






Article

Advanced Efficient Feature Selection Integrating Augmented Extreme Learning Machine and Particle Swarm Optimization for Predicting Nitrogen Use Efficiency and Yield in Corn

Josselin Bontemps^{1,*} , Isa Ebtehaj¹ , Gabriel Deslauriers², Alain N. Rousseau³ , Hossein Bonakdari⁴  and Jacynthe Dessureault-Rompré¹ 

¹ Department of Soils and Agri-Food Engineering, Université Laval, Quebec, QC G1V 0A6, Canada; isa.ebtehaj.1@ulaval.ca (I.E.); jacynthe.dessureault-rompre@fsaa.ulaval.ca (J.D.-R.)

² PleineTerre Inc., 169 Rue St-Jacques, Napierville, Quebec, QC J0J 1L0, Canada; gdeslauriers@pleineterre.com

³ Institut National de la Recherche Scientifique–Eau Terre Environnement Research Centre, Quebec, QC G1K 9A9, Canada; alain.rousseau@inrs.ca

⁴ Department of Civil Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada; hossein.bonakdari@uottawa.ca

* Correspondence: josselin.bontemps.1@ulaval.ca

Abstract: Efficient nitrogen management is crucial for improving corn productivity while minimizing environmental impacts. This study evaluates the response of corn to nitrogen fertilization using three key metrics: yield; nitrogen harvest index (NHI); and agronomic nitrogen use efficiency (ANUE). This experiment was conducted over three years (2021–2023) across 84 sites in Quebec, Canada, with five nitrogen treatments applied post-emergence (0, 50, 100, 150, 200 kg N/ha) and initial nitrogen applied at seeding (30 to 60 kg/ha). In addition, various soil health indicators, including physical, chemical, and biochemical properties, were monitored to understand their interaction with nitrogen use efficiency. Machine learning techniques, such as augmented extreme learning machine (AELM) and particle swarm optimization (PSO), were employed to optimize nitrogen recommendations by identifying the most relevant features for predicting yield and nitrogen use efficiency (NUE). The results highlight that integrating soil health indicators such as enzyme activities (β -glucosidase [BG] and N-acetyl- β -D-glucosaminidase [NAG]) and soil proteins into nitrogen management models improves prediction accuracy, leading to enhanced productivity and environmental sustainability. These findings suggest that advanced data-driven approaches can significantly contribute to more precise and sustainable nitrogen fertilization strategies.

Keywords: augmented extreme learning machine (AELM); feature selection; particle swarm optimization (PSO); nitrogen use efficiency (NUE); soil health indicators; corn yield



Academic Editors: Eugenija Bakšienė and Audrius Kačergius

Received: 19 December 2024

Revised: 13 January 2025

Accepted: 17 January 2025

Published: 20 January 2025

Citation: Bontemps, J.; Ebtehaj, I.; Deslauriers, G.; Rousseau, A.N.; Bonakdari, H.; Dessureault-Rompré, J. Advanced Efficient Feature Selection Integrating Augmented Extreme Learning Machine and Particle Swarm Optimization for Predicting Nitrogen Use Efficiency and Yield in Corn. *Agronomy* **2025**, *15*, 244. <https://doi.org/10.3390/agronomy15010244>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Efficient agricultural management practices are essential to ensure both high productivity and environmental sustainability, particularly in the context of nitrogen fertilization for crops such as corn. Corn (*Zea mays* L.) is an important crop in Quebec in terms of cultivated area and farm revenues. In 2020, 459,100 hectares were planted with corn (grain corn and silage corn). Nitrogen application rates are up to 240 kg/ha, significantly higher than the recommended maximum rate of 170 kg/ha [1]. This over-application of nitrogen has economic and environmental consequences, necessitating more efficient nitrogen management strategies. Approximately 50% of the nitrogen applied is in the form of urea, which

has a nitrous oxide (N₂O) emission factor of 1.1% of the applied amount [2]. This practice has a significant environmental impact, with a potential contribution of up to 100,000 tons of CO₂ equivalent emissions from corn fields in Quebec [2]. However, this estimate may be conservative, as nitrogen fertilization significantly increases N₂O emissions, especially when application rates exceed crop requirements, due to the exponential—rather than linear—relationship between nitrogen inputs and emissions [3,4].

The environmental consequences of excessive nitrogen fertilization extend beyond greenhouse gas emissions. Nitrogen losses occur through various pathways, including ammonia volatilization, denitrification, and nitrate leaching, leading to significant air and water pollution and soil degradation [5–7]. These losses are often not directly associated with a proportional increase in crop yield, highlighting the inefficiency of over-fertilization [5]. Recent studies have demonstrated that optimal nitrogen fertilization techniques, such as site-specific nutrient management, have reduced nitrogen application rates by 30–60% without sacrificing crop yields while substantially decreasing environmental impacts [8,9].

Although advancements in nitrogen management exist, a pressing need exists to incorporate soil health metrics into fertilization models adequately. Indeed, despite significant advances in fertilization technologies, current nitrogen recommendation models often fail to account for the influence of soil health on crop response to fertilization. Machine learning (ML) has emerged as a powerful tool for developing decision support systems that optimize fertilizer application and nutrient management. These techniques increase crop yields while reducing costs and environmental impacts [10]. By processing large volumes of complex, non-linear data from diverse sources, ML models build predictive frameworks for agricultural productivity [10]. In corn production, ML methodologies have been successfully applied to estimate the Economic Optimum Nitrogen Rate (EONR), addressing challenges associated with the interactions between genotype, environment, and management practices [11]. Furthermore, incorporating mechanistic hydrological features—such as water table-adjusted available water capacity and the ratio of in-season rainfall to available water capacity—enables ML models to better account for the effects of soil hydrology on nitrogen dynamics [11]. Along with sensor technology advances, ML offers cost-effective and comprehensive solutions for nutrient assessment and agricultural decision-making [10]. Soil health, including physical, chemical, biochemical, and microbiological properties, is critical in determining how effectively crops utilize nitrogen [12]. Key indicators of soil health, such as soil organic matter, microbial biomass, and soil respiration, respond rapidly to changes in soil management practices and are essential for optimizing nitrogen use efficiency [13,14]. For example, biologically healthy soils have provided higher corn yields per fertilizer unit [15]. Among these management practices, tillage is critical in shaping nitrogen dynamics by influencing soil structure, microbial activity, and nitrogen mineralization rates [16]. Conventional Plowing is associated with more significant risks of nitrogen losses through leaching or volatilization. Conversely, Reduced Tillage and Direct Drilling are associated with improved soil organic matter retention and reduced nitrogen losses [17]. Furthermore, enzyme activities, such as β -glucosidase (BG) and N-acetyl- β -D-glucosaminidase (NAG), are susceptible to soil management practices, serving as valuable indicators of how tillage affects soil nitrogen cycling and microbial activity [18–20].

