)) relative to accumulative harvest area of the northern, northeastern, and central regions, respectively. The corresponding linear function-fitted accumulative R&M outlay (Equation (4)) in relation to accumulative harvest area are depicted in Figures 5a-5c.

In Figures 2a-2c, the minimum accumulative R&M outlays of power function fitting curve of three geographical regions approach zero. The finding indicates that the power function-fitted accumulative R&M cost estimation model is applicable to small-, moderate-, and large-scale harvest areas.

In Figures 3a-3c, the minimum polynomial function-fitted accumulative R&M expenditures of the North and Central Plains approach zero, while that of the northeastern region is negative. The negative R&M outlay of the Northeast is attributable to scarcity of repairs and maintenance during early machine service life. The polynomial function-fitted R&M cost estimation model is thus suitable for small- to large-scale farmland in the northern and central regions but unfit for small-scale harvest areas in the

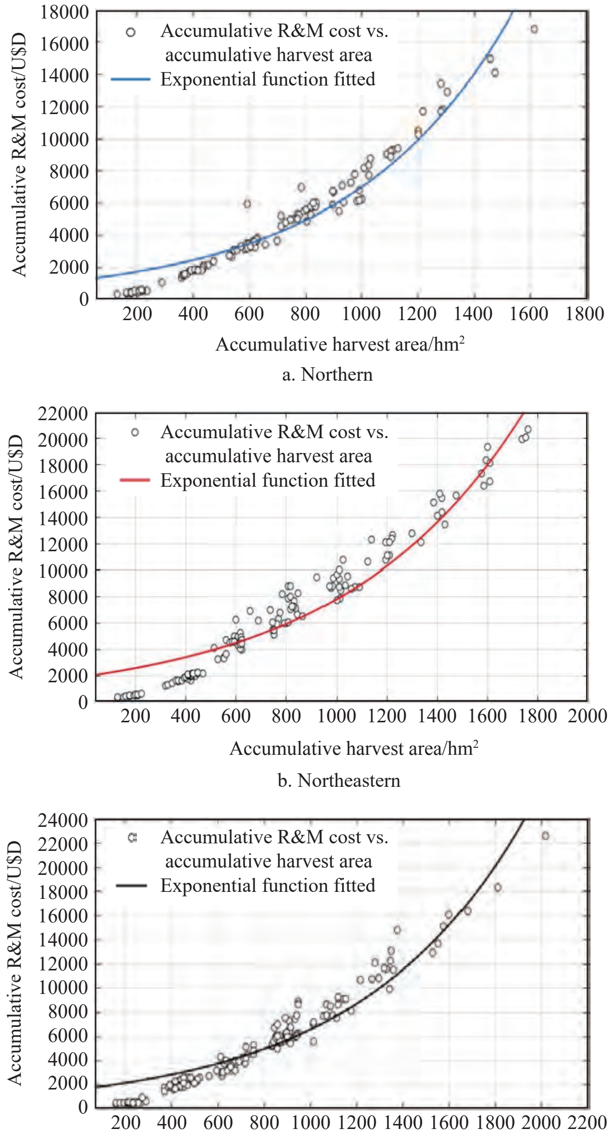


a. Northern

b. Northeastern

c. Central regions

Figure 3 Polynomial function-fitted accumulative R& M cost relative to accumulative harvest areas



a. Northern

b. Northeastern

c. Central regions

Figure 4 Exponential function-fitted accumulative R&M cost relative to accumulative harvest areas

Northeast.

In Figures 4a-4c, the minimum exponential function-fitted R&M accumulative cost of rice combine harvesters of three geographical regions is around USD 2000. The excessive minimum R&M expenditure is contrary to logic, rendering the exponential-fitted R&M cost estimation model impractical. Meanwhile, due to the non-linearity of R&M outlay of agricultural machinery, the linear function curve fitting model is non-ideal for estimation of R&M cost, as shown in Figures 5a-5c.

Table 2 lists the curve fitting coefficients (a, b, and c) of power, polynomial, exponential, and linear functions by geographical regions (Figures 2-5). In the table, the power-function coefficients of determination (R^2) of the three geographical regions are 0.9775-0.9806; and those (R^2) of polynomial, exponential, and linear functions are 0.9776-0.9820; 0.9132-0.9348; and 0.9465-0.9725. The large R^2 indicate high predictive accuracy of the mathematical functions.

Source: curve-fitting function based on survey data. RMSE : root mean squared errors.

In Table 2, in addition to the straightforwardness of power function model, its root mean squared errors (RMSE) for the three geographical regions are comparably small. Besides, the power function-fitted model is commonly used in estimation of R&M expenses of agricultural machinery[15,18,20,21].

2.2 Second stage: Algorithmic scheme evolution and training

In the second stage, multilayer feed-forward algorithm is adopted as the processing algorithm of the ANN-based estimation model. In the ANN training, Levenberg-Marquardt backpropagation learning is used as the training algorithm.

2.2.1 Multilayer feed-forward algorithm

Multilayer feed-forward algorithm is used as the processing algorithm to estimate the R&M cost, which is a function of geographical region (R_E), harvest area (x), and power-function curve fitting coefficients (a, b) (Table 2). Figure 6 illustrates the schematic of ANN-based R&M cost estimation model, consisting of three layers: input (R), hidden (S), and output (T) layers.

