International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(2): 488-503

Original Article | ISSN (0): 2582-631X

DOI: 10.47857/irjms.2025.v06i02.03126

An Optimal Fuzzy System for Nitrogen Fertilizer Recommendation Using Probability Density Function Coupled with Whale Optimization Algorithm

Uditendu Sarkar¹, Gouravmoy Banerjee², Indrajit Ghosh^{3*}

¹National Informatics Centre, Ministry of Electronics and Information Technology, Government of India, West Bengal State Centre, Kolkata, West Bengal, India, ²Eklavya Model Residential School, Swayem, Namok, Mangan, Sikkim, India, ³Department of Computer Science, Agro-Computing Research Laboratory, Ananda Chandra College, Jalpaiguri, West Bengal, India. *Corresponding Author's Email: ighosh2002@gmail.com

Abstract

In most countries, the site-specific, precise application of N fertilizer in an optimal dose is one of the most challenging tasks for sustainable agriculture. Several recommendation systems have been proposed for N fertilizer as an alternative to the scarce and expensive soil experts. However, none of them exhibited an impressive performance. In this article, we have proposed a highly efficient optimal fuzzy system (OFS) with a novel architecture to recommend crop-specific optimal doses of N fertilizer based on site-specific soil and climatic data. In our proposed OFS, the fuzzy membership functions of each variable were replaced by the respective probability density functions, the rule base was redefined, and finally, conflict resolution was achieved using the probabilistic fuzzy logic controller approach. The output probability density function was optimized with the most popular whale optimization algorithm (WOA). Such an innovative approach to designing an N fertilizer recommendation system has never been well thought out so far. Our designed OFS was empirically validated in terms of four statistical metrics: the co-efficient of determination (R^2), the Nash-Sutcliffe efficiency (*NSE*), the root means squared error (*RMSE*), and the mean absolute error (*MAE*) against three varieties of paddy and two varieties of potato cultivated in the Gangetic alluvial plain in West Bengal, India. It was further compared to the other latest N fertilizer recommendation systems designed for different crops worldwide. The study revealed that our system (with R^2 ranging from 0.9628 to 0.9880) outperformed all other systems (with R^2 ranging from 0.1900 to 0.8400).

Keywords: Fertilizer Recommendation System, Fuzzy Hybrid Systems, Nitrogen Fertilizer, Optimal Fuzzy System, Probability Density Function, Whale Optimization Algorithm.

Introduction

Soil nutrients provide the proper nourishment for the crops, leading to healthy and vigorous plant growth. Plants can reach their full potential and produce higher yields when adequate nutrients are available in the soil. Soil with nutrient deficiencies leads to stunted plant growth and lower crop productivity. Therefore, maintaining an adequate quantity of soil nutrients is crucial for improving crop productivity. However, over time, subsequent cultivation or inefficient fertilizer management strategies deplete the soil's nutrients. As a remedy, various chemical fertilizers are applied to the soil to restore the soil's nutritional level for better crop yield. Nitrogen (N) is the foremost nutrient that boosts plant development and promotes yield growth (1). N deficiency slows crop growth, decreases the harvest index, and ultimately leads to a reduction

in global production (2). On the other hand, the increasing application of N induces many serious environmental disorders, such as water eutrophication, inhalable particulate matter formation, etc., along with an adverse effect on farmers' health (3, 4). An adequate quantity of N fertilizer in the soil promotes strong vegetative development and healthy leaves. The site-specific precise application of N fertilizer is the only strategy capable of optimizing the overall use of N (5). Therefore, recommending the optimal dose of N fertilizers is very significant for growing healthy crops, increasing yields, and maintaining a sustainable environment (6). However, in most countries, due to the scarcity of soil experts and a lack of adequate knowledge, farmers apply N fertilizers at a higher dose than required, with the misconception that increased application of N

This is an Open Access article distributed under the terms of the Creative Commons Attribution CC BY license (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(Received 08th November 2024; Accepted 15th April 2025; Published 30th April 2025)

increases crop yields, which affects the sustainability of the agroecosystem and raises production costs (7). The earlier Ν recommendation approaches were based on soil N testing and plant tissue testing (8). Other physical and analytical methods have also been proposed in the literature (9). These methods required chemical analysis of the soil samples, comprehensive and precise knowledge of soil science, and extensive mathematical calculations, which are quite challenging tasks for rural smallholder cultivators. Several agricultural simulation models (ASMs) were recognized as useful tools in agricultural research to develop alternative strategies for fertility management and to quantify the relationships between crops, management practices, and the environment (10, 11). The simulation precision of ASMs were further improved using remote sensing data for in-season crops and made them efficient for recommending N fertilizers (12). However, these models are plagued by some serious issues like data acquisition, assimilation algorithms, and model accuracy (13). As an alternative, researchers have developed various N fertilizer recommendation systems that heavily rely on machine learning (ML) strategies, providing a cutting-edge approach to handling challenging linear or non-linear problems. The application of ML techniques for recommending fertilizers was found to be highly viable for growing crop yields in corn production in China (14). Several MLbased models were suggested for the N fertilizer recommendation, but most of them were designed using decision tree-based regressors. A simple regression model was proposed to recommend an economically optimum rate of N for corn using soil data and climatic parameters (15). Four regression models: linear regressor (LR), ridge regressor (RR), least absolute shrinkage and selection operator (LASSO) regressor, and gradient-boosted tree regressor (GBT) were designed for corn N fertilizer recommendation (16). Bayesian regressor (BR) and random forest (RF) regressor models were suggested for the recommendation of side dressing N rates in corn (17). For rice, two regressors, RF and support vector machine (SVM), were used for modelling in-season N topdressing rate recommendation (18). A comparative study was carried out with four ML