In this context, many nitrogen response metrics evaluate how crops react to nitrogen fertilization, each with advantages and limitations [21–30]. Yield is one of the most widely used metrics, offering a direct measure of crop productivity by indicating the amount of grain produced per unit area. However, yield alone fails to account for the efficiency of nitrogen use, particularly the difference between fertilized and unfertilized plots. This limits our capacity to assess the actual effect of nitrogen inputs on crop performance. In contrast, agronomic nitrogen use efficiency (ANUE) considers the additional grain yield

produced per unit of nitrogen applied, providing a more accurate assessment of how effectively nitrogen contributes to increased productivity [22]. By comparing the results from fertilized and unfertilized plots, ANUE offers a more nuanced and complementary perspective on the benefits of nitrogen application. Other metrics, such as the nitrogen harvest index (NHI), provide insight into nitrogen partitioning within the plant, showing the proportion of total plant nitrogen allocated to the grain at harvest [23]. Beyond these metrics, a few have been proposed to address different aspects of nitrogen efficiency. The Apparent Nitrogen Recovery Efficiency (NRE), for instance, evaluates how much of the applied nitrogen is taken up by the plant, offering insights into nitrogen uptake efficiency across the whole plant [22]. The Physiological Nitrogen Use Efficiency (PNUE) measures how effectively nitrogen taken up by the plant is converted into grain, helping to assess nitrogen's physiological use within the crop [22]. Another critical measure is the Economic Optimal Nitrogen Rate (EONR), which calculates the amount of nitrogen required for maximum economic returns by balancing fertilization costs and crop yield [27]. All these metrics, while applicable under certain conditions, may not be entirely suited to the specific conditions of this study, which focuses on nitrogen's effects on grain yield, nitrogen allocation, and overall nitrogen use efficiency. Therefore, yield, NHI, and ANUE were selected as the most appropriate metrics to capture these aspects in our experimental setup.

Integrating soil health indicators into nitrogen management models can significantly improve fertilization practices by balancing crop productivity with environmental sustainability. Conservation agriculture practices, such as no-till systems and diverse crop rotations, increase potentially mineralizable nitrogen, which is key to improving corn yield and nitrogen use efficiency (NUE) [31]. Furthermore, incorporating soil health indicators like soil respiration and mineralizable nitrogen indicators into nitrogen recommendations can optimize economic returns and environmental outcomes [32].

This study aims to explore the interaction between soil health indicators and nitrogen fertilization in corn by applying advanced feature selection techniques. Feature selection, a key concept in ML, helps identify the most relevant variables for predicting crop performance and nitrogen use efficiency. Feature selection has long been a fundamental topic in ML research. In this context, a feature represents a metric that conveys pertinent and distinguishing details about a data item [33]. Identifying the appropriate features represents an essential stage in the development of an ML model, and this step can significantly improve the model's accuracy, rationalize its complexity, and increase its interpretability. A comprehensive set of features may include attributes that are redundant, irrelevant, weakly relevant but not redundant, and powerfully relevant [34]. Thus, the feature selection process optimizes model performance and simplifies the model, making it more accessible and easier to understand.

Through ML methods such as augmented extreme learning machine (AELM) and particle swarm optimization (PSO), we aim to refine models that predict corn yield and NUE based on comprehensive soil health and environmental variables. In doing so, we strive to develop accurate and environmentally friendly recommendations. Moreover, this research has the potential to contribute to the broader effort of sustainable agricultural management in the face of climate change and resource depletion.

2. Materials and Methods

2.1. Study Area and Experimental Design

This experiment was conducted between 2021 and 2023 across a total of 84 sites in major grain corn-producing regions of Quebec, Canada: Montérégie; Outaouais; Laurentides; Lanaudière; Centre-du-Québec; Estrie; and Capitale Nationale (Figure 1). Each site was subjected to a nitrogen fertilization trial, with five side dressing nitrogen treatments applied:

0; 50; 100; 150; and 200 kg N/ha. Additionally, 30 to 60 kg N/ha were uniformly applied at seeding across all experimental units. The experimental layout followed a randomized complete block design with three blocks per site, resulting in 15 experimental units per site.

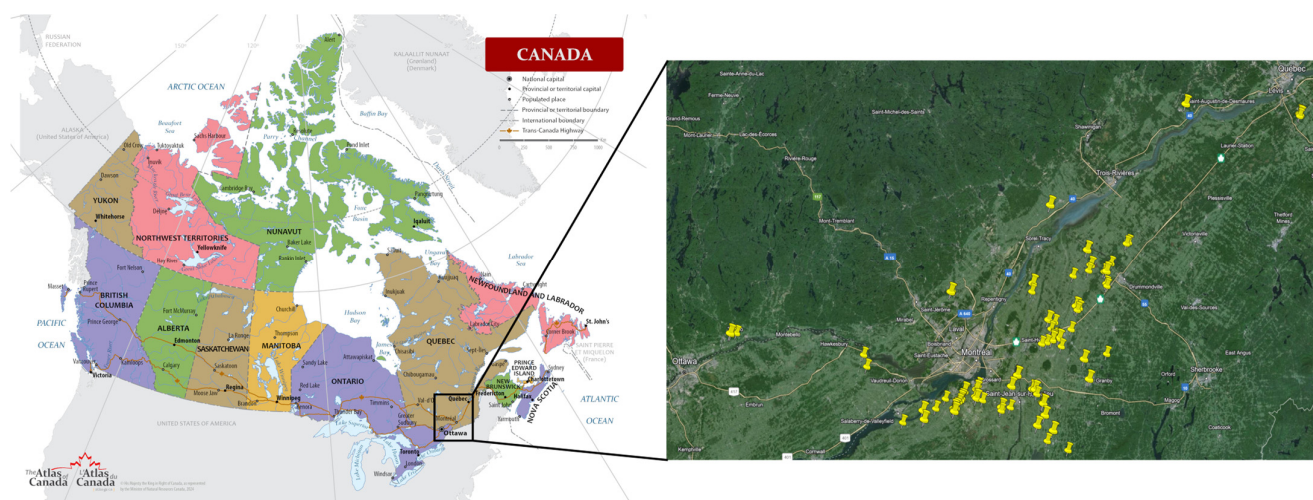


Figure 1. Geographical location of the experimental sites in Southern Quebec, Canada.

