RESEARCH ARTICLE

The evaluation method of livestock manure resource utilization management model based on random forest algorithm

Buyu Wang*, Jialei Du, Youxin Yu

College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, China

Received: September 21, 2023; **accepted:** November 16, 2023.

The utilization of livestock and poultry manure resources has become a crucial aspect of agricultural development due to the growth of the social economy and increased environmental protection awareness. However, at present, the resource of utilization effect of livestock manure is insufficient, resulting in a large amount of livestock manure resources wasted. In order to analyze the evaluation effect of the management model on the resource utilization of livestock and poultry manure, this study took livestock and poultry manure resources as an independent variable and used the management model and random forest algorithm as the dependent variables. IBM SPSS software was employed to conduct various analysis methods to evaluate the connection between variables to assess the outcomes of the management model in utilizing livestock manure resources. The Pearson coefficient results showed that the correlation coefficients between the independent variable livestock manure and the utilization management model and the random forest algorithm were -0.067 and 0.275, respectively. In addition, the reliability analysis results showed that the reliability values of standard variance, actual variance, and error variance were 0.751, 0.058, and 0.693, respectively. The results suggested that the use of a management model based on the random forest algorithm positively impacted the evaluation of livestock and poultry manure resource utilization. This model could be utilized to improve the effectiveness of using these resources.

Keywords: livestock manure resource; utilization management model; random forest algorithm.

***Corresponding author:** Buyu Wang, College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China. Email: bywang08@126.com.

Introduction

Resource utilization of animal manure is an important task in agricultural production, and its importance is mainly reflected in the following aspects including (1) animal manure has the potential to function as an organic fertilizer resource, delivering the necessary nutrients essential for crop growth, and augmenting crop yield and quality [1, 2]; (2) livestock and poultry manure contains a large number of organic substances and microorganisms, which can improve soil structure, improve soil fertility, and promote crop growth [3]; (3)the usage of livestock and poultry waste as a resource can mitigate environmental pollution, minimize the impacts of agricultural production on the environment, and safeguard the ecological environment [4]; (4) the use of livestock and poultry manure as a resource can enhance agricultural productivity, decrease production expenses, and encourage sustainable agricultural growth. Research has found that livestock manure is a valuable resource that, if handled properly, can be a good alternative to artificial fertilizers. To avoid environmental pollution, it is essential to treat animal waste properly to reduce odors, nutrient loss, and emissions [5]. In addition, numerous studies have found that poor manure management can have a negative impact on Asia's rapidly growing livestock industry. Lasser *et al*. found that waterways were often contaminated with improperly disposed fecal matter. High concentrations of nutrients in rivers and lakes could cause ecological imbalances as it promoted abnormal growth of algae and aquatic plants [6]. In addition, Hernandez's team found that different pathogens were prevalent in livestock manure and could be direct discharged into streams and infiltration into groundwater posing considerable health risks to humans and animals [7]. A comprehensive assessment of existing manure management strategies is needed to develop new strategies that local livestock producers can adopt.

To enhance management and utilization of livestock manure, a simulation model has been suggested to measure dynamic changes in pasture ecosystem, animal productivity, and ecosystem dynamics, although its technical complexity demands high proficiency [8]. A unique feature of the model is its ability to reflect seasonal and interannual temporal fluctuations in feed quality and quantity, as well as the possible effects of these changes on energy and nutrient intake of livestock [9]. Other studies also found that resource utilization management model could be used to manage and utilize livestock and poultry manure [10]. Resource utilization management model is a mathematical model used to describe and predict the process of resource utilization, which typically involves the input, output, transformation, and consumption of resources. This approach can aid managers in developing optimized resource utilization plans, enhancing resource utilization efficiency, and safeguarding the environment by predicting resource utilization efficiency and environmental impact [11]. The resource utilization management model is widely applied in agricultural production, environmental protection, energy utilization, and other fields. It also can help managers to optimize resource utilization plans and improve resource utilization efficiency [12]. Yang used the resource utilization management model to evaluate the efficiency of water resource utilization in China. The evaluation results led to a proposal of strategies aimed at enhancing utilization efficiency of water resources, thereby improving the overall water resource utilization efficiency [13]. In addition, in order to improve the efficiency of waste management and utilization, the scientists proposed the use of resource utilization management model to manage waste, so as to improve the sustainable use of waste [14]. At present, the resource utilization management model has been applied to the management of livestock and poultry manure. However, traditional management methods struggle to address the complexity and diversity of such manure [15]. The random forest algorithm is an ensemble learning method, which improves the accuracy and stability of prediction by constructing and integrating multiple decision trees [16]. This approach has numerous benefits including a strong generalization ability and the capacity to effectively avoid overfitting issues, which make it a popular choice for medical diagnosis, financial risk assessment, environmental monitoring, and other applications [17]. In the management model of livestock manure resource utilization, random forest can be used to predict the effect of different management methods, so as to help managers choose the best management plan.