For the input layer ($[P]_{(1\times i)}$), the geographical regions ($p_{1,1}$) are Thailand's northern (N), northeastern (NE), and central (C) regions; and the harvest area (x; $p_{1,2}$) is in hectare (hm²). The power-function curve fitting coefficients (a, b; $p_{1,3}$, $p_{1,4}$) are obtained from the

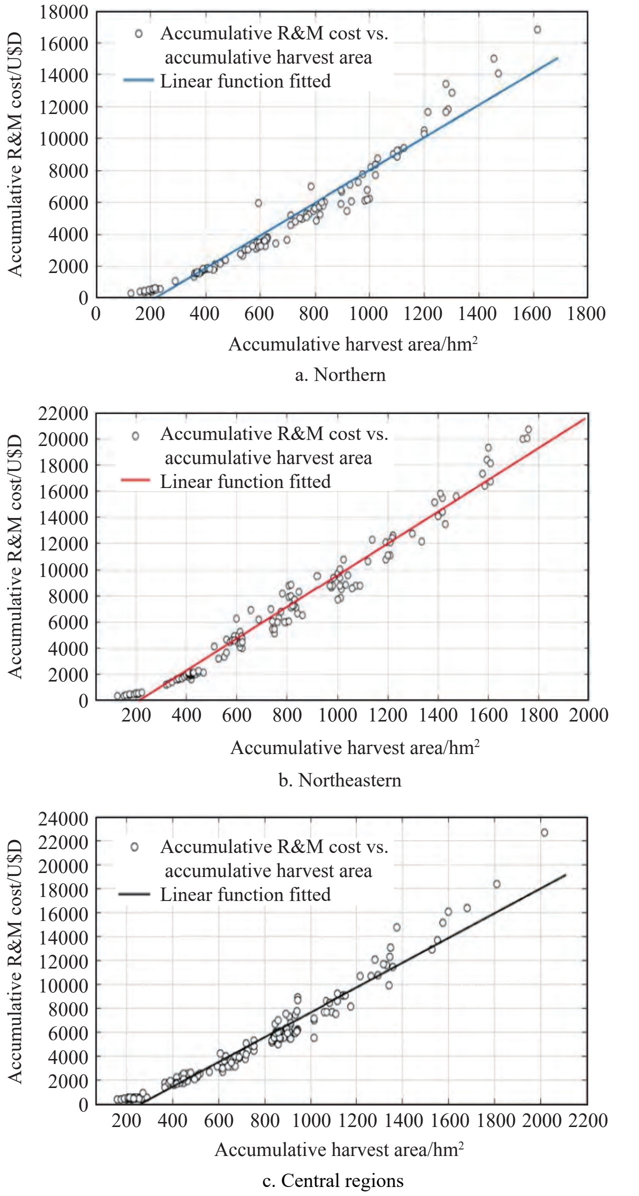


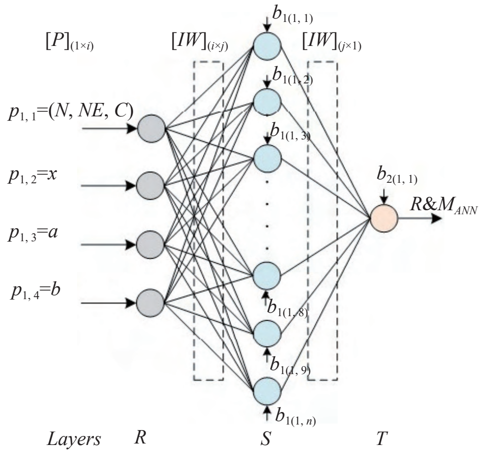
Figure 5 Linear function-fitted accumulative R&M cost relative to accumulative harvest areas

Table 2 Curve fitting coefficients by geographical region

Regions	Equations	R^2	*RMSE*	Curve Fitting Coefficients *a*	*b*	*c*
North (*N*)	$y=ax^b+c$	0.9801	521	0.0968	1.636	–
	$y=ax^2+bx+c$	0.9803	520.3	0.0046	3.487	–331.8
	$y=ae^{bx}+c$	0.9348	942.1	1226	0.0017	–
	$y=ax+b$	0.9500	825.3	10.21	–2197	–
Northeastern (*NE*)	$y=ax^b+c$	0.9806	746.6	0.4274	1.442	–
	$y=ax^2+bx+c$	0.9820	722	0.0024	7.839	–1203
	$y=ae^{bx}+c$	0.9205	1511	1962	0.0013	–
	$y=ax+b$	0.9725	887.9	12.15	–2566	–
Central (*C*)	$y=ax^b+c$	0.9775	571.9	0.1216	1.593	–
	$y=ax^2+bx+c$	0.9776	572.7	0.0035	4.417	–627.4
	$y=ae^{bx}+c$	0.9132	1123	2093	0.001 19	–
	$y=ax+b$	0.9465	881.3	10.33	–2655	–

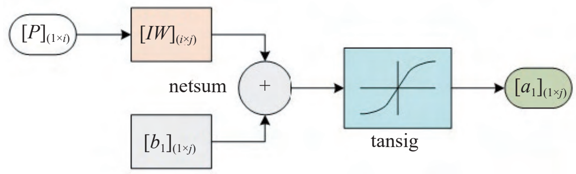
mathematical model. Hyperbolic tangent sigmoid transfer function (*tansig* in Figure 7) is the activation function of the input layer.

For the input-side hidden layer ($[IW]_{(i\times j)}$), the number of neurons is iteratively optimized by ANN (i.e., multilayer feed-forward algorithm) based on type or complexity of experimentation[28]. The relationships are expressed in Equations (5)-(6) and graphically depicted in Figure 7.