techniques: SVM, RF, GBT, and RR, for a sitespecific recommendation of economically optimal N for canola (19). To select the most suitable model for recommending an economically optimum rate of N for corn, a performance analysis was conducted with eight regressor models, where the decision tree-based model was empirically found to be the best (20). To date, the highest accuracy was reported using an RF regressor with $R^2 = 0.8400$ against corn (20). However, there are two significant shortcomings of prior methodologies: such regressor-based systems require a large amount of data, which is challenging to acquire in a timely manner, and the results are very difficult to interpret, and their accuracy is comparatively very low (21). On the contrary, fuzzy logic systems are the most comprehensive and flexible approach to use in the context of predicting the economically optimal rate of N fertilizer (22). However, a few systems were reported that employ fuzzy systems for recommending N fertilizer. A fuzzy system was suggested for N recommendation in Greece, where the sensitivity of the system against different input parameters was studied but did not report the accuracy of the system (23). Another system was designed for recommending the spatial variable application of N using soil, precipitation, and plant data (24). A fuzzy system was implemented that used soil properties such as electrical conductivity and soil organic matter as inputs (25). For recommending the optimum dose of N fertilizer for paddy fields in Malaysia, a fuzzy system was proposed that used aerial images captured by drones, and the input variables were suggested by the agricultural experts (26). To determine the land-specific exact fertilizer needs of eight crops in two agro-climatic zones in India, a fuzzy decision support system was developed to improve crop productivity with minimum consumption of fertilizer using exhaustive field measurements and laboratory analysis (27). A fuzzy logic control approach was used to design a fuzzy system to estimate nutrient levels and application timing for various growth stages of mango trees in Malaysia (28). As advancement, an improved version of the fuzzy system was developed using the type-II fuzzy set, which deals with fertilization by primary

fertilizers and is applied to wheat cultivation in

Pakistan (29). In addition to their few benefits,

the major drawback of such fuzzy systems is the need for trial and error in designing the appropriate membership functions, which results in lower precision due to non-optimal output (30). To achieve better accuracy with limited sitespecific data, the adoption of alternative strategies for designing an innovative fuzzy recommendation system for N fertilizer is highly significant.

The aim of this paper is to propose a novel architecture of an optimal fuzzy system (OFS) to design an N fertilizer recommendation system with better accuracy using soil and climatic parameters. In our proposed system, the output fuzzy membership functions of a classical fuzzy system have been redefined in terms of probability density functions (p.d.f.) hybridized with the whale optimization algorithm (WOA) as the most popular nature-inspired optimization technique. Such an innovative approach has not been well thought out until now to achieve the optimal output of a fuzzy system. The performance of our designed system was empirically validated using four well-known performance metrics: the co-efficient of determination (R^2) , the Nash-Sutcliffe efficiency (NSE), the root mean squared error (RMSE), and the mean absolute error (MAE) against three paddy varieties and two potato varieties grown in the Gangetic alluvial plain in West Bengal, India. It was further compared to the other latest N fertilizer recommendation systems designed for other crops worldwide. The study revealed that our system outperformed all other systems in terms of R^2 (ranging from 0.9628 to 0.9880).

Methodology Theoretical Foundations

Since our proposed OFS is an enhancement of the classical fuzzy system, the preliminary concepts of the classical fuzzy set and system are described for a short overview. Fuzzy set theory was first introduced by L. A. Zadeh to handle uncertainties in real-world and it is a precise logic of imprecision and approximate reasoning (31). It is the most comprehensive and flexible method to incorporate expert knowledge into a recommendation system and allows to implement the human reasoning in computers and makes precise inferences or predictions based on imprecise data, particularly where the mathematical model of a process does

not exist. In fuzzy set theory, a parameter whose values are represented in linguistic terms is called a linguistic variable, and the linguistic terms are called fuzzy sets. A linguistic variable (e.g., temperature) is a word in a natural language and its attributes (e.g. low, medium, high) are a set of words called linguistic values. Each parameter (input and output) is given membership in several fuzzy sets. The membership functions are used in the fuzzy system for the appropriate transfer of numerical variables to linguistic variables. They specify the degree of membership of a variable in a fuzzy set. Membership close to zero implies weak membership, and a membership close to one implies strong membership of x in fuzzy set F. Triangular, trapezoidal, sigmoid, Gaussian, bell, and other functions are widely used in the design of fuzzy membership functions. The choices for these functions are context-dependent and subjective in nature. The classical fuzzy system consists of four basic components, a fuzzifier, the rule (knowledge) base, a decision-making logic (inference engine), and a defuzzifier (32). The first step in developing a fuzzy system is to construct appropriate fuzzy sets for the input and output variables. The fuzzifier converts the crisp inputs into membership values in all fuzzy sets. Because membership functions overlap, nonzero membership values exist in more than one fuzzy set. The rule base consists of a predefined set of ifthen rules dictating the outcome(s). The rules serve the purpose of mapping the input fuzzy sets to the output fuzzy sets. As a consequence of overlap in the fuzzy sets, multiple rules are likely to be activated by the same set of input variables. When more than one rule is activated, the decisionmaking logic resolves conflicts between rules and constructs the resulting output membership function, representing the decision to be taken. The strengths of the activated rules and the membership values of the output variables in the output fuzzy sets are used to construct the output membership function (33).

Our Proposed Optimal Fuzzy System

The classical fuzzy systems suffer from some major limitations (30). To overcome these shortcomings, we have proposed a novel architecture of an advanced fuzzy system, where the probabilistic view of fuzzy methods has been incorporated. The vagueness represented by a fuzzy membership function can be modelled probabilistically by distributions over intended meanings, and the advantages of replacing fuzzy membership values with probabilities were well justified in the literature (34-37). Therefore, instead of fuzzy membership, the probability of a variable being a member of a given class was well thought out. In our OFS, (i) the fuzzy membership function of each linguistic term corresponding to a fuzzy set of a variable is replaced by a p.d.f. that classifies the input into respective classes (33), (ii) the rule base is redefined, and (iii) the conflict resolution is made by optimizing the p.d.f. of the output variable using the most popular meta-heuristic whale optimization algorithm (WOA). To avoid complexity, the data space of each input variable is classified into three classes: *low (l), medium (m)*, and *high (h)*. Researchers may consider more classes based on the nature of the variables and their data spaces in an application domain. For an input variable *x*, the probabilities of a data point ($x = x_0$) in the respective classes are defined by three triangular p.d.f.s, as defined by equations [1], [2], and [3].