Livestock manure has genuine fertilizer value, which leads many farmers, horticulturists, and gardeners to utilize it for crop fertilization. Treated manure is stored in open ponds or cesspools, while untreated manure is discharged into fields as manure. Animal manure provides a significant source of agricultural nutrients including nitrogen, which contributes to the emission of N_2O into the soil. As livestock numbers increase to meet human demand for animal products, managing manure to reduce N_2O emissions become critical. However, challenges still exist in the management of resource utilization of livestock and poultry manure including low resource utilization efficiency and serious environmental pollution. This study aims to improve resource utilization efficiency, reduce environmental pollution, and promote sustainable agricultural development through the management model of livestock manure resource utilization based on the random forest algorithm. The results of this study would provide a scientific based new evaluation method for the resource utilization management model of livestock and poultry manure to improve the efficiency and effect of the resource utilization of livestock and poultry manure and to provide reference for model evaluation in other fields. The random forest algorithm-based model evaluation method should have high practicality and application value and could assess not only livestock manure resource utilization management models, but also models in other fields.

Materials and Methods

Data collection

This study was based on original data analysis results provided by different participants within the region of Inner Mongolia, China. The study included 150 participants from small family farms, 45 from medium farms, 30 from large farms, and 15 from livestock-related businesses and research institutions. The geographical distribution of participants covered different climatic conditions and agricultural production environments to ensure the breadth and applicability of the study. Through the data collection, the participants' education level, livestock breeding experience, and other factors were also included to ensure the diversity and accuracy of the data. The information provided by the participants in this study mainly included five different types including (1) the composition of feces including the content of organic substances, nitrogen, phosphorus, potassium, and other nutrient elements as well as the types and contents of harmful substances such as hydrogen sulfide and ammonia; (2) the manure treatment method, which included the manure treatment technology and process such as aerobic fermentation, anaerobic fermentation, drying treatment, *etc*.; (3) the resource utilization methods, which mainly included the conversion of manure into organic fertilizer, inorganic fertilizer, biological gas, and other different resource utilization methods as well as the quality, output, economic benefits, and other information after conversion into these resources; (4) the experience and knowledge of the participants on the recycling of livestock and poultry waste as well as the difficulties and challenges in implementing the recycling process; (5) the impact of waste treatment and resource utilization on the environment such as the impact on air, water, soil, and other environmental factors. Through the collection and analysis of these information, we could better understand the resource utilization effect of livestock and poultry manure and provided references for improving the resource utilization efficiency of manure and protecting the environment.

Data processing

All data were processed by using SPSS software (IBM, Armonk, New York, USA) for descriptive statistics, correlation coefficient, regression analysis, one-way ANOVA, graph analysis, T-test, reliability analysis, and control charts related to the variables. Livestock manure resources were set as independent variables, while random forest algorithm and management model were dependent variables. The specific correlation between the three variables was examined.