Note: The three layers: input (R), hidden (S), and output (T) layers. $p_{1,1}$=the geographical regions, $p_{1,2}=x$ (harvest area), $p_{1,3}=a$ (the power-function curve fitting coefficient a), $p_{1,4}=b$ (the power-function curve fitting coefficient b), $[P]_{(1\times i)}$=input layer matrix, $[IW]_{(i\times j)}$=input weight matrix, $b1_{(1\times j)}$=input layer bias, $[LW]_{(j\times 1)}$=layer weight matrix, $b2_{(1\times 1)}$=output layer bias, $R\&M_{ANN}$=R&M cost estimated by ANN.

Figure 6 Schematic of ANN-based R&M cost estimation model.



Note: $[n_1]_{(1\times j)}$ = the summation of input weight matrix $[IW]_{(i\times j)}$ multiplied by input data $[P]_{(1\times i)}$ (i.e., R_E, x, a, b) plus input layer bias $[b_1]_{(1\times j)}$; and $[a_1]_{(1\times j)}$ is the result of hyperbolic tangent sigmoid transfer function of $[n_1]_{(1\times j)}$

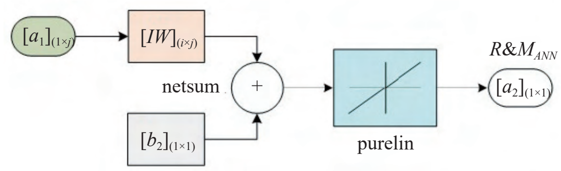
Figure 7 Schematic of input weight layer ($[IW]_{(i\times j)}$) of ANN-based R&M cost estimation model

$$[n_1]_{(1\times j)} = \sum_{j=1}^{n}[P]_{(1\times i)}[IW]_{(i\times j)} + [b_1]_{(1\times j)} \tag{5}$$

$$[a_1]_{(1\times j)} = \tan sig([n_1]_{(1\times j)}) \tag{6}$$

where, $[n_1]_{(1\times j)}$ is the summation of input weight matrix $[IW]_{(i\times j)}$ multiplied by input data $[P]_{(1\times i)}$ (i.e., R_E, x, a, b) plus input layer bias $[b_1]_{(1\times j)}$; and $[a_1]_{(1\times j)}$ is the result of hyperbolic tangent sigmoid transfer function of $[n_1]_{(1\times j)}$ (Figure 7).

For the output-side hidden layer ($[LW]_{(j\times 1)}$), the relationships are expressed in Equations (7)-(8) and graphically depicted in



Note: $[n_2]_{(1\times 1)}$ is the summation of layer weight matrix $[LW]_{(j\times 1)}$ multiplied by $[a_1]_{(1\times j)}$ of the input layer plus output layer bias $[b_2]_{(1\times 1)}$, and $[a_2]_{(1\times 1)}$ is the result of linear transfer function of $[n_2]_{(1\times 1)}$.

Figure 8 Schematic of output weight layer of ANN-based R&M cost estimation model

Figure 8. The R&M cost of rice combine harvesters (the ANN output) is calculated by Equation (9), in which linear transfer function (*purelin* in Figure 8) is the activation function.

$$[n_2]_{(1\times1)}=\sum_{j=1}^{n}[a_1]_{(1\times j)}\cdot[LW]_{(j\times1)}+[b_2]_{(1\times1)} \tag{7}$$

Substituting the hyperbolic tangent sigmoid transfer function (*tan sig*) in a_1,

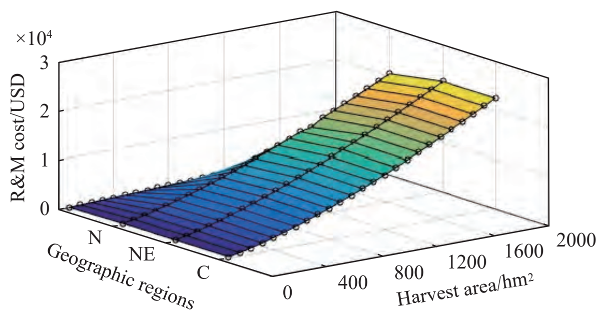
$$[n_2]_{(1\times1)}=\sum_{j=1}^{n}\left[\left[tansig\left([n_1]_{(1\times j)}\right)\right]\cdot[LW]_{(j\times1)}\right]+[b_2]_{(1\times1)} \tag{8}$$

$$R\&M_{ANN}=[a_2]_{(1\times1)}=purlin([n_2]_{(1\times1)}) \tag{9}$$

where, $[n_2]_{(1\times1)}$ is the summation of layer weight matrix $[LW]_{(j\times1)}$ multiplied by $[a_1]_{(1\times j)}$ of the input layer plus output layer bias $[b_2]_{(1\times1)}$, and $[a_2]_{(1\times1)}$ is the result of linear transfer function of $[n_2]_{(1\times1)}$.

2.2.2 Training the ANN estimation model

In the ANN training, Levenberg-Marquardt backpropagation learning is used as the training algorithm. The training datasets include geographical regions, harvest area, power-function coefficients (*a, b*), and target R&M cost based on the conventional power function model (i.e., output), as illustrated in Figure 9. In the figure, the geographical regions are the northern (N), northeastern (NE), and central (C) regions. The harvest area is varied between 0, 400, 800, 1200, 1600 and 2000 hm². Based on the field survey, the maximum accumulative harvest area is around 2000 hm².



Note: *N*, *NE*, *C* denote northern, northeastern, and central regions.