$$P_{l}(x_{0}) = \begin{cases} 1, & x_{0} = x_{min} \\ \frac{x_{mid} - x_{0}}{x_{mid} - x_{min}}, & x_{min} < x_{0} < x_{mid} \end{cases}$$
[1]

$$P_{m}(x_{0}) = \begin{cases} 0, & x_{0} = x_{mid} \\ 0, & x_{0} = x_{min} \\ \frac{x_{0} - x_{min}}{x_{mid} - x_{min}}, & x_{min} < x_{0} < x_{mid} \\ \frac{x_{max} - x_{0}}{x_{max} - x_{mid}}, & x_{mid} < x_{0} < x_{max} \\ 0, & x_{0} \ge x_{max} \\ 0, & x_{0} \le x_{max} \end{cases}$$
[2]

$$P_{h}(x_{0}) = \begin{cases} 0, & x_{0} \le x_{mid} \\ \frac{x_{0} - x_{mid}}{x_{max} - x_{mid}}, & x_{mid} < x_{0} < x_{max} \\ 1, & x_{0} = x_{max} \end{cases}$$
[3]

Where x_{min} and x_{max} are the minimum and maximum values of x observed in its data space, and $x_{mid} = (x_{max} - x_{min})/2$. The choice of three triangular p.d.f.s holds the basic condition of probability:

$$\sum_{i=l,m,h} P(class \ i \ ; x = x_0) = 1 \quad : \ \forall x_0$$
[4]

Therefore, the probability $P_i(x_0)$ of an input data point x_0 being classified in the class *i* is defined by: $P_i(x_0) = P(class \ i; x = x_0)$ [5]

During the coding process, the categorization probabilities $P_i(x_0)$: (i = l, m, h) of an input data point x_0 are computed. Similarly, for multiple inputs, the categorization probabilities of all the inputs are obtained. The rule base consists of a collection of '*if*-

if (Class i; $x = x_0$) AND (Class i; $y = y_0$) then (Class i; z)

then the structure of the rule is defined as:

then' rules nearly identical to a fuzzy rule base. However, instead of fuzzy linguistic terms, class labels are used. Suppose x and y are two input variables and z is the output variable,

[6]

probabilities $P_i(x_0)$ and $P_i(y_0)$ for input values x_0

and y_0 are computed for each rule. The class

probabilities are then multiplied together for

each output function under the assumption of

independence of classification (38), and the

output p.d.f. for the *n*-th rule is given by:

Where *i* represents the class labels (*low, medium, high*), and x_0 and y_0 are the input data points of *x* and *y*, respectively. Similar to the method adopted for input label classification, the output variable *z* is also classified into three labels: *low* (*l*), *medium* (*m*), and *high* (*h*), as defined by equations [1], [2], and [3]. In the conflict resolution stage, the class

$$P_n(z) = \sum_{k=l,m,h} \sum_{j=l,m,h} \sum_{i=l,m,h} P_i(x_0) P_j(y_0) P_k(z)$$
[7]

When *r* rules are activated simultaneously for input values x_0 and y_0 , the resulting p.d.f. of the output variable *z* is obtained by:

$$P(z) = \sum_{n=1}^{r} P_n(z)$$
 [8]

For more than two input variables, the conflict resolution is made by calculating the class probabilities of all the input variables and then multiplying them with the mean of the corresponding p.d.f. of the output variable z. The mean of output z is a weighted average of the mean of the z_i 's weighted by the probabilities P_i 's. Thus, the output crisp value z_n is given by:

$$z_n = \sum_{k=l,m,h} \sum_{j=l,m,h} \sum_{i=l,m,h} P_i(x_0) P_j(y_0) \alpha_k(z)$$
[9]

Where α_l , α_m , and α_h are the means of the three p.d.f.s, P_l (*z*), P_m (*z*), and P_h (*z*) of the output variable *z*.

Equation [9] defines that using the correlation product encoding method, the value (z_n) of the output variable z can be obtained in terms of the class probabilities of the input variables and the mean values α_l , α_m , and α_h of the output p.d.f.s, $P_l(z)$, P_m (z), and P_h (z). The class probabilities of all the input variables are obtained using equations [1], [2], and [3]. However, in order to get the optimal output Z_n , precise determination of the values of α_l , α_m , and α_h is a challenging problem in such decisionmaking systems designed using fuzzy probabilistic controller approaches. The optimization process is capable of solving this problem. Several applications of different optimization algorithms for designing fuzzy logic systems have been reported in the literature (39-42). We attempted different nature-inspired algorithms to get the optimum value of the output z_n where the WOA outperformed the others. An outstanding performance of WOA has already been recognized in diverse domains to solve optimization problems including handwritten Arabic optical character recognition (43), imagery segmentation (44), battlefield simulation (45) etc. However, the application of the WOA for designing an OFS has never been reported yet. The WOA can mimic the hunting strategy of humpback whales (46). These whales hunt fish using a technique known as bubble net feeding. The bubble net feeding strategy primarily employs two distinct strategies: surrounding the prey (target) and subsequently narrowing the circle towards it, and spiral assault. The algorithm proceeds by applying various permutations of whale position to reach the final global maxima (46). Initially, we considered 25 whales distributed over the search space. The value of maximum iteration was set to 500, and the values of lower and upper bounds were set to 0 and 1500, respectively. The standard reference value of the N fertilizer (N_{STCR}) was calculated using the soil test crop response (STCR) method (47). The STCR method suggests the exact quantity of fertilizer required for a particular variety of crop (48). The Indian Council for Agricultural Research (ICAR) recognized this model as a standardized method with the main objective of helping the farmers achieve the target yield to its full potential through the application of the optimum quantity of fertilizers. The standard reference value (NSTCR) was considered the target, and the coefficient of determination (R^2) was used as the fitness function for optimization. The major advantages of R^2 are that it is dimensionless, stable, and bound between 0 and 1 (49).

System Validation

In order to validate the performance of our OFS in field applications, it was implemented against three varieties of paddy (IET-4094, IET-4097, and BORO-4789) and two varieties of potato (Kufri Jyoti High Yielding (KJ-HY) and Kufri Jyoti Low Yielding (KJ-LY)), cultivated as the major crops in the study area comprising of three districts, Burdwan, Hooghly, and Nadia in the state of West Bengal, India. The average yield rates of paddy and potato in this area are 3194 and 35848 kg/ha, respectively (50). The description of the study area, the source and nature of the field data, and the empirical results are discussed below.

Study Area

The three districts, Burdwan, Hooghly, and Nadia, are situated in the Gangetic alluvial plain in West Bengal and spread over 7024 sq. km. between latitudes 22.47°N and 23.82°N and longitudes 86.80°E and 88.69°E, as presented in Figure 1.