Cumulative precipitation and corn heat units (CHU) were recorded weekly and aggregated monthly from May to October for each site to account for climatic variables. An initial site characterization was performed in June before the beginning of this experiment. This characterization was conducted once per block for cost efficiency by randomly sampling soil from each block at 0 to 20 cm (Table 1).

Table 1. Inputs (1 to 38) and outputs (1 to 3) definitions.

Variable	Definition	Input/Output No.
N_Rate	Side dress N rate applied (kg/ha)	In1
Texture	Group of Soil Texture (G1, G2, G3)	In2
Tillage	Type of tillage (Conventional Plowing; Reduced Tillage; Direct Drilling)	In3
pH	Soil pH	In4
OrgMat	Organic Matter (%)	In5
P_M3	P Mehich-3 (kg/ha)	In6
K_M3	K Mehich-3 (kg/ha)	In7
Ca_M3	Ca Mehich-3 (kg/ha)	In8
Mg_M3	Mg Mehich-3 (kg/ha)	In9
Al_M3	Al Mehich-3 (mg/kg)	In10
Init_NO3	Initial soil nitrates (mg/kg)	In11
Tassel_NO3	Soil nitrates at the tasselling stage (mg/kg)	In12
Harvest_NO3	Soil nitrates at harvest (mg/kg)	In13
BD_0_15	Bulk Density 0–15 cm (g/cm ³)	In14
BD_15_30	Bulk Density 15–30 cm (g/cm ³)	In15
BD_30_45	Bulk Density 30–45 cm (g/cm ³)	In16
BG	β -glucosidase enzyme activity at harvest (nmol/g dry soil/h)	In17
Init_BG	β -glucosidase enzyme activity pre-side dress (nmol/g dry soil/h)	In18
NAG	N-acetyl- β -glucosaminidase enzyme activity at harvest (nmol/g dry soil/h)	In19
Init_NAG	N-acetyl- β -glucosaminidase enzyme activity pre-side dress (nmol/g dry soil/h)	In20
FDA	Fluorescein diacetate activity at harvest (μ g/h/kg dry soil)	In21

Table 1. Cont.

Variable	Definition	Input/Output No.
Init_FDA	Fluorescein diacetate activity pre-side dress ($\mu\text{g}/\text{h}/\text{kg}$ dry soil)	In22
RM	24 h Microbial respiration at harvest ($\text{mg C-CO}_2/\text{kg}$ dry soil/h)	In23
Init_RM	24 h Microbial respiration pre-side dress ($\text{mg C-CO}_2/\text{kg}$ dry soil/h)	In24
PRT	ACE Soil Proteins at harvest (g/kg dry soil)	In25
Init_PRT	ACE Soil Proteins pre-side dress (g/kg dry soil)	In26
CHU_May	Accumulated Corn Heat Units (May)	In27
CHU_June	Accumulated Corn Heat Units (June)	In28
CHU_July	Accumulated Corn Heat Units (July)	In29
CHU_August	Accumulated Corn Heat Units (August)	In30
CHU_September	Accumulated Corn Heat Units (September)	In31
CHU_October	Accumulated Corn Heat Units (October)	In32
Rain_May	Accumulated Rain in mm (May)	In33
Rain_June	Accumulated Rain in mm (June)	In34
Rain_July	Accumulated Rain in mm (July)	In35
Rain_August	Accumulated Rain in mm (August)	In36
Rain_September	Accumulated Rain in mm (September)	In37
Rain_October	Accumulated Rain in mm (October)	In38
Yield	Dry Yield (Humidity: 14.5%) (kg/ha)	Out1
NHI	Nitrogen Harvest Index (%)	Out2
ANUE	Agronomic Nitrogen Use Efficiency (kg grain/ kg N applied)	Out3

Physical indicators included the percentages of clay, silt, and sand used to determine soil texture, which was categorized into texture groups (G1 (47 sites), G2 (24 sites), and G3 (13 sites) corresponding to fine, medium, and coarse texture, respectively) according to CRAAQ [1]. Bulk density was also measured at three depths: 0–15 cm; 15–30 cm; and 30–45 cm. Chemical indicators included pH, organic matter content, and nutrient concentrations (P, K, Mg, Ca, and Al), all measured once per block in June. Soil nitrate (NO_3^-) concentrations were measured for each block during the post-emergence stage before the nitrogen application (between the V5 and V7 stages). Subsequently, nitrate concentrations were assessed for each treatment at the tasseling stage and harvest. The type of tillage at each site was also recorded as Conventional Plowing (8 sites), Reduced Tillage (57 sites), or Direct Drilling (19 sites) (Table 1).

Biochemical indicators included ACE soil proteins, enzyme activities associated with the nitrogen and carbon cycles (N-acetyl- β -D-glucosaminidase (NAG), β -Glucosidase (BG)), the potential of total microbial activity (fluorescein diacetate activity, FDA), and microbial respiration (RM). They were measured at harvest for each experimental unit and at the initial characterization phase before nitrogen application, with samples taken once per block (Table 1).

To evaluate crop response to nitrogen fertilization, three key metrics were calculated (Table 1): yield; nitrogen harvest index (NHI); and agronomic nitrogen use efficiency (ANUE).

- Yield was measured as the dry grain mass at 14.5% moisture content (kg/ha) harvested from each plot. This serves as a direct indicator of crop productivity and the overall effectiveness of nitrogen fertilization. Yield is one of the most used metrics in agricultural studies as it provides a straightforward measure of the output in relation to the input;

- N Harvest Index (*NHI*) was calculated as the ratio of nitrogen content in the grain (N_g) to the total nitrogen content in the aerial biomass (N_a) at harvest. This index offers insights into the efficiency with which the plant allocates nitrogen to grain production, which is particularly important for crops like corn, where the grain is the primary product. A higher *NHI* indicates that a more significant proportion of the nitrogen taken up by the plant is stored in the grain, reflecting better nitrogen use for reproductive purposes [23] and calculated as follows:

$$NHI = N_g / N_a \quad (1)$$

- Agronomic nitrogen use efficiency (ANUE) was calculated using the following formula:

$$ANUE = (Y_F - Y_{UF}) / N_{Rate} \quad (2)$$

where Y_F is the grain yield from the fertilized plots; Y_{UF} is the grain yield from the control (unfertilized) plots, and N_{Rate} is the amount of nitrogen applied in the fertilized plots [22]. This metric measures the additional grain yield obtained per unit of nitrogen applied, providing a valuable metric for assessing both agronomic and economic efficiency of nitrogen use. ANUE highlights how much extra yield is produced because of nitrogen fertilization, which is critical in determining the optimal fertilizer application rate.