Figure 9 Training datasets of the proposed ANN estimation scheme

Figures 10a-10c respectively illustrate the power-function coefficients (*a*,*b*) relative to harvest area of the northern, northeastern, and central regions. The coefficients (*a*, *b*) are 0.0968 and 1.636; 0.4274 and 1.442; and 0.1216 and 1.593 for the North, Northeast, and Central Plains, respectively.

In Figures 6-8, the matrix size of input data $[P]_{(1\times i)}$ is [1×4] (i.e., R_E, *x, a, b*). After a series of trial and error, the number of neurons of 10 (*n*=10) is selected, given large R^2 (0.999 78) and optimal response time. The matrix size of the input weight $[IW]_{(i\times j)}$ is thus [4×10] and that of bias $[b_1]_{(1\times j)}$ is [1×10]. The matrix size of the output layer weight $[LW]_{(j\times1)}$ is [10×1] and that of $[b_2]_{(1\times1)}$ is [1×1]. In training the ANN, Levenberg-Marquardt learning algorithm is used to vary *IW* and *LW* of the ANN. The ANN algorithmic output is the R&M cost of rice combine harvesters. The aforesaid relationships are expressed in Equations (10)-(14).

$$[P_i]_{(1\times i)}=\begin{bmatrix}p_{1,1} & p_{1,2} & \cdots & p_{1,i}\end{bmatrix}_{(1\times4)} \tag{10}$$

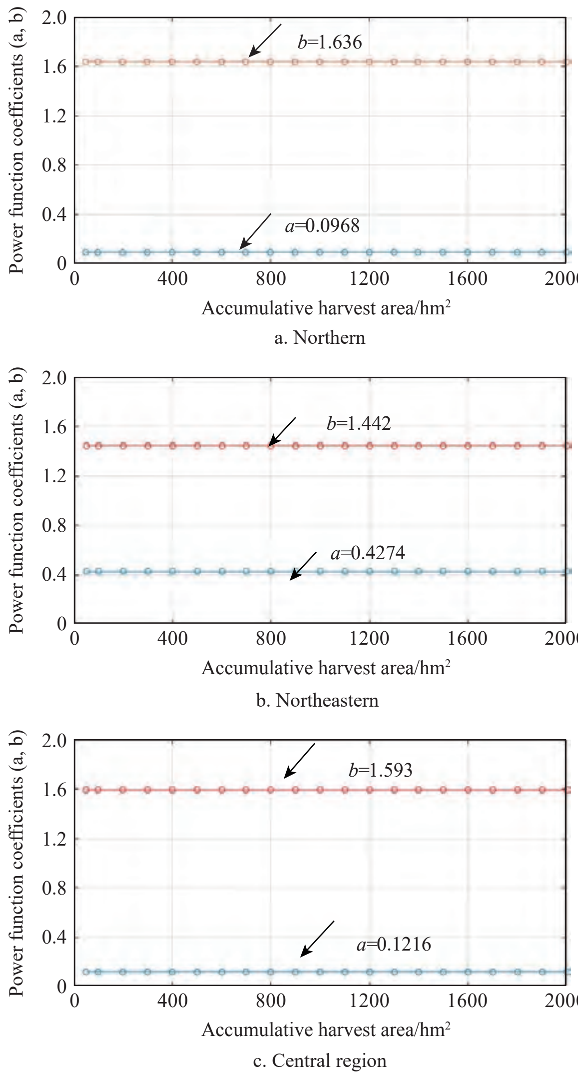


Figure 10 Power-function coefficients (*a*, *b*) relative to harvest areas

$$[IW]_{(i\times j)}=\begin{bmatrix}iw_{1,1} & iw_{1,2} & \cdots & iw_{1,n}\\ iw_{2,1} & iw_{2,2} & \cdots & iw_{2,n}\\ iw_{3,1} & iw_{3,2} & \cdots & iw_{3,n}\\ iw_{4,1} & iw_{4,2} & \cdots & iw_{4,n}\end{bmatrix}_{(4\times10)} \tag{11}$$

$$[b_1]_{(1\times j)}=\begin{bmatrix}b_{1,1} & b_{1,2} & \cdots & b_{1,j}\end{bmatrix}_{(1\times10)} \tag{12}$$

$$[LW]_{(j\times1)}=\begin{bmatrix}lw_{1,1}\\ lw_{2,1}\\ \vdots\\ lw_{j,1}\end{bmatrix}_{(10\times1)} \tag{13}$$

$$[b_2]_{(1\times1)}=[b_{2(1,1)}]_{(1\times1)} \tag{14}$$

Figure 11 shows the internal validation result of post-training ANN algorithmic scheme whose R^2 is 0.999 78, indicating very high predictive accuracy of the algorithmic scheme. Figure 12a illustrates the geometry of the post-training ANN algorithmic scheme to estimate the R&M cost, and the Matlab/Simulink procedural scheme of the ANN-based estimation model is depicted in Figure 12b.