Figure 1: Study Area

This region was considered for study because it is one of the most agriculturally productive areas of the state, having approximately 69% of the land under cultivation (51), the main source of livelihood for the majority of the people in this region is agriculture, this area has a diverse fertilizer consumption index from low to high (Zscore ranging from -0.323 to 1.765) (52), this area consists of varied soil types such as clayey, clayey loam, loamy, loamy sandy, and gravelly loam (51), paddy and potato are the two major cash crops of this area, and nearly 1.1783 and 0.2239 million hectares are under cultivation for paddy and potato, respectively (53). The diversified characteristics of this region further inspired us to select this area for study.

Data Collection

The five parameters; the nitrogen content of the soil (N), the measure of the soil acidity (pH), the electrical conductivity of the soil (EC), the soil organic carbon (OC), and the average rainfall (Ra) were considered the input variables of the OFS. The performance was validated using authentic

datasets containing all these input parameters for the study area. The dataset of soil parameters was collected from the Soil Health Card Data repository provided by the Department of Agriculture and Farmers Welfare, Government of India (54). A total of 922, 9042, and 1599 samples were collected from three districts: Burdwan, Hooghly, and Nadia, respectively. The dataset for average rainfall was collected from the Climate Research Unit, University of East Anglia, UK dataset (version 4.07) (55). All the data collected from these two sources were merged together to make a single dataset.

Probability Density Functions

The p.d.f.s of each of the five input variables (N, pH, EC, OC, and Ra) were constructed from the dataset. Three triangular p.d.f.s, defined by equations [1-3], determined the class probabilities (P_i) of data points to classify in the respective classes (*low, medium, and high*). For the existing dataset, the shape of the p.d.f.s of each input variable and the output variable N_R , their degree of overlap, and data spaces are presented in Figures 2(A-F).



Figure 2A: The p.d.f.s of Soil Nitrogen (*N*)



Figure 2B: The p.d.f.s of soil pH (pH)



Figure 2E: The p.d.f.s of average rainfall (Ra)

Rule Base

The rule base consists of rules with an antecedentconsequent or *if-then* structure. The antecedent (*if part*) is formed by all possible combinations of premises obtained from five input variables with three class labels (*low, medium,* and *high*). The premises are combined using a union operator





Figure 2F: The p.d.f.s of the recommended N (NR)

(AND). The consequent part is obtained by the class probability of the output variable. For five input variables, a complete set of 125 rules was framed. As the rule base contains the complete set of all possible rules, the dependency on rule optimization by a human expert is circumvented. For example, the structure of a rule for five input variables, *N*, *pH*, *EC*, *OC*, and *Ra*, is defined as:

Where *i*, *j*, and *k* define any one of the class labels from *low*, *medium*, or *high*, and N_R is the output variable. The output p.d.f. for the above rule is given by:

$$P(N_R) = \sum_{k=l,m,h} \sum_{j=l,m,h} \sum_{i=l,m,h} P_i(N) P_j(pH) P_j(EC) P_i(OC) P_j(Ra) P_k(N_R)$$
[11]

If r rules are activated simultaneously, then the resulting p.d.f. of the output variable N_R is obtained by:

$$P(N_R) = \sum_{n=1}^{r} P_n(N_R)$$

Performance Metrics Used for Validation

A significant part of the model-building process is the evaluation of the performance of the model. The performance was evaluated in terms of four well-accepted statistical metrics: the coefficient of determination (R^2), the Nash-Sutcliffe efficiency (*NSE*), the root mean squared error (*RMSE*), and the mean absolute error (*MAE*). R^2 and *NSE* assess the correctness of the model, while *RMSE* and *MAE* project the errors in the predictions. The definition of these four metrics is provided in (56):

[12]

$$R^{2} = \frac{\left[\sum_{i=1}^{n} N_{Ri} - \overline{N_{R}} (N_{STCRi} - \overline{N_{STCR}})\right]^{2}}{\sum_{i=1}^{n} N_{Ri} - \overline{N_{R}}^{2} \sum_{i=1}^{n} (N_{STCRi} - \overline{N_{STCR}})^{2}}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n} (N_{Ri} - \overline{N_{STCR}})^{2}}{\sum_{i=1}^{n} (N_{Ri} - \overline{N_{R}})^{2}}$$
[13]

Vol 6 | Issue 2

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \sum_{i=1}^{n} (N_{Ri} - N_{STCRi})^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |N_{Ri} - N_{STCRi}|$$
[1]

Where N_{Ri} is the recommended dose of nitrogen and N_{STCRi} is the reference value of nitrogen for *i*-th record in the dataset. $\overline{N_R}$ is the mean of the recommended doses of nitrogen, $\overline{N_{STCR}}$ is the mean of the reference values of nitrogen, and n is the total of records in the dataset. The outcome of the



Figure 3A: Convergence Plot for Paddy (IET-4094)



Figure 3C: Convergence Plot for Paddy (BORO-4789)



convergence study of the optimization is presented by means of the convergence graphs, where the horizontal axis represents the number of iterations and the vertical axis defines the values of the fitness function (R^2). The convergence graphs for five crops are presented in Figures 3 (A-E).



Figure 3B: Convergence Plot for Paddy (IET-4097)



Figure 3D: Convergence Plot for Potato (KJ-HY)



Figure 3E: Convergence Plot for Potato (KJ-HY)

The maximum number of iterations was set to 500, and the lower and upper limits of the search space of WOA were set to 0 and 1500, respectively. The convergence graphs depict that the values of R^2 gradually converged towards optimum values with an increasing number of iterations, and the values of the fitness function (R^2) , and finally, at saturation, the best possible values were achieved.

Results

To study the performance of the system, it was deployed against three varieties of paddy (IET-

targeted crops using four popular and well accepted metrics (*R*², *NSE*, *RMSE*, and *MAE*). Table 1 summarizes the empirical values of these four metrics obtained against five crops.