Utilization management model

The resource utilization management model is a mathematical model used to describe and predict the process of resource utilization, which specifically accounts for the input, output, transformation, and consumption of resources. By considering the environmental impact and efficiency of resource utilization, managers can optimize plans and improve resource utilization while also protecting the environment. The resource utilization management model can be divided into production model, allocation model, use model, and resource management model. Each model has its specific application scenario and parameter settings. Through the establishment and application of resource utilization management model, the process of resource utilization can be better understood and managed, leading to sustainable resource utilization and promoting the sustainable development of society and the economy. The process of establishing the utilization management model in this study was included several steps below. First, the goal and scope of the model were defined. The data related to resource utilization was then collected and analyzed. Based on data analysis, problems and improvement opportunities in resource utilization were identified. Then the specific strategies and action plans for the identified problems and opportunities were formulated by ensuring that the strategy was aligned with the organization's strategic objectives and values. The strategies and action plans developed were then put into practice. After implementing improvement measures, the changes in resource utilization were continuously monitored and evaluated, which provided feedback and adjustments to the model based on the results of monitoring and evaluation. It might involve fine-tuning the policy, introducing new improvements, and so on. It was also ensured that the model would continue to adapt to the changing and evolving needs of the organization. An established utilization management model would help organizations to better manage and optimize resource usage, improve efficiency, and reduce costs.

Random forest algorithm

A random forest algorithm is an ensemble learning algorithm that uses multiple decision trees to predict outcomes. Each decision tree is built based on different training data sets and feature subsets, so there is some independence between them. The random forest algorithm can get the final prediction result by averaging or voting the results of multiple decision trees, thus improving the accuracy and stability of prediction. The scikit-learn (https://scikit-learn.org/stable/index.html), a popular machine learning library available in Python and containing a wealth of machine learning features, was employed to implement random forest algorithm in this study. Because the development and maintenance of the scikit-learn package is a global activity, its code can be used in many cities and countries around the world. The general steps of the development process of random forest algorithm were as follows. First, the relevant data of manure resources and management and utilization mode including the type, quantity, quality, and other information of manure resources provided by participants were collected. All collected data were pre-processed to ensure the quality and usability. Then the features related to the manure resource and management and utilization model from the pre-treated data were selected to reduce the complexity of the model and improve the forecasting ability of the model. Thet random forest algorithm was then used to train the selected features and establish the relationship model between the manure resources and the management and utilization model. Finally, the established model was evaluated and used to predict the new manure resources and management and utilization model, thus optimizing the utilization efficiency and management effect of manure resources.

Results and discussion

Descriptive statistical analysis

The descriptive statistical analysis describes the overall results of the study, including the number of observations, the minimum, the maximum, the mean of each variable, and the standard deviation. In this study, the livestock manure resource acted as an independent variable with a minimum value of 1.000 and maximum value of 5.000, and mean and standard deviation of 1.82 and 0.97, respectively. The utilization management model was a dependent variable with a minimum and maximum value of 1.00 and 4.00, while the mean and standard deviation were 1.9900 and 0.79766, respectively. Another

Table 1. Pearson correlation coefficient among variables.

	Livestock manure resource	Utilization management model	Random forest algorithm
Livestock manure resource	1	-0.067	0.275**
	-	0.507	0.006
	100	100	100
Utilization management model	-0.067	1	0.014
	0.507	-	0.893
	100	100	100
Random forest algorithm	0.275**	0.014	1
	0.006	0.893	-
	100	100	100

Note: **: Correlation was significant at the $P < 0.01$ (2-tailed).

dependent variable was the random forest algorithm, which had a minimum value of 1.00, a maximum value of 4.00, a mean value of 1.8700, and a standard deviation of 0.81. The minimum value and maximum value are two important indicators used to describe the distribution range of data in descriptive statistical analysis. The minimum value stands for the smallest value in the data set, signifying the lower limit of data distribution, while the maximum value is the largest value in the data set, representing the upper limit of the data distribution [18]. Minimum and maximum values can help to understand the distribution range of the data and thus better understand the characteristics and properties of the data [19]. In addition, a statistical comparison was conducted between the mean values of the animal manure resource group, the utilization management model group, and the random forest algorithm group. The mean value of the animal manure resource group was found to be significantly different from that of the utilization management model group, with a *P* value less than 0.05 indicating statistical significance. The *P* value between the animal manure resource group and the random forest algorithm group was also less than 0.05, indicating a statistically significant. There was a significant difference between the management model group and the random forest algorithm group ($P < 0.05$).