In estimation of R&M cost (Figure 12b), a geographical region (North, Northeast, or Central Plains) and corresponding region-specific coefficients (*Coef a* (*N*), *Coef a* (*Ne*), *Coef a* (*C*), *Coef b* (*N*), *Coef b* (*Ne*), *Coef b* (*C*)) are manually selected using switches 1-6, where *N*, *NE*, and *C* denote the northern, northeastern, and central

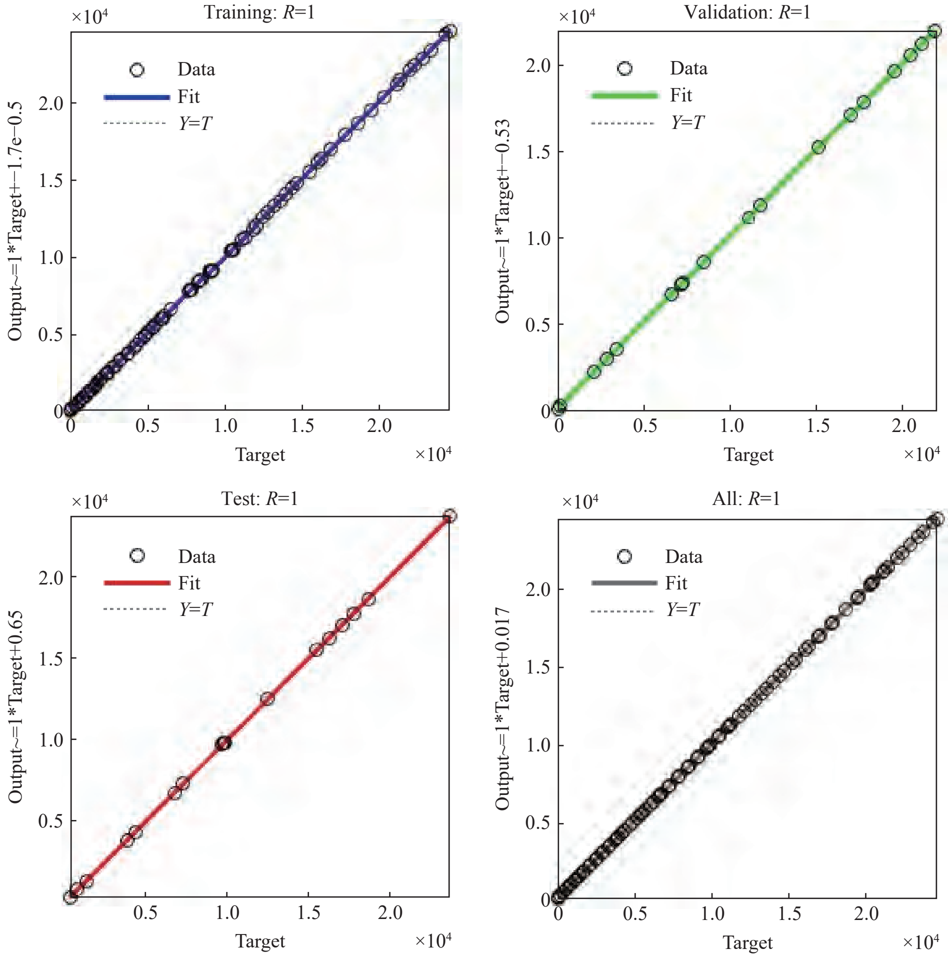
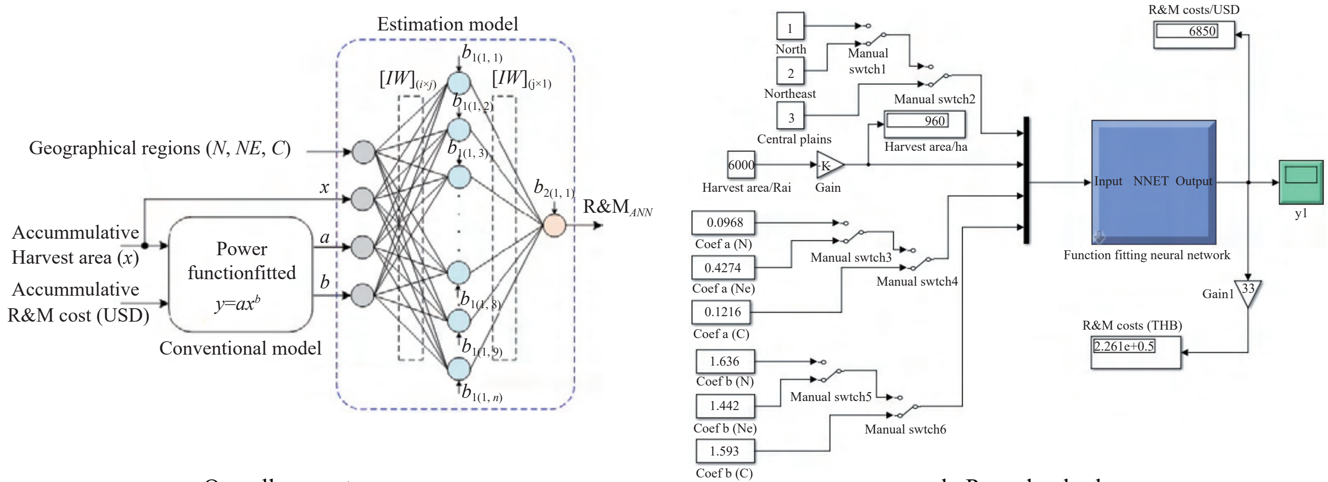


Figure 11 Internal validation of post-training ANN algorithmic scheme



a. Overall geometry b. Procedural scheme

Figure 12 Schematic of the post-training ANN algorithmic scheme for R&M cost estimation

regions. The harvest area is varied between 0-2000 hm². The ANN algorithmic scheme is also applicable to a harvest area larger than 2000 hm². The power-function coefficients (a, b) of the North, Northeast, and Central Plains are 0.0968 and 1.636, 0.4274 and 1.442, 0.1216 and 1.593, respectively (Figure 10).