Table 1:	Empirical	Values of Four	Metrics	Obtained	against Five	Crops
----------	-----------	----------------	---------	----------	--------------	-------

Crops	R ²	NSE	RMSE (kg/ha)	MAE (kg/ha)
Paddy (IET-4094)	0.9650	0.9627	12.4229	8.5884
Paddy (IET-4097)	0.9628	0.9581	9.5839	6.5011
Paddy (BORO-4789)	0.9685	0.9616	55.2985	40.4300
Potato (KJ-HY)	0.9880	0.9876	12.6227	6.9934
Potato (KJ-LY)	0.9810	0.9810	12.3918	6.3800

For comprehensive visualization, several easy-tounderstand graphical tools have been suggested in the literature. Among them, the scatter plots and box plots are very popular and widely used. The scatter plots are very useful graphical tools for projecting the relationship between two variables within the same interval. The two-dimensional data points scattered within the coordinate space help to determine the strength of the correlation between the two variables. The box plot is another graphical representation of statistical information about two or more variables placed side-by-side. On the other hand, a box plot typically displays the

median along with the upper and lower hinge points, or boundaries, of the data points. The upper and lower hinges are indicated by whiskers or end markers, and any value beyond these points is treated as an outlier. The box plots help in visualizing the central tendency and variability of two or more groups of variables and facilitate to understand the relationship between them. In addition to Table 1, Figures 4 and 5 present the scatter plots and box plots that represent the system recommended doses of N fertilizer (N_R) against the reference values of N fertilizer (N_{STCR}).



Figure 4A: Scatter Plot for Paddy (IET-4094)



600

Vol 6 | Issue 2



Figure 5C: Box Plot for Paddy (BORO-4789) Figure 5D: Box Plot for Potato (KJ-HY)



Figure 5E: Box Plot for Potato (KJ-LY)

The coefficient of correlation (R^2) is a popular regression metric used to measure the degree of collinearity between the actual and recommended doses of an output variable (57). The values of R^2 range from 0 to 1, and a higher value of R^2 indicates better performance of a system (49). It is evident from Table 1 that, when measured in terms of R^2 , the best performance was reported against potato (KJ-HY) (with $R^2 = 0.9880$). On the other hand, the least performance was observed against paddy (IET-4097) (with $R^2 = 0.9628$). A lower but comparable value of R^2 was obtained for potato (KJ-LY). Comparatively lower performances were achieved for other two crops, paddy (IET-4094) and paddy (BORO-4789), in terms of R^2 . The results of the experiment reveal that the system performed exceptionally well in terms of percentage (96.28 to 98.80%). The maximum variation in performances for five different crops was negligible (only 2.61%), which signifies the robustness of our proposed system. Nash-Sutcliffe efficiency (NSE) is another dimensionless metric that incorporates measurement uncertainty and provides a clear indication of how well the output values match the reference values (57). Similar to R^2 , the values of *NSE* range from 0 to 1 and have the same significance as R^2 . When measured in terms of NSE, an identical ranking of system performances was observed against potato (KJ-HY), potato (KJ-LY), and paddy (IET-4094). However, the paddy (BORO-4789) secured the fourth position, followed by the paddy (IET-4094). In terms of NSE, the performance of the system is very substantial (ranging from 95.81 to 98.76%), and the maximum variation in performance was observed to be only 2.95%. Root mean squared error (RMSE) and mean absolute error (MAE) are

498

standard and well-established error metrics used for performance analysis. *RMSE* and *MAE* have the same unit as the output variable and are very convenient to interpret the errors in the recommendation (57). The lowest value of RMSE (9.5839 kg/ha) was observed for paddy (IET-4094), whereas, the highest value of RMSE (55.2985 kg/ha) was achieved against paddy (BORO-4789). In terms of RMSE, a nearly equitable performance was reported against the other three crops: paddy (IET-4094), potato (KJ-HY), and potato (KJ-LY). We observed an exact identical ranking in system performances against these five crops, based on MAE data. For other crops, except paddy (BORO-4789), the maximum differences in RMSE and MAE were 3.0388 and 2.2084 kg/ha, respectively. This observation indicates that, in most cases, the system is very reliable. It is interesting to note that in the higher range of N_{STCR}, the recommended doses (N_R) for paddy (IET-4094) were higher than the reference values of *N*_{STCR}, as shown in Figure 4(A). It signifies that the system's performance degrades at the higher range of NSTCR. A possible reason is that a negative skewness is observed in the dataset for the two input variables, pH and Ra, which might adversely affect the optimization process at the higher range of NSTCR. A similar observation is also reflected in the scatter plots of paddy (BORO-4789), potato (KJ-HY) and potato (KJ-LY), as presented in Figures 4(C-E). However, the scatter plot obtained for paddy (IET-4097) is somewhat different due to some unknown reason. Hopefully, our future research will be able to investigate the root cause. In most cases, the system's limitation was that, at the higher ranges, it recommended higher doses of N fertilizer (N_R) than the reference values of N

fertilizer (N_{STCR}). A similar observation is also validated by the box plots (Figures 5(A-E)). Our future attempt will be to overcome this limitation by better tuning the optimization parameters. The scatter plots and box plots (Figures 4 and 5) reveal that the doses of N fertilizer recommended by our proposed system are consistent with the standard doses suggested by STCR (48). We studied the

efficiency of our system with four other optimization techniques, such as Particle Swarm Optimization (PSO), Ant-lion Optimization (ALO), Grey Wolf Optimization (GWO), and Firefly Optimization (FFO). The empirical values of the fitness function (R^2) obtained against five crops for each of the four optimization techniques are presented in Table 2.

Table 2: The Empirical Values of the Fitness Function (R^2) Obtained against Five Crops for Each Optimization Technique

Optimization Techniques	Crops	R ²
Particle Swarm Optimization	Paddy (IET-4094)	0.9430
	Paddy (IET-4097)	0.9387
	Paddy (BORO-4789)	0.9432
	Potato (KJ-HY)	0.9612
	Potato (KJ-LY)	0.9634
Ant-lion Optimization	Paddy (IET-4094)	0.9601
	Paddy (IET-4097)	0.9587
	Paddy (BORO-4789)	0.9571
	Potato (KJ-HY)	0.9652
	Potato (KJ-LY)	0.9675
Grey Wolf Optimization	Paddy (IET-4094)	0.9629
	Paddy (IET-4097)	0.9605
	Paddy (BORO-4789)	0.9570
	Potato (KJ-HY)	0.9811
	Potato (KJ-LY)	0.9705
Firefly Optimization	Paddy (IET-4094)	0.9547
	Paddy (IET-4097)	0.9563
	Paddy (BORO-4789)	0.9592
	Potato (KJ-HY)	0.9689
	Potato (KJ-LY)	0.9681

Table 2 reveals that the values of the fitness function (R^2) obtained using the other four popular optimization techniques were less than those obtained using the Whale Optimization Algorithm as presented in Table 1. This observation demonstrates why WOA was the optimal selection for this application.