The correlation coefficients between variables, including dependent and independent variables were described in Table 1, which included the Pearson correlation value, the significant levels, and the number of observations. It was noteworthy that the livestock manure, considered an independent variable, exhibited a negative relation with the utilization management model with a rate of -0.067 revealing an insignificant correlation between the two. However, a positive and significant relation between the random forest algorithm model and livestock manure with the rate level of 0.275 and significant level of 0.006 were observed, which showed 100% significant level with each other. The result demonstrated that the utilization management model was another independent variable, which showed the positive relationship with the random forest algorithm rate of 0.014.

ANOVA analysis of single factor

The variance analysis of the management model and the random forest algorithm was carried out in the inter-group and intra-group, respectively. The results showed that the sum of squares between the groups of the utilization management model was 2.534, while the sum of squares within the groups was 60.456 with the mean squares of 0.633 and 0.636, respectively. The F value of utilization management model was 0.995, which indicated a positive correlation between inter- and intra-group of utilization management model with a significance level of 0.414. The random forest algorithm demonstrated as a dependent variable with a

Table 2. The total variance interpretation values for each component using principal component analysis (PCA).

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %
1	1.281	42.685	42.685	1.281	42.685	42.685
2	1.006	33.541	76.225	1.006	33.541	76.225
3	0.713	23.775	100.000			

Note: The components 1, 2, and 3 were livestock manure group, utilization management model group, and random forest algorithm group, respectively.

sum of squares reaching 9.401 and 55.909 in inter- and intra-group, respectively, an F value of 3.994, and a significance level of 0.005, indicating a significant difference. In addition, the model summary results showed that the R value of the model was 0.277, the R variance was 0.077, the adjusted R variance was 0.058, and the standard error of the estimated value was 0.78839. All results suggested that model summary was related to dependent and independent variables. A variance analysis was performed for regression, residual, and total value with the sum of squares values of 5.018, 60.292, and 65.310, respectively. The results also demonstrated the mean square value of regression as 2.509, the residual value of 0.622, F value of 4.037, the significance value of 0.021, and the significance level as 2%.

Regression analysis between variables

Regression analysis was performed on the independent and dependent variables. The beta value of livestock manure resources was 0.230, the standard error was 0.081, and the T-test result of $P < 0.01$, indicating a very significant difference. On the other hand, the beta value for the utilization management model was 0.033 with a standard error of 0.100 and a standardization coefficient of 0.032. The T-test value was 0.329 with the significance level of 7%. The total variance interpretation values for each component was shown in Table 2. The results showed the variance percentage and cumulative percentage for the initial eigenvalues and the extracted sum of squares values. The results demonstrated that the livestock manure group had a total principal component analysis (PCA) value of 1.281, while the management mode group and random forest algorithm group had total PCA values of 1.006 and 0.713, respectively. The variance percentages for the different groups of livestock manure, utilization management model, and random forest algorithm were 42.685, 33.541, and 23.775, respectively, while the cumulative percentage values were 42.685, 76.225, and 100.00, respectively, indicating the initial characteristic value of each component. The results also showed that the variance percentage of the extracted square values were 42.685 and 33.541, and the cumulative percentage values were 42.685 and 76.225 for livestock manure and utilization management model groups, respectively. The component matrix values of the animal manure resource group, the utilization management model group, and the random forest algorithm group were further analyzed. The value of the component matrix represented the contribution of the component to each variable. The results showed that the values of matrix component 1 were 0.804, -0.154, and 0.781 for livestock manure, utilization management model, and random forest algorithm groups, respectively, indicating that the contribution degree of each component to the total variance was 80.4%, -15.4%, and 78.1%, respectively. The component matrix values 2 were -0.043, 0.974, and 0.236 for each group, respectively, indicating that the contribution degree of each component to the total variance was -4.3%, 97.4%, and 23.6%, respectively. The chi-square test is a statistical technique commonly applied to verify the independence between two or more categorical variables [20]. The larger the chi-square value, the less independence between the two variables, indicating a correlation between variables. On the other hand, the smaller the chi-square value, the greater the independence between the two variables, indicating no correlation between them [21]. The chi-square values of livestock manure, utilization management model, and random forest algorithm groups were 76.5, 35.12, and 36.4, respectively. Therefore, the overall significance value was 0, indicating the significant level of correlation between the groups.