These metrics provide a detailed assessment of corn's productivity (via yield) and nitrogen-use efficiency (via *NHI* and ANUE) under different fertilization treatments. This multidimensional evaluation allows for a more comprehensive understanding of the impact of nitrogen on crop performance, considering not only the total output but also how efficiently the plant utilizes nitrogen.

2.2. Augmented Extreme Learning Machine (AELM)

Extreme learning machine (ELM) [35] is an algorithm for single-hidden layer feed-forward neural networks (SLFFNNs). ELM is known for its ability to learn extremely fast compared to traditional gradient-based learning algorithms. ELM's design principle separates the learning process into two distinct steps: random initialization of hidden neurons and a closed-form solution for output weights. This approach eliminates the need for iterative tuning of parameters, making the learning process highly efficient. Indeed, ELM randomly assigns the input weights and biases and then analytically determines the output weights, making the learning process extremely fast.

A conceptual framework of ELM is provided in Figure 2. Based on this figure, by defining the number of input variables (NIV) and the number of hidden neurons (NHN), the $NHN \times NIV$ input weights matrix and the $NHN \times 1$ bias matrix of hidden neurons dimensions are randomly produced. Using these two matrices along the activation function, a $NIV \times NHN$ matrix is generated. Finally, the output weight is calculated analytically using this latter matrix and the matrix of the target samples. To be more precise, the mathematical formulation of the ELM is described in the next paragraph.

Imagine a SLFFNN with NIV input features, a single output neuron, and NHN hidden neurons. The input matrix X dimensions $NS \times NIV$ with NS representing the number of samples is transformed to the hidden layer using a randomly initialized bias matrix of hidden neurons (BHN) with dimensions of $1 \times NHN$ and an input weight matrix (lnW) with dimensions $NIV \times NHN$ [35].

$$M = f(X \times lnW + BHN) \quad (3)$$

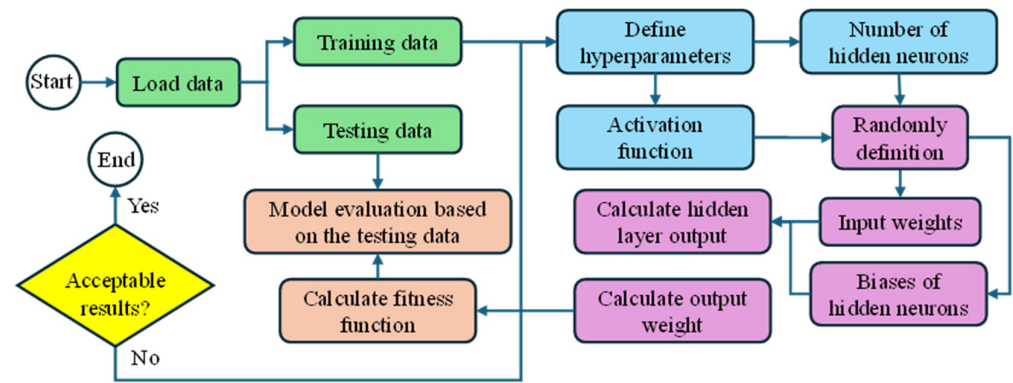


Figure 2. Conceptual framework of the ELM.

In this context, BHN represents a bias vector of dimensions NHN that is also randomly allocated, and M denotes the hidden layer output matrix with dimensions $NIV \times NHN$.

The output weights O , which link the hidden layer to the output layer, have dimensions $NHN \times 1$. Instead of being learned via backpropagation in conventional feedforward neural network (FFNN), these weights are calculated analytically by solving the following linear system [35]:

$$MO = Y \quad (4)$$

Here, Y is the target output matrix with dimensions $N \times 1$.

One usually employs the Moore–Penrose pseudoinverse of M to find the solution for M , symbolized as M^+ . This approach yields the least-squares solution to the previously mentioned linear system [35].

$$O = M^+Y \quad (5)$$

The Moore–Penrose pseudoinverse (MPPI) finds the best-fit solution to a linear system by minimizing the sum of the squared differences (errors) between the observed and predicted values. It works by extending the concept of an inverse matrix to non-square or singular matrices. The MPPI can handle cases where the matrix has no regular inverse, providing the least-squares solution that minimizes the error. The pseudoinverse M^+ is calculated as follows [36]:

$$M^+ = \left(M^T M\right)^{-1} M^T \quad (6)$$

ELM offers numerous advantages, making it a powerful tool for various ML applications. One of its key benefits is that it avoids becoming stuck in local minima, ensuring a more reliable convergence to a global optimum compared to traditional neural networks [36]. This is particularly advantageous in complex problem spaces where local minima impede learning. ELM also provides excellent controllability [37], allowing users to adjust parameters to optimize performance for specific tasks efficiently. Its fast learning rate is another significant advantage [38,39], as ELM can complete training much quicker than conventional methods, making it suitable for real-time applications. Furthermore, ELM boasts high accuracy in its predictions, enabling it to perform well across various datasets and problem types. Its high generalization capability ensures that the model performs well on training and unseen data, reducing the risk of overfitting [40]. The learning process of ELM requires only a single iteration [41], simplifying the training procedure and reducing computational complexity. This aspect, combined with the need for low user intervention [42], makes ELM a user-friendly option that does not demand extensive tuning or expert knowledge to achieve good performance. Lastly, ELM is known for its robustness [43], maintaining stability and effectiveness even in the presence of noise or outliers in the data. These advantages collectively make ELM an appealing choice for many

ML practitioners. However, there is a drawback to this method that can negatively affect its performance.

Based on the provided details, the two matrices, input weights and biases of hidden neurons (BHN), are initialized randomly. In the case of an ELM with at least one input feature ($NIV = 1$), 66% of the parameters are assigned randomly [41]. This random assignment can cause issues, necessitating a more refined method. To tackle this problem, this study introduces a repetitive approach called the Augmented Version of the ELM (AELM). This new method aims to enhance the reliability and performance of the ELM by addressing the randomness in parameter initialization. Given the highly rapid training time of the ELM, the classical ELM is executed repeatedly in an iterative process. The final model is selected based on its best performance on unseen data.