Discussion

Our proposed system was compared with the existing nitrogen guidance systems, which were designed using various conventional agronomic models such as linear regressor, ridge regressor, gradient boosted trees, least absolute shrinkage and selection operator (LASSO), support vector regressor, etc., to solve Ν fertilizer recommendation problems for different crops worldwide. The conventional models whose performance was evaluated in terms of R^2 were selected for a fair comparison. The targeted crops, the conventional models used, and the reported values of R^2 were available from the respective papers and are summarized in Table 3. The findings of the comparative study are presented in Figure 6 using a comparative graph representing the R^2 values of the different models of various N fertilizer recommendation systems.

Table 3: Details of the Conventional Models Used, the Targeted Crops, and the Reported R²

Models	Designed by	Conventional models used	Targeted crops	R^2
Model-1	Qin <i>et al.</i> (16)	Linear regressor	Corn	0.1900
Model-2		Ridge regressor	Corn	0.3500

Model-3		Gradient boosted trees	Corn	0.3700
Model-4		LASSO	Corn	0.4200
Model-5	Puntel <i>et al.</i> (15)	Quadratic regressor	Corn	0.4500
Model-6	Zhang <i>et al.</i> (18)	Support vector regressor	Rice	0.6800
Model-7		Random forest regressor	Rice	0.7200
Model-8	Ransom <i>et al.</i> (20)	Random forest regressor	Corn	0.8400
Model-9	Sarkar <i>et al.</i>	Optimized fuzzy model	Paddy (IET-4094)	0.9650
Model-10	(proposed OFS)		Paddy (IET-4097)	0.9628
Model-11			Paddy (BORO-4789)	0.9685
Model-12			Potato (KJ-HY)	0.9880
Model-13			Potato (KJ-LY)	0.9810



Figure 6: Comparative Graph

The comparative graph depicts that in terms of R^2 , Model-12 was the best-performing model (with R^2 = 0.9880), and Model-1 was the worst. The Model-13 was the second-best (with R^2 = 0.9810), whereas the Model-11 secured the third position, followed by the Model-9 and the Model-8. The values of R^2 for the rest of the systems were very small and were not comparable with others. The overall performance of our designed OFS (with R^2 ranging from 0.9628 to 0.9880) is better than other systems (with R^2 ranging from 0.1900 to 0.8400). The results depicted in Figure 6 confirm the superiority of our proposed system over all other existing systems designed for N fertilizer recommendations.

Conclusion

Site-specific application of N fertilizer in an optimal dose is one of the significant challenges in precision agriculture. Several machine learningbased systems have been proposed as an alternative to soil scientists for N fertilizer recommendations. Nevertheless, one of the major drawbacks of these systems is their low precision. In this study, we have proposed an innovative optimal fuzzy system that can recommend the precise dose of N fertilizer. A nature-inspired meta-heuristic optimization algorithm has overwhelmed the painful trial-and-error methods of obtaining the optimal output of a classical fuzzy system. We have proposed a unique architecture for a highly effective optimal fuzzy system, combining it with the whale optimization algorithm to recommend the ideal dose of N fertilizer. Empirical studies using field data revealed that our proposed system with high accuracy outperformed all the existing systems. The significant contribution of our proposed system is that it can recommend the precise dose of N fertilizer based on the site-specific soil and climatic parameters while sidestepping the intervention of expensive and scarce soil scientists. Hopefully, such a system will benefit farmers by applying N fertilizer at an optimal dose to achieve their best yield, conserve the agroecosystem, and ultimately promote sustainable agriculture.

Abbrevations

ASM: Agricultural simulation model, BR: Bayesian regressor, EC: Electrical conductivity, GBT: Gradient-boosted tree, ICAR: Indian Council for Agricultural Research, KJ-HY: Kufri Jyoti High Yielding, KJ-LY: Kufri Jyoti Low Yielding, LASSO: Least absolute shrinkage and selection operator, LR: Linear regression, MAE: Mean absolute error, ML: Machine learning, N: Nitrogen, N_R : Recommended nitrogen, NSE: Nash-Sutcliffe efficiency, *N*_{STCR}: Standard reference value, OC: Organic carbon, OFS: Optimal fuzzy system, p.d.f.: Probability density functions, Ra: Average rainfall, RF: Random forest, RMSE: Root means squared error, RR: Ridge regressor, STCR: Soil test crop response, SVM: Support vector machine, WOA: Whale optimization algorithm

Acknowledgement

None.

Author Contributions

Uditendu Sarkar: Conceive, Experiment Design, Experiments perform, data analyzation, figures and/or tables prepare, draft preparation, review, Gouravmoy Banerjee: Conceive, Experiment Design, Experiments perform, data analyzation, figures and/or tables prepare, draft preparation, review, Indrajit Ghosh: Conceive, Experiment Design, Experiments perform, data analyzation, figures and/or tables prepare, draft preparation, review.

Conflict of Interest

The authors declare no conflict of interest.

Ethics Approval

Not Required.

Funding

No funding was received for the current research.

References

- 1. Fu Z, Zhang K, Zhang J, Zhang Y, Cao Q, Tian Y, Zhu Y, Cao W, Liu X. Optimizing nitrogen application and sowing date can improve environmental sustainability and economic benefit in wheat-rice rotation. Agricultural Systems. 2023; 204:103536.
- 2. Fathi A, Zeidali E. Conservation tillage and nitrogen fertilizer: a review of corn growth and yield and weed management. Central Asian Journal of Plant Science Innovation. 2021; 1(3):121-142.
- 3. Nabavi-Pelesaraei A, Rafiee S, Mohtasebi SS, Hosseinzadeh-Bandbafha H, Chau KW. Integration of artificial intelligence methods and life cycle

assessment to predict energy output and environmental impacts of paddy production. Science of the total environment. 2018; 631:1279-1294.