Table 3. Reliability test.

standard variance	0.751
True variance	0.058
Error variance	0.693
Common Inter-Item Correlation	0.078
Scale reliability	0.202
Scale reliability (Unbiased)	0.218

Reliability and analysis of single sample

Reliability tests were carried out in the study. Table 3 showed the statistical analysis of reliability for different test results. The standard variance value was 0.751, while the actual variance value was 0.058 and the error variance reliability value was 0.693. The results suggested that the scale reliability was 0.202 and unbiased scale reliability was 0.218, thereby establishing the overall findings as reliable for analysis. Table 4 demonstrated a one-sample test analysis of each variable including the dependent and independent variables. The results demonstrated the t value, df, significance, mean differences, and 95% confidence intervals. The livestock and poultry manure resources group was independent variable with t value of 18.603, significance level of 0.000, and mean difference value of 1.820. The management model was used as the dependent variable with the t value of 24.948, the significance level of 0.000, and the mean difference of 1.990. Another dependent variable, the random forest algorithm, demonstrated the t value, significance level, and mean difference of 24.948, 0.000, and 1.990, respectively. The results indicated that there were significant differences between the animal manure resources group, the utilization management model group, and the random forest algorithm group.

Conclusion

This study examined the method for evaluating the management model of livestock manure resource utilization based on the random forest algorithm. The relationship between the management model of animal manure resource utilization and the random forest algorithm and the evaluation effect of the model were analyzed. The correlation coefficients between the independent variable, livestock manure, and the utilization management mode and the random forest algorithm were -0.067 and 0.275, respectively, indicating that the utilization management mode and the random forest algorithm had different effects on the resource utilization of livestock manure. Meanwhile, the reliability analysis showed that the reliability values of standard variance, actual variance, and error variance were 0.751, 0.058, and 0.693, respectively, indicating that the prediction results of the model had high stability and reliability. In summary, the resource management model for livestock manure utilization, built with the random forest algorithm, had significant practicality and value for application. This study provided a scientific basis for the resource utilization and management of livestock and poultry manures. Through analysis of the relationship between the management model of livestock manure resource utilization and the random forest algorithm, the factors affecting the resource utilization of livestock manure could be better understood, and scientific decision basis could be provided for managers. In the future, the model will be further studied and optimized to improve the prediction accuracy and stability of the model and provide strong support for maximizing the utilization of resources and environmental protection. At the same time, we will also explore the application of other machine learning methods in the resource utilization management

Table 4. One-sample test.

	Test value = 0					
	T	df	Significance (2-tailed)	Mean difference	95% Confidence interval of the difference	
					Lower	Upper
Livestock manure resource	18.603	99	0.000	1.82000	1.6259	2.0141
Utilization management model	24.948	99	0.000	1.99000	1.8317	2.1483
Random forest algorithm	23.023	99	0.000	1.87000	1.7088	2.0312

of livestock and poultry manures, in order to find a more effective and accurate model evaluation method.

Acknowledgement

This work was supported by the Program for Improving the Scientific Research Ability of Youth Teachers of Inner Mongolia Agricultural University (Grant number: BR220148), the Science and Technology Major Project of Inner Mongolia (Grant number: 2020ZD0004).