3 Results and discussion

This section is concerned with the validation results of the proposed ANN algorithmic scheme. In the validation stage (i.e., the third stage), the ANN-based R&M cost estimations are calculated, and the results are compared with those of mathematical estimation model.

Table 3 lists the simulation parameters of the conventional power-function curve-fitting and ANN-based estimation models of

Table 3 Simulation parameters of power-function curve fitting and ANN-based estimation models

Geographical region	Harvest area (x)/hm²	Curve fitting coefficients	
		a	b
North (N)	400-2000	0.0968	1.636
Northeast (NE)	400-2000	0.4274	1.442
Central (C)	400-2000	0.1216	1.593

the North (N), Northeast (NE), and Central Plains (C). The harvest area (x) is varied between 400, 800, 1200, 1600, and 2000 hm^2.

3.1 R&M cost estimation using conventional curve-fitting model

The power-function curve-fitting R&M cost estimation of N, NE, and C are respectively expressed in Equations (15)-(17). The differences in the power-function coefficients (a, b) between the northern (N), northeastern (NE), and central (C) regions are attributable to the topographic dissimilarity of the three geographical regions.

$$R\&M_{(N)} = 0.0968x^{1.636} \tag{15}$$

$$R\&M_{(NE)} = 0.4274x^{1.442} \tag{16}$$

$$R\&M_{(C)} = 0.1216x^{1.593} \tag{17}$$

Table 4 summarizes the estimated R&M outlay (USD) of the three geographical regions using the conventional curve fitting model. In the table, given the harvest area (x) of 400 hm^2, the estimated R&M costs using conventional curve-fitting model are 1749.2 USD, 2415 USD, and 1698 USD for the northern, northeastern, and central regions, respectively. With 800 hm^2 harvest area, the corresponding R&M costs are 5436 USD, 6562 USD, and 5123.3 USD. The R&M expenditure is positively correlated to the size of harvest area. Figure 13 illustrates the R&M cost estimations in relation to harvest area using the conventional R&M models for the three geographical regions (Equations (15)-(17)).

Table 4 R&M cost estimations using the conventional power-function curve-fitting model

Harvest area (x)/hm^2	R&M Cost Estimation (USD)		
	North (N)	Northeast (NE)	Central (C)
400	1749.2	2415	1698
800	5436.6	6562	5123.3
1 200	10 554	11 776	9773.9
1600	16 879	17 830	15 456
2000	24 342	24 599	22 053

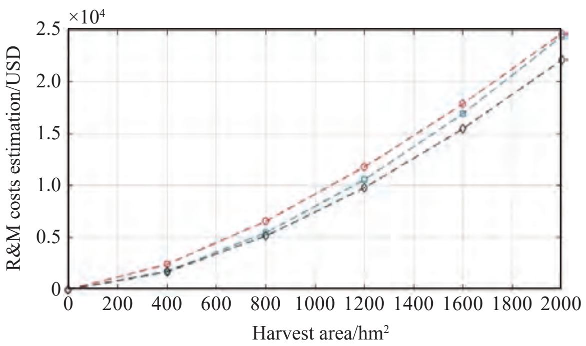


Figure 13 R&M cost estimations relative to harvest area using the conventional power-function curve-fitting R&M cost models of northern, northeastern, and central regions

3.2 R&M cost estimation using the proposed ANN-based model

Table 5 presents the R&M cost estimations of the three geographical regions using the ANN-based estimation model, based on the simulation parameters in Table 3. Given the harvest area (x) of 400 hm^2, the estimated R&M costs using ANN-based model are 1749 USD; 2416 USD; and 1699 USD for the northern, northeastern, and central regions. With 800 hm^2 harvest area, the corresponding R&M costs are USD 5437 USD; 6563 USD; and 5124 USD. The R&M expenditure and harvest area size are positively correlated. The R&M cost estimations relative to harvest area using the ANN-based estimation model by geographical region are shown in Figure 14.

Table 5 R&M cost estimations using the proposed ANN-based model

Harvest area (x)/hm^2	R&M Cost Estimation/USD		
	North (N)	Northeast (NE)	Central (C)
400	1749	2416	1699
800	5437	6563	5124
1200	10 550	11 780	9744
1600	16 900	17 830	15 460
2000	24 340	24 600	22 050

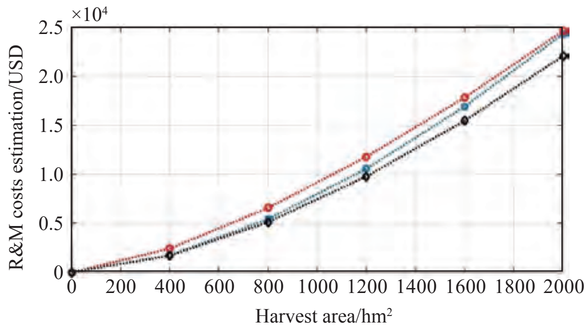


Figure 14 R&M cost estimations relative to harvest area using the proposed ANN-based estimation model of northern, northeastern, and central regions

3.3 Comparison between the conventional and ANN-based R&M cost estimations

To validate the ANN-based estimation model, the R&M cost estimations of the conventional power-function curve-fitting and ANN-based models are compared. The percentage error between R&M cost estimations of the conventional curve-fitting and ANN models is determined by

$$Error = \frac{R\&M_{CONV} - R\&M_{ANN}}{R\&M_{CONV}} \times 100\% \tag{18}$$

where, $R\&M_{CONV}$ and $R\&M_{ANN}$ are the R&M cost estimations of the conventional and ANN-based models.