- 4. Luo H, Guan Q, Lin J, Wang Q, Yang L, Tan Z, Wang N. Air pollution characteristics and human health risks in key cities of northwest China. Journal of Environmental Management. 2020; 269:110791.
- 5. Kyveryga PM, Blackmer AM, Morris TF. Disaggregating model bias and variability when calculating economic optimum rates of nitrogen fertilization for corn. Agronomy Journal. 2007; 99:1048–1056.
- De Lara A, Mieno T, Luck JD, Puntel LA. Predicting site-specific economic optimal nitrogen rate using machine learning methods and on-farm precision experimentation. Precision Agriculture. 2023; 24(5):1792-812.
- 7. Djaman K, Mel VC, Ametonou FY, El-Namaky R, Diallo MD, Koudahe K. Effect of nitrogen fertilizer dose and application timing on yield and nitrogen use efficiency of irrigated hybrid rice under semi-arid conditions. Journal of Agricultural Science and Food Research. 2018; 9(2):2-7.
- Puntel LA, Sawyer JE, Barker DW, Dietzel R, Poffenbarger H, Castellano MJ, Moore KJ, Thorburn P, Archontoulis SV. Modeling long-term corn yield response to nitrogen rate and crop rotation. Frontiers in plant science. 2016; 7:1630.
- 9. Samal SK, Prasad L, Kumar R. How to apply fertilizers, based on soil test? A step by step guide. Food and Scientific Reports. 2020; 1(6):51-52.
- 10. Doltra J, Gallejones P, Olesen JE, Hansen S, Frøseth RB, Krauss M, Stalenga J, Jończyk K, Martínez-Fernández A, Pacini GC. Simulating soil fertility management effects on crop yield and soil nitrogen dynamics in field trials under organic farming in Europe. Field crops research. 2019; 233:1-11.
- Adeboye OB, Schultz B, Adeboye AP, Adekalu KO, Osunbitan JA. Application of the AquaCrop model in decision support for optimization of nitrogen fertilizer and water productivity of soybeans. Information Processing in Agriculture. 202; 8(3):419-36.
- 12. Jin Z, Prasad R, Shriver J, Zhuang Q. Crop model-and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system. Precision Agriculture. 2017; 18:779-800.
- 13. Kasampalis DA, Alexandridis TK, Deva C, Challinor A, Moshou D, Zalidis G. Contribution of remote sensing on crop models: a review. Journal of Imaging. 2018 Mar 23; 4(4):52.
- 14. Dong Y, Fu Z, Peng Y, Zheng Y, Yan H, Li X. Precision fertilization method of field crops based on the Wavelet-BP neural network in China. Journal of Cleaner Production. 2020; 246:118735.
- 15. Puntel LA, Pagani A, Archontoulis SV. Development of a nitrogen recommendation tool for corn considering static and dynamic variables. European Journal of Agronomy. 2019; 105:189-99.
- 16. Qin Z, Myers DB, Ransom CJ, Kitchen NR, Liang SZ, Camberato JJ, Carter PR, Ferguson RB, Fernandez FG, Franzen DW, Laboski CA. Application of machine learning methodologies for predicting corn economic optimal nitrogen rate. Agronomy Journal. 2018; 110(6):2596-607.

- 17. Sulik J, Banger K, Janovicek K, Nasielski J, Deen B. Comparing Random Forest to Bayesian Networks as nitrogen management decision support systems. Agronomy Journal. 2023; 115(3):1431-46.
- 18. Zhang J, Fu Z, Zhang K, Li J, Cao Q, Tian Y, Zhu Y, Cao W, Liu X. Optimizing rice in-season nitrogen topdressing by coupling experimental and modeling data with machine learning algorithms. Computers and Electronics in Agriculture. 2023; 209:107858.
- 19. Wen G, Ma BL, Vanasse A, Caldwell CD, Smith DL. Optimizing machine learning-based site-specific nitrogen application recommendations for canola production. Field Crops Research. 2022; 288:108707.
- 20. Ransom CJ, Kitchen NR, Camberato JJ, Carter PR, Ferguson RB, Fernández FG, Franzen DW, Laboski CA, Myers DB, Nafziger ED, Sawyer JE. Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations. Computers and electronics in agriculture. 2019; 164:104872.
- 21. Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and electronics in agriculture. 2018; 151:61-9.
- 22. Assimakopoulos JH, Kalivas DP, Kollias VJ. A GISbased fuzzy classification for mapping the agricultural soils for N-fertilizers use. Science of the Total Environment. 2003; 309(1-3):19-33.
- 23. Papadopoulos A, Kalivas D, Hatzichristos T. Decision support system for nitrogen fertilization using fuzzy theory. Computers and electronics in agriculture. 2011; 78(2):130-9.
- 24. Bouroubi Y, Tremblay N, Vigneault P, Bélec C, Panneton B, Guillaume S. Fuzzy logic approach for spatially variable nitrogen fertilization of corn based on soil, crop and precipitation information. In: Murgante B, Gervasi O, Iglesias A, Taniar D, Apduhan BO. editors. ICCSA: International Conference on Computational Science and Its Applications, 20-23 June 2011, Santander, Spain. Berlin: Springer, 2011. p. 356-368. https://doi.org/10.1007/978-3-642-21928-3_25
- 25. Kweon G. Delineation of site-specific productivity zones using soil properties and topographic attributes with a fuzzy logic system. Biosystems engineering. 2012; 112(4):261-77.
- 26. Abd Razak MS, Abdul-Rahman S, Mutalib S, Abd Aziz Z. Nitrogen Fertilizer Recommender for Paddy Fields. In: Mohamed A, Berry MW, Yap BW, editors. Soft Computing in Data Science. Singapore: Springer; 2017. p. 230–240. https://doi.org/10.1007/978-981-10-7242-0_20
- 27. Prabakaran G, Vaithiyanathan D, Ganesan M. Fuzzy decision support system for improving the crop productivity and efficient use of fertilizers. Computers and electronics in agriculture. 2018; 150:88-97.
- Nooriman WM, Abdullah AH, Abdul Rahim N, Tan ESMM. Fuzzy logic based prediction of micronutrients demand for harumanis mango growth cycles. J Phys: Conf Ser. 2021 Nov; 2107(1):012048.