References

1. Dong H, Mangino J, McAllister T, Hatfield JL, Johnson DE, Lassey KR, *et al*. 2006. Emissions from livestock and manure management. Embrapa Meio Ambiente-Capítulo em livro científico (ALICE). 4:1-89.
2. Li J, Akdeniz N, Kim HHM, Gates RS, Wang X, Wang K. 2021. Quantification of sustainable animal manure utilization strategies in Hangzhou, China. Agr Syst. 191:103150.
3. Hieu Vu N, Lambertz C, Gauly M. 2016. Factors influencing milk yield, quality and revenue of dairy farms in southern Vietnam. Asian J Anim Sci. 10:290-299.
4. Cockburn M. 2020. Review: Application and prospective discussion of machine learning for the management of dairy farms. Animals-Basel. 10:1690.
5. Bobbo T, Biffani S, Taccioli C, Penasa M. 2021. Comparison of machine learning methods to predict udder health status based on somatic cell counts in dairy cows. Sci Rep-UK. 1:59.
6. Lasser J, Matzhold C, Egger-Danner C. Fuerst-Waltl B. 2021. We integrate diverse data sources to predict disease risk in dairy cattle—a machine learning approach. J Anim Sci. 99:294.
7. Hernández BC, Lopez-Villalobos N, Vignes M. 2021. Identifying health status in grazing dairy cows from milk mid-infrared spectroscopy using machine learning methods. Animals-Basel. 11:2154
8. Habeeb AAM, Gad AE, Atta MAA. 2016. Changes in body weight gain and blood hormonal levels in relation to change in age of egyptian male buffaloes calves from birthing to puberty. Adv Appl Physiol. 1:43-48.
9. Han XP, Hubbert B, Hubbert ME, Reinhardt CD. 2016. Overview of the beef cattle industry in China: The widening deficit between demand and output in a vicious circle. J Fish Livest Prod. 4:190.
10. Velez FF, Colman S, Kauffman L, Ruetsch C, Anastassopoulos K. 2021. Real-world reduction in healthcare resource utilization following treatment of opioid use disorder with reSET-O, a novel prescription digital therapeutic. Expert Rev Pharm Out. 21(1):69-76.
11. Deng P, Xu G, Bing J, Xu C, Jia J. 2020. Evaluation method of rain–flood resource utilization availability and its application in the Hanjiang River Basin. Water Supply. 20(8):3557-3575.
12. Kumar RG, Kesinger MR, Juengst SB, Brooks MM, Fabio A, Dams-O'Connor K, *et al*. 2020. Effects of hospital-acquired pneumonia on long-term recovery and hospital resource utilization following moderate to severe traumatic brain injury. J Trauma Acute Care. 88(4):491-500.
13. Yang Y. 2021. Evaluation of China's water-resource utilization efficiency based on a DEA-Tobit two-stage model. Water Supply. 21(4):1764-1777.
14. Nižetić S, Djilali N, Papadopoulos A, Rodrigues JJ. 2019. Smart technologies for promotion of energy efficiency, utilization of sustainable resources and waste management. J Clean Prod. 231(5):565-591.
15. Zhang J, Liu J, Evrendilek F, Zhang X, Buyukada M. 2019. TG-FTIR and Py-GC/MS analyses of pyrolysis behaviors and products of cattle manure in CO_2 and N_2 atmospheres: kinetic, thermodynamic, and machine-learning models. Energ Convers Manage. 195(6):346-359.
16. Schonlau M, Zou RY. 2020. The random forest algorithm for statistical learning. The Stata Journal: Promoting communications on statistics and Stata. 20(1):3-29.
17. Abdulkareem NM, Abdulazeez AM. 2021. Machine learning classification based on Radom Forest Algorithm: A review. Int J Sci Bus. 5(2):128-142.
18. Hammond KJ, Crompton LA, Bannink A, Dijkstra J. 2016. Review of current in vivo measurement techniques for quantifying enteric methane emission from ruminants. Anim Feed Sci Technol. 219:13-30

19. Gwaza DS, Momoh OM. 2016. Endangered indigenous cattle breeds of Nigeria a case for their conservation and management. World Sci News. 30:68-88.
20. Talha M, Azeem S, Sohail M, Javed A. 2020. Mediating effects of reflexivity of top management team between team processes and decision performance. Azerbaijan J Educ Stud. 1:105-119.
21. Wang G, Hua L, Squires V. 2017. Development impacts on beef and mutton production from the pastoral and agro-pastoral systems in China and the economic and cultural factors that influence it. Livest Res Rural Dev. 29(10):1-19.