In Table 6, the error between the conventional curve-fitting and ANN-based R&M cost estimation models of the northern region (N) is between –0.003 80%-0.001 24%. The percentage errors between both estimation models for the Northeast (NE) and Central Plains (C)arebetween0.000 41%-0.000 04%;and–0.003 06%-0.000 584 6 %. The overall error is below 1%, indicating good agreement between the conventional and ANN-based estimation models.

Table 6 Percentage error of R&M cost estimations between conventional power-function curve-fitting and ANN-based models

Harvest area (x) /hm^2	Error of R&M Cost Estimation/%		
	North (N)	Northeast (NE)	Central (C)
400	–0.000 11	0.000 41	0.000 58
800	0.000 07	0.000 15	0.000 13
1200	–0.000 38	0.000 34	–0.003 06
1600	0.001 24	0.000 00	0.000 25
2000	–0.000 08	0.000 04	–0.000 14

Figure 15 compares the R&M outlay estimations of the conventional and ANN-based models of the northern region. The results of both models are in good agreement and consistent with Table 6. Likewise, the R&M expenditure estimations of the Northeast are also in good agreement, as shown in Figure 16. In Figure 17, the estimated R&M costs of both estimation models of the central region are agreeable. In essence, the results validate the applicability of the ANN-based model to estimating the R&M outlays of agricultural machinery, i.e., rice combine harvesters.

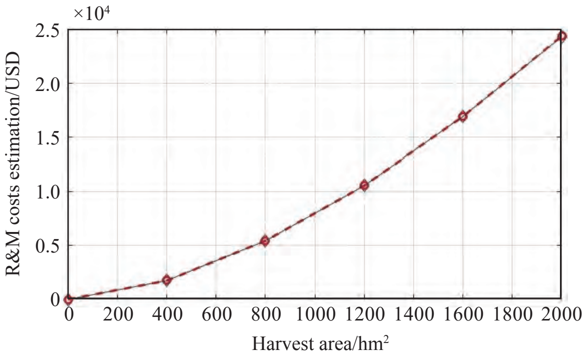


Figure 15 Comparison between R&M cost estimations of conventional and ANN-based models of the north region

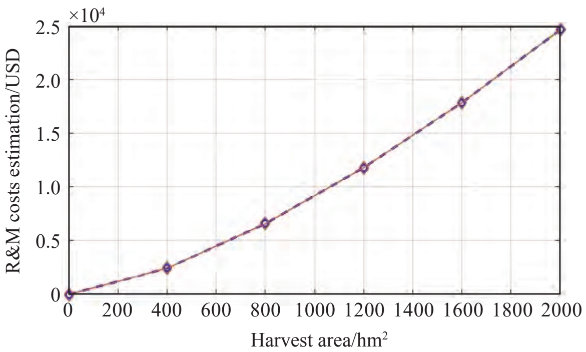


Figure 16 Comparison between R&M cost estimations of conventional and ANN-based models of the northeastern region

The comparison between the estimated R&M costs of both estimation models validates the suitability of the proposed ANN-based estimation model as an alternative to the conventional mathematical estimation model, as evidenced by the percentage error of less than 1%. Unlike the mathematical model which demands a large amount of data, the proposed ANN model requires a substantially smaller amount of data for accurate estimation and thereby lower data-collection budget. Besides, data update is more convenient for the ANN-based estimation model. The proposed ANN model also incorporates curve-fitting coefficients into the model to improve the estimation accuracy. Specifically, the proposed ANN-based model is useful for pricing the service rate of agricultural machinery in an efficient and reasonable fashion.

4 Conclusions

This research proposes an ANN-based R&M cost estimation model for agricultural machinery with high estimation accuracy and small data requirement. In the study, the proposed ANN estimation model is implemented to estimate the R&M costs using a sample of locally-made rice combine harvesters in the rice-growing regions of Thailand. The model inputs are geographical regions, size of harvest area, and curve fitting coefficients related to historical cost data; and the ANN output is the estimated R&M cost. The ANN algorithmic scheme uses multilayer feed-forward and Levenberg-Marquardt backpropagation learning algorithms. To validate, the ANN-based R&M cost estimations are calculated, and results are compared with those of conventional mathematical estimation model. The results show that the percentage error between the conventional and ANN-based estimation models is below 1%, indicating high estimation accuracy of the proposed ANN model. The proposed ANN model requires a substantially smaller amount of data and, with the inclusion of curve-fitting coefficients in the model, can achieve improved estimation accuracy. The ANN-based model is beneficial for pricing the service rate of agricultural machinery in an efficient and reasonable manner. Moreover, with minor modifications, the ANN-based R&M cost estimation model is also applicable to other geographical areas and combine harvesters or tractors of different countries of origin.

Acknowledgements

This research was supported by the Fundamental Fund of Khon Kaen University (KKU). Sincere appreciation goes to the faculty of Engineering, KKU for administrative support.

[References]

[1] Spokas L, Steponavicius D. Fuel consumption during cereal and rape harvesting and methods of its reduction. Journal of Food Agriculture & Environment, 2011; 9: 257–263.

[2] Lips M, Burose F. Repair and maintenance costs for agricultural machines. International Journal of Agricultural Management, 2012; 1: 2047–3710.

[3] Pflueger B. How to calculate machinery ownership and operating costs. SDSU Farm Financial Management Specialist. 2009.