- 29. Ashraf A, Akram M, Sarwar M. Type-II Fuzzy Decision Support System for Fertilizer. The Scientific World Journal. 2014; 2014(1):695815.
- Pawar PM, Ganguli R. Structural Health Monitoring Using Genetic Fuzzy Systems. London: Springer; 2011. http://link.springer.com/10.1007/978-0-85729-907-9
- 31. Zadeh LA. Fuzzy sets. Information and Control. 1965; 8(3): 338-353.
- 32. Berenji HR. Fuzzy Logic Controllers. In: Yager RR, Zadeh LA, editors. An Introduction to Fuzzy Logic Applications in Intelligent Systems. Boston, MA: Springer US; 1992. p. 69–96. https://doi.org/10.1007/978-1-4615-3640-6_4
- 33. Laviolette M, Seaman JW, Barrett JD, Woodall WH. A
- probabilistic and statistical view of fuzzy methods. Technometrics. 1995; 37(3):249-61.
- 34. Hisdal E. Infinite-valued logic based on two-valued logic and probability. Part 1.1. Difficulties with present-day fuzzy-set theory and their resolution in the TEE model. International journal of manmachine studies. 1986; 25(1):89-111.
- 35. Hisdal E. Infinite-valued logic based on two-valued logic and probability. Part 1.2. Different sources of fuzziness. International journal of man-machine studies. 1986; 25(2):113-38.
- 36. Hisdal E. Are grades of membership probabilities? Fuzzy Sets and Systems. 1988; 25(3):325-48.
- Cheeseman P. Probabilistic versus fuzzy reasoning. Machine Intelligence and Pattern Recognition 1986; 4:85-102. doi: 10.1016/B978-0-444-70058-2.50011-5.
- Douglas Barrett J, Woodall WH. A probabilistic alternative to fuzzy logic controllers. IIE transactions. 1997; 29(6):459-67.
- 39. Fang G, Kwok NM, Ha Q. Automatic Fuzzy Membership Function Tuning Using the Particle Swarm Optimization. In: 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application. 2008. p. 324–8. https://ieeexplore.ieee.org/abstract/document/47 56789
- 40. Zhang HX, Wang F, Zhang B. Genetic optimization of fuzzy membership functions. In: 2009 International Conference on Wavelet Analysis and Pattern Recognition. 2009. p. 465–70. https://ieeexplore.ieee.org/abstract/document/52 07463
- 41. Lagunes ML, Castillo O, Soria J. Optimization of membership function parameters for fuzzy controllers of an autonomous mobile robot using the firefly algorithm. In: Castillo O, Melin P, Kacprzyk J, editors. Fuzzy logic augmentation of neural and optimization algorithms: theoretical aspects and real applications. Cham: Springer; 2018. p. 199-206. https://doi.org/10.1007/978-3-319-71008-2_16
- 42. Nikolić M, Šelmić M, Macura D, Ćalić J. Bee colony optimization metaheuristic for fuzzy membership functions tuning. Expert Systems with Applications. 2020; 158:113601.
- 43. Sahlol AT, Abd Elaziz M, Al-Qaness MA, Kim S. Handwritten Arabic optical character recognition approach based on hybrid whale optimization algorithm with neighbourhood rough set. IEEE Access. 2020; 8:23011-21.

- 44. Yan Z, Zhang J, Yang Z, Tang J. Kapur's entropy for underwater multilevel thresholding image segmentation based on whale optimization algorithm. IEEE access. 2020; 9:41294-319.
- 45. Duan Y, Liu C, Li S. Battlefield target grouping by a hybridization of an improved whale optimization algorithm and affinity propagation. IEEE Access. 2021; 9:46448-61.
- 46. Mirjalili S, Lewis A. The whale optimization algorithm. Advances in engineering software. 2016; 95:51-67.
- 47. Ramamoorthy B, Velayutham M. Soil test crop response correlation work in India. World Soil Resources Reports, FAO, Rome. 1971; 41:96-102.
- 48. Bidhan Chandra Krishi Viswavidyalaya. Soil Test Crop Response. BCKV: Kalyani, Nadia; 2020. https://www.bckv.edu.in/index.php/en/aicrp-onsoil-test-crop-response-correlation-en
- 49. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peerj Computer Science. 2021; 7:e623.
- 50. Department of Agriculture and Farmers Welfare, Government of India. District-wise, season-wise crop production statistics from 1997. DAFW: New Delhi, India; 2020. https://data.gov.in/resource/districtwise-season-wise-crop-production-statistics-1997
- 51. Indian Council for Agricultural Research Central Research Institute for Dryland Agriculture. District Plan. ICAR - CRIDA: New Delhi, India; 2012. https://www.icar-crida.res.in/CP-2012/

- 52. Rukhsana, Alam A. Levels of Agriculture Development and Crop Diversification: A District-Wise Panel Data Analysis in West Bengal. In: Rukhsana, Alam A, editors. Agriculture, Environment and Sustainable Development: Experiences and Case Studies. Cham: Springer International Publishing; 2022. p. 105–17. https://doi.org/10.1007/978-3-031-10406-0_7
- 53. Mandal S, Burman D, Mandal UK, Sharma PC. Agricultural Production and Marketing of Major Food Crops and Spices in West Bengal - Status and Strategies. ICAR-CSSRI; 2021. p. 45. https://cssri.res.in/PSSU/Publications/Folder%20a nd%20Bulletins/2021/2021-1%20-Dr.%20Subhasis%20Mandal%20-%20Policy%20Paper.pdf
- 54. Department of Agriculture and Farmers Welfare, Government of India. Soil Health Card. DAFW: New Delhi, India; 2017. https://soilhealth.dac.gov.in/home
- 55. Harris I, Osborn TJ, Jones P, Lister D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific data. 2020 Apr 3; 7(1):109.
- 56. Zhao L, Zhao X, Zhou H, Wang X, Xing X. Prediction model for daily reference crop evapotranspiration based on hybrid algorithm and principal components analysis in Southwest China. Computers and Electronics in Agriculture. 2021; 190:106424.
- 57. Moriasi DN, Gitau MW, Pai N, Daggupati P. Hydrologic and water quality models: Performance measures and evaluation criteria. Transactions of the ASABE. 2015; 58(6):1763-85.