2.3. Particle Swarm Optimization (PSO)

The particle swarm optimization (PSO) algorithm [43] is a population-based, stochastic optimization technique [44]. It is modeled after the social behaviors of animals, such as fish schools and bird flocks [45]. In their quest for food, these swarms follow a cooperative strategy where individuals continuously modify their search patterns based on interactions with other swarm members and personal experiences. The fundamental principle of the PSO algorithm is inspired by artificial life theory and uses swarm intelligence to explore vast solution spaces effectively during the optimization process.

To examine the behavior of artificial systems with lifelike characteristics within PSO as a swarm-based algorithm, the following fundamental principles have been considered for effective computer simulation [46,47]:

1. **Adaptability:** The swarm should be capable of modifying its behavior when such changes are warranted;
2. **Diverse Response:** The swarm should avoid restricting its search path to a limited area to maximize resource acquisition;
3. **Proximity:** The swarm should efficiently perform time-based calculations and navigate simple spaces;
4. **Quality Recognition:** The swarm should be able to detect and respond to quality variations in the environment;
5. **Stability:** The swarm should maintain consistent behavior despite minor environmental changes.

It is important to note that the principles of adaptability and stability are two sides of the same coin. The PSO algorithm updates particle positions and velocities in response to environmental changes. This ensures that PSO meets the principles of proximity and quality. Additionally, PSO swarms have unrestricted movement, enabling them to search for optimal solutions within the solution space continually. Particles in the PSO algorithm maintain a stable trajectory while adapting to environmental changes. As a result, PSO-based particle swarms adhere to all five specified principles.

Each group member is called a particle in the PSO as a swarm-based search technique [48,49]. This particle represents a potential solution to the optimization problem at hand, and it can remember both the optimal velocity of the swarm and its position within the search space. In each generation, the updated position of every particle is computed by altering the velocity along each dimension. This adjustment is achieved by integrating the information from all the particles in the swarm. Particles continuously adjust their states within the D-dimensional search space until they reach an optimal or balanced condition or exceed computational limits. The objective function serves to link all the dimensions within the problem space.

Consider N as the size of the swarm. Each particle's position and velocity are represented by $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ and $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, respectively. The best position found by the entire swarm is indicated as $P_g = (p_{g1}, p_{g2}, \dots, p_{gD})$. On the other hand, the best position found by each particle, or the particle's optimal position, is denoted as $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$. These positions are pbest (personal best) for P_i and gbest (global best) for P_g . The following formula determines the optimal position for everyone [43]:

$$P_{i,t+1}^d = \begin{cases} x_{i,t+1}^d & \text{if } f(X_{i,t+1}) < f(P_{i,t}) \\ P_{i,t}^d & \text{otherwise} \end{cases} \quad (7)$$

The velocity and position vectors are updated as follows [43]:

$$v_{i,t+1}^d = w \times v_{i,t}^d + c_1 \times \text{rand} \times (p_{i,t}^d - x_{i,t}^d) + c_2 \times \text{rand} \times (p_{g,t}^d - x_{i,t}^d) \quad (8)$$

$$x_{i,t+1}^d = x_{i,t}^d + v_{i,t+1}^d \quad (9)$$

Here, t represents the iteration number (ranging from 1 to the maximum iteration number $- 1$), and i indicates the particle number (ranging from 1 to N). The term rand signifies a random value in the interval $[0, 1)$, while c_1 and c_2 are personal and global learning coefficients, respectively. P_g and P_i are known as "gbest" and "pbest", respectively, and w denotes the inertia weight. Equation (8) consists of three components. The first component represents the influence of the particle's previous velocity. The second component is based on the particle's own experience and knowledge. The experiences of other particles influence the third component in the swarm and the current particle's experience. This algorithm has a significant advantage over other minimization strategies because the large number of swarming particles helps the method resist the local minimum problem [50]. More detailed information about the PSO can be found in the literature studies [51,52].

2.4. AELMPSO-Based Feature Selection Approach

In general, feature selection is utilized to identify the most impactful relevant variables for constructing models. The primary motivations for employing feature selection are (1) simplifying models to make them easier for researchers and users to interpret [53], (2) mitigating the curse of dimensionality, (3) reducing CPU usage, and (4) minimizing overfitting to improve generalizability [54].

This study selected 38 independent variables to estimate yield, NHI, and ANUE (see Table 1). The optimal sub-features from 1 to 38 were calculated using the binomial coefficient C , Equation (10), which represents the number of ways to choose M elements from a set of N elements without regard to the selection order.

$$C = \frac{N!}{M!(N - M)!} \quad (10)$$

The "!" (factorial) denotes the product of all positive integers up to a given number so that the $N!$, $M!$, and $(N - M)!$ are the factorials of N , M , and $N - M$, respectively.

Utilizing the previously mentioned equation, Table 2 displays the number of distinct sub-features for feature sets ranging from 1 to 38. A total of 5.49756×10^{11} sub-features must be evaluated to identify the optimal subsets for feature counts between 1 and 38. Consequently, using feature selection techniques is necessary to efficiently navigate and determine the best sub-features from this vast number of possible combinations. The best subsets ranging from one to 38 features are selected using the developed AELMPSO-based feature selection method to determine the optimal combinations of these features.

Table 2. Calculation of all possible sub-features for 1 to 38 features.

No. of Inputs	No. of Sub-Features	No. of Inputs	No. of Sub-Features	No. of Inputs	No. of Sub-Features	No. of Inputs	No. of Sub-Features
1	3.90×10^1	11	1.68×10^9	21	6.24×10^{10}	31	6.15×10^7
2	7.41×10^2	12	3.91×10^9	22	5.10×10^{10}	32	1.54×10^7
3	9.14×10^3	13	8.12×10^9	23	3.77×10^{10}	33	3.26×10^6
4	8.23×10^4	14	1.51×10^{10}	24	2.51×10^{10}	34	5.76×10^5
5	5.76×10^5	15	2.51×10^{10}	25	1.51×10^{10}	35	8.23×10^4
6	3.26×10^6	16	3.77×10^{10}	26	8.12×10^9	36	9.14×10^3
7	1.54×10^7	17	5.10×10^{10}	27	3.91×10^9	37	7.41×10^2
8	6.15×10^7	18	6.24×10^{10}	28	1.68×10^9	38	3.90×10^1
9	2.12×10^8	19	6.89×10^{10}	29	6.36×10^8	Sum	5.50×10^{11}
10	6.36×10^8	20	6.89×10^{10}	30	2.12×10^8		

This subsection presents the developed AELMPSO techniques for feature selection in predicting ANUE and yield in corn. Figure 3 reveals the flowchart of the developed AELMPSO for feature selection. The modeling process begins by loading the training and testing datasets. The samples for estimating yield, NHI, and ANUE are 1171, 1142, and 934, respectively. Eighty percent of all samples were selected for model calibration (training dataset), while the remaining 20% were reserved for model validation (testing dataset) [40]. Additionally, k-fold cross-validation is utilized to prevent overfitting and assess the generalizability of the proposed models. The value of k is set to 5. In this method, all samples were randomly divided into five groups. Four groups were used for training for each iteration of the modeling process, while the remaining group was used for testing. This procedure is repeated five times, ensuring that each sample serves as a test sample at least once.

The next step was to define the parameters for PSO and AELM, the two techniques used to identify the optimal feature combinations. For the PSO evolutionary algorithm, the parameters that need to be set before modeling include inertia weight, population size, number of iterations, and global and personal learning coefficients (c_1 and c_2 in Equation (8)). Grid search was used to determine the optimal values of the hyperparameters. It is a systematic hyperparameter optimization technique to find a model's optimal set of parameters. It exhaustively searches through a predefined set of hyperparameter combinations to evaluate their performance. The grid search process involves three main steps. First, the hyperparameter grid is defined by specifying a range of values for each hyperparameter. Next, the model is trained and evaluated for every possible combination of hyperparameters within the grid, with each combination assessed based on a chosen performance metric. Finally, the hyperparameter set that delivers the best performance on the validation dataset is selected as the optimal configuration. This systematic approach ensures the identification of the most effective hyperparameters for the model. Through a grid search process, these parameters were set as follows: inertia weight at 0.9; population size at 200; number of iterations at 100; and both learning coefficients (c_1 and c_2 in Equation (8)) at 2.05. These values were determined using the grid search approach. Grid search is a systematic approach for hyperparameter optimization in machine learning. It involves specifying a set of hyperparameters and a range of values for each one and then evaluating the model's performance for each combination of these values. The goal is to identify the combination of hyperparameters that results in the best performance according to predefined metrics, such as accuracy or mean squared error.

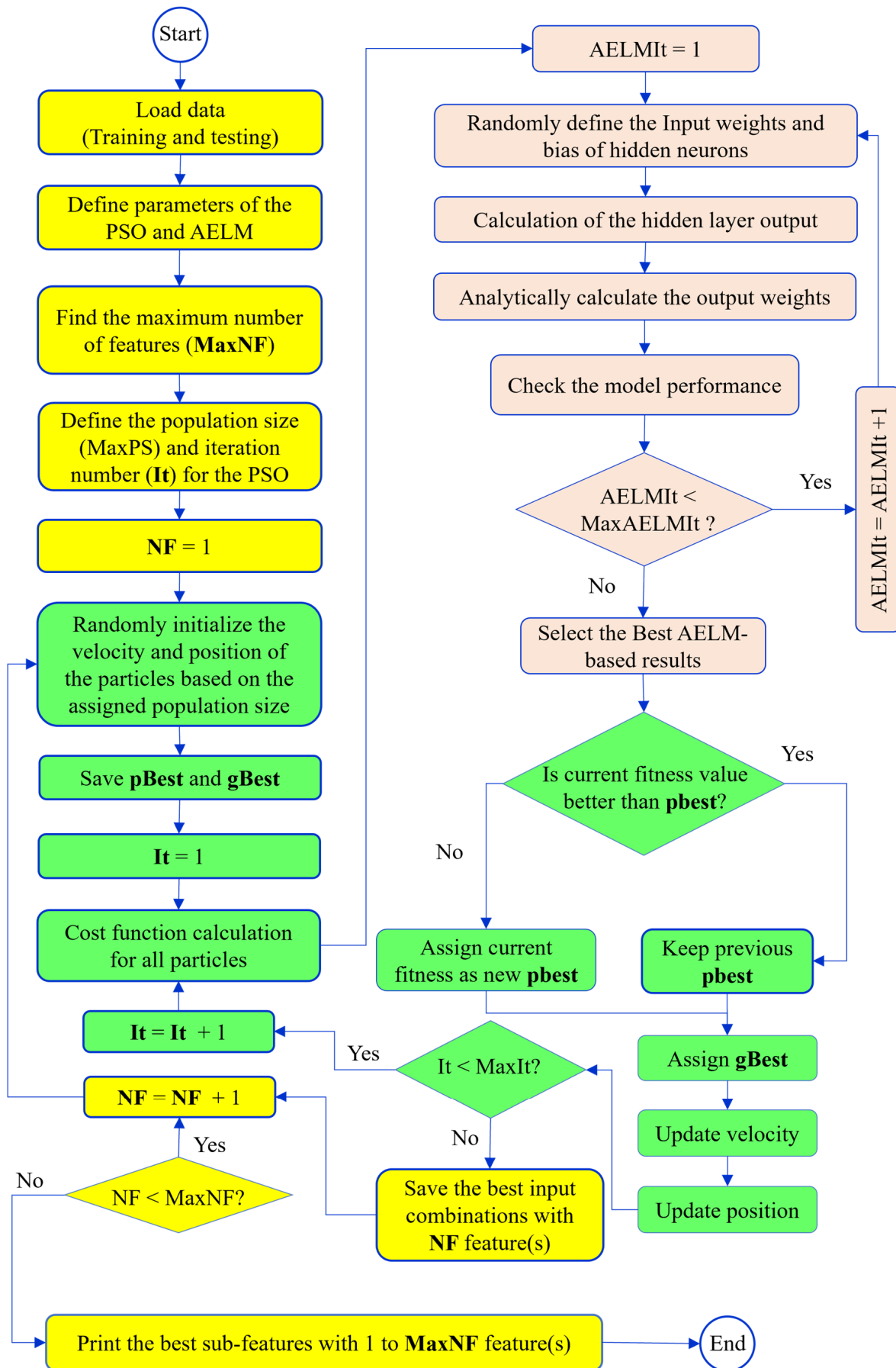


Figure 3. Flowchart of the developed AELMPSO for feature selection.

Furthermore, the AELM parameters included the activation function, the number of hidden neurons, and the number of iterations. The activation function chosen for this study was the Sigmoid function, which has displayed superior performance in the literature [39,55,56]. Its performance was also compared with other functions, such as the tangent hyperbolic, radial basis function, triangular function, hard limit, and sine function; it was found to be more effective. The number of iterations for the AELM was set to 1000, determined through a trial-and-error process. The number of hidden neurons was calculated based on the following equation [57,58]:

$$NHN_{max} \leq \frac{NTrS}{NIV + 2} \quad (11)$$

where NHN_{max} is the maximum number of allowable hidden neurons; $NTrS$ is the number of training samples, and NIV is the number of input variables.

After defining the parameters, the leading training loop began, iterating from NF to MaxNF, where NF represents the number of features. In this study, NF is set to 38, so the process was repeated 38 times. The following explanations pertain to a single iteration. The first step involved the random initialization of the velocity and position of the particles based on the predefined population size. Following this, the best personal and global particles were selected, as the cost function values from this step, which were randomly initialized, are better than the initial cost function set to infinity. After the initialization process in the PSO, the algorithm's main evolutionary optimization steps begin. The AELM was called upon to calculate the cost function within the algorithm. As previously mentioned, the AELM also involved an iteration process, denoted by AELMI_t. In the first step of the AELM, the input weights and biases of the hidden neurons were randomly initialized. The hidden layer output matrix was calculated considering these two matrices and the activation function. Subsequently, the Moore-Penrose pseudoinverse of this newly generated matrix and the output variable matrix were used to calculate the output weight matrix. The performance of the model was then evaluated. This process was repeated until the maximum iteration count defined for AELM (MaxAELMI_t) was reached. Then, the best results of the AELM were reported as the cost function for this iteration.

In the next step of the evolutionary optimization of the PSO, the fitness of the current iteration was compared with the existing personal best (pBest). If the current fitness was lower, the personal best was updated; otherwise, the current personal best remained unchanged. Based on the new pBest, the global best (gBest) was also recalculated. Next, the velocities and corresponding positions of the particles were updated. This process continued for all iterations specified by the iteration number for the PSO (i.e., MaxIt). Once this iterative process was complete, the feature selection process for identifying the best sub-feature set with NF features was concluded. Subsequently, the entire process was repeated for each value of NF until MaxNF was reached.

2.5. Quantitative Metrics for Model Assessment

The statistical metrics used in this study included the coefficient of determination (R^2), Nash–Sutcliffe Efficiency (NSE), Percentage BIAS (PBIAS), and Normalized Root Mean Square Error (NRMSE). The R^2 metric measures the proportion of the variance in the dependent variable that is predictable from the independent variables. An R^2 value closer to 1 indicates a better fit of the model to the data. NSE assesses the predictive power of machine learning models. It ranges from $-\infty$ to 1, with values closer to 1 indicating better predictive performance. Negative values suggest that the mean observed value is a better predictor than the model. PBIAS measures the average tendency of the simulated data to be larger or smaller than their observed counterparts. It is expressed as a percentage. A PBIAS

of 0 indicates an ideal model, with negative values indicating overestimation and positive values indicating underestimation. RMSE measures the average magnitude of the errors between predicted and observed values without considering their direction. It clarifies the precision of the model, with lower values indicating better model performance. NRMSE is the RMSE normalized by the average of the observed data, allowing for comparison across different datasets or models. Lower values indicate better model performance and allow for a more straightforward interpretation of model accuracy in relation to the observed data. These metrics collectively provide a comprehensive evaluation of the model's performance [59–61], capturing various aspects of accuracy, precision, bias, and efficiency. Using these diverse metrics, this study ensured a robust assessment of the model's predictive capabilities, facilitating a thorough understanding of its strengths and limitations. The mathematical definitions of these metrics are as follows, and their descriptive performance is provided in Table 3 [62].

$$R^2 = \frac{\sum_{i=1}^{NS} (O_i - \bar{O})^2 (P_i - \bar{P})^2}{\sum_{i=1}^{NS} (O_i - \bar{O})^2 \sum_{i=1}^{NS} (P_i - \bar{P})^2} \quad (12)$$

$$NSE = 1 - \frac{\sum_{i=1}^{NS} (O_i - P_i)^2}{\sum_{i=1}^{NS} (O_i - \bar{O})^2} \quad (13)$$

$$PBIAS = 100 \times \frac{\sum_{i=1}^{NS} (O_i - P_i)}{\sum_{i=1}^{NS} (O_i)} \quad (14)$$

$$NRMSE = \frac{\sum_{i=1}^{NS} (O_i - P_i)^2}{\sum_{i=1}^{NS} O_i} \quad (15)$$

Table 3. Descriptive performance of statistical metrics.

	Index	R^2	NSE	PBIAS	NRMSE
Descriptive performance	Unsatisfactory	$R^2 < 0.5$	$NSE < 0.4$	$PBIAS > \pm 25\%$	$30\% < NRMSE$
	Acceptable	-	$0.4 < NSE < 0.5$	-	-
	Satisfactory	$0.5 < R^2 < 0.6$	$0.5 < NSE < 0.65$	$\pm 15\% > PBIAS > \pm 25\%$	$20\% < NRMSE < 30\%$
	Good	$0.6 < R^2 < 0.7$	$0.65 < NSE < 0.75$	$\pm 10\% > PBIAS > \pm 5\%$	$10\% < NRMSE < 20\%$
	Very Good	$0.7 < R^2 < 1$	$0.75 < NSE < 1$	$\pm 5\% > PBIAS$	$NRMSE < 10\%$

The “-” symbol indicates that no threshold is defined for the corresponding metric at the specified performance level. For example, “Acceptable” performance does not have predefined thresholds for R^2 , PBIAS, or NRMSE.

3. Results

3.1. Descriptive Statistics

Given the detailed nature of the descriptive statistics, both the table of descriptive statistics (Supplementary Table S1) and the correlation matrices (Supplementary Figure S1) for each output variable (yield, NHI, and ANUE) are included in the Supplementary Materials.

The dataset displayed a broad range of variability and skewness across inputs, highlighting several key characteristics relevant to model performance. For instance, P_M3 exhibited a mean of 82.23 kg/ha with a standard deviation of 81.09 kg/ha, reflecting substantial variability and a high kurtosis of 10.18, indicative of extreme values. Harvest_NO3, with a mean of 20.59 mg/kg and kurtosis of 16.06, displayed significant right-skewness, suggesting the presence of potential outliers that may impact predictive stability. Factors such as Init_FDA exhibited high variability among the input variables, with a mean of 20,659.75 $\mu\text{g}/\text{h}/\text{kg}$ and a standard deviation of 14,690.79 $\mu\text{g}/\text{h}/\text{kg}$. Similarly, BG showed a mean value of 18.46 nmol/g/h with a standard deviation of 16.04 nmol/g/h and a skew-

ness of 2.07, emphasizing its right-skewed distribution with kurtosis of 5.37. These values highlight the variability and the influence of extreme values in soil biochemical activity. Conversely, variables such as BD_0_15 (mean: 1.33 g/cm³, standard deviation: 0.15 g/cm³) and RM (mean: 3.70 mg C-CO₂/kg/h, standard deviation: 1.29 mg C-CO₂/kg/h) showed relative stability across samples, reflecting consistent properties within the dataset. Environmental variables such as cumulative heat units (CHU) and rainfall varied across months. For instance, Rain_June had a mean of 109.57 mm with a standard deviation of 59.19 mm. At the same time, CHU_July displayed a mean of 782.91 and a standard deviation of 74.10, demonstrating variability in climatic factors that necessarily affect crop development (Supplementary Table S1).

Yield, one of the metrics, exhibited a mean of 11,419.32 kg/ha with a standard deviation of 2460.29 kg/ha, indicating moderate variability and a slight left-skewness (skewness: -0.80). NHI demonstrated a mean of 0.53 with notable variability (range of values: 0.06 to 0.91 and standard deviation: 0.19) and slight left-skewness (-0.60), suggesting consistent nitrogen allocation efficiency. ANUE also exhibited notable variability with a mean of 23.18 kg of grain/kg N, a standard deviation of 22.09 kg of grain/kg N, a range of values from -80.96 kg of grain/kg N to 145.83 kg of grain/kg N, and a high kurtosis of 3.87, highlighting potential outliers and a concentration of values at the higher end (Supplementary Table S1).

The generally low correlations observed (e.g., $r < 0.3$ for most variable pairs) suggest minimal linear associations, which supports the decision to apply advanced modeling techniques to capture meaningful patterns within this diverse dataset (Supplementary Figure S1).

3.2. Feature Selection Results

Increasing the number of input variables, or features, in an ML model introduces several potential risks that might impact performance and efficiency. The most significant risk is overfitting, where the model learns the noise, random fluctuations in the training data, and underlying patterns. This tends to occur when the model is too complex relative to the informativeness and amount of training data. This leads it to capture spurious relationships that do not generalize well to new data. Additionally, more input variables increase model complexity, potentially resulting in longer training times and requiring more computational resources, which is particularly challenging with large datasets or limited computational resources. Another risk is diminishing returns; as more variables are added, the incremental gain in performance can decrease, and after a certain point, additional features may contribute little to improving performance and might even degrade it due to increased noise and complexity. Not all features are equally informative; some might be redundant, introducing unnecessary complexity into the model and increasing the time to identify and remove such redundancies. Furthermore, more features can complicate the interpretability of the model, as models with fewer input variables are generally easier to explain and understand. After evaluating the input combinations with higher features, a maximum of eight features for each output variable was considered to mitigate these risks.

Figure 4 reveals the results of the feature selection for models using three to eight features (feature definitions are presented in Table 1). For yield, the inputs for M4 to M8 include input values 1 (N_Rate) and 37 (Rain_September), indicating these inputs might be significant for the output. Moreover, input 36 (Rain_August) is consistently included from M7 to M8, suggesting its potential importance as the model complexity increases.

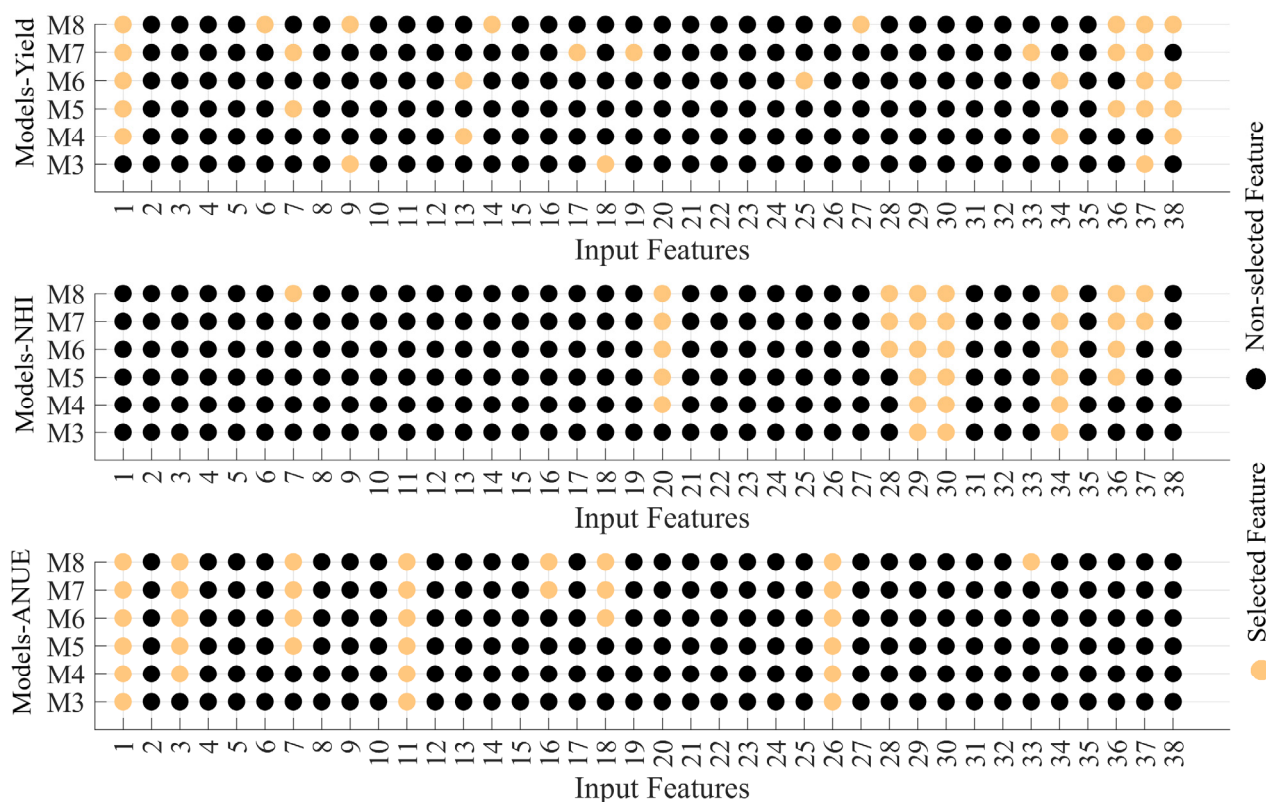


Figure 4. Results of the feature selection.

For NHI, the complexity also increases from M3 to M8, with each model using more inputs. Additionally, Inputs 29 (CHU_July), 30 (CHU_August), and 34 (Rain_June) are consistently present from M3 to M8, indicating their importance for NHI. Inputs