[4] Calcante A, Fontanini L, Mazzetto F. Repair and maintenance costs of 4WD tractors and self propelled combine harvesters in Italy. Journal of Agricultural Engineering, 2013; 44: 141–147.

[5] Calcante A, Fontanini L, Mazzetto F. Repair and Maintenance Costs of 4WD Tractors in Northern Italy. Transactions of the ASABE, 2013b; 56: 356–362.

[6] Sopegno A, Calco A, Berruto R, Busato P, Bocthic D. A web mobile application for agricultural machinery cost analysis. Computers and Electronics in Agriculture, 2016; 130: 158–168.

[7] Ismail M I, El Pebrian D. The characteristics of the repair and maintenance costs distribution of rice combine harvester in Malaysian paddy fields. Agricultural Engineering International, 2019; 20: 132–138.

[8] Rahman A, Latifunnahar M, Alam M M. Financial management for custom hire service of tractor in Bangladesh. Int J Agric & Biol Eng, 2013; 6(3): 28–33.

[9] Zhang Z, Zhang Z H, Wang X M, Liu H, Wang Y J, Wang W J. Models for economic evaluation of multi-purpose apple harvest platform and software development. Int J Agric & Biol Eng, 2019; 12(1): 74–83.

[10] Frank L. Coefficients of repair and maintenance costs for axial and transverse combine harvesters in Argentina. Spanish Journal of Agricultural Research, 2003; 1: 81–97.

[11] ASAE EP496.3. Agricultural machinery management. In: ASAE Standard, 2006; 384–390. Am Soc Agric Biol Eng, St. Joseph, MI, USA.

[12] Lips M. Repair and maintenance costs of hill-farm tractors and transporters for upland mechanization. Journal of Agricultural Engineering, 2013; 44: 133–138.

[13] Masek J, Novak P, Pavlicek T. Evaluation of combine harvester fuel consumption and operation costs. In: Malinovska L, Osaducks V (eds): Proceeding of the 14th International Scientific Conference: Engineering for Rural Development, Jelgava, May 20-22, 2015; 78–83.

[14] Al-Suhaibani SA, Wahby MF. Farm tractor breakdown classification, Journal of the Saudi Society of Agricultural Sciences. Journal of the Saudi Society of Agricultural Sciences, 2017; 16: 294–298.

[15] Rashidi M, Ranibar I. Prediction of repair and maintenance costs of universal 650 tractors based on accumulated usage hours. Agricultural Engineering Research Journal, 2016; 6: 10–14.

[16] Ranjbar I, Rashidi M, Khabbaz BG. Prediction of repair and maintenance

costs of two-wheel drive tractors in Iran. In Proc. of the XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR), Quebec, Canada, Jun. 13-17, 2010; 1–10.
[17] Keshavarzpour F. Prediction of repair and maintenance costs of John Deere 3140 tractors. World Engineering & Applied Sciences Journal, 2011; 2: 11–14.
[18] Khodabakhshian R, Shakeri M. Prediction of repair and maintenance costs of farm tractors by using of preventive maintenance. International Journal of Agricultural Sciences, 2011; (3): 39–44.
[19] Pishbin S. Using mathematics modeling to estimate repair costs of tractors. Advances in Environmental Biology, 2014; 8: 646–651.
[20] Dahab M H, Mohamed A O, Kheiry A N O. Repair and maintenance costs estimation as affected by hours of use and age of agricultural tractor in New Halfa Area-Sudan. International Journal of Agriculture Innovations and Research, 2016; 4: 824–830.
[21] Gautam AK, Shrivastava AK. Development of mathematical model for repair and maintenance of some of the farm tractors of JNKVV, Jabalour, India. Advances in Crop Science and Technology, 2017; 5: 228–3.
[22] Ranjbar I, Abbaspour-Fard M H, Ajabshir Y, Rohani A. Prediction of John Deere Tractor Repair and Maintenance Costs by Using Two Different Structures of MLP Artificial Neural Network. Journal of Development and Agricultural Economics, 2009; 22: 86–97.
[23] Rohani A, Abbaspour-Fard M H, Abdolahpour S. Prediction of tractor repair and maintenance costs using artificial neural network. Expert Systems with Applications, 2011; 38: 8999–9007.
[24] Azim H, Farshid K, Nasser SP. Machinery cost prediction based on a new neural network method. Journal of Basic and Applied Scientific Research, 2013; 3: 484–491.
[25] Kumar U A. Comparison of Neural Networks and Regression Analysis: A New Insight. Expert Systems with Applications, 2005; 29: 424–430.
[26] Nawi N M, Khan A, Rehman M Z. A New Levenberg Marquardt Based Back Propagation Algorithm Trained with Cuckoo Search. The 4th International Conference on Electrical Engineering and Informatics (ICEEI 2013).
[27] Sapna S, Tamilarasi A, Kumar M P. Backpropagation learning algorithm. Based on levenberg marquardt algorithm. Computer Science & Information Technology (CS & IT), 2012; 393–398.
[28] Lavrakas P J. Encyclopedia of Survey Research Methods. Thousand Oaks, California, USA: SAGE Publications, Inc. Press. 2008. pp. 196-197.
[29] Persistence Market Research firm (PMR). (2018): Combine Harvester Market: Global Industry Analysis 2013-2017 and Forecast 2018-2026. Avaible at https:// www.persistencemarketresearch.com/ samples/20209. Accessed on [2019-08-12].
[30] Gupta N. Artificial neural network. Network Complex Syst, 2013; 3: 24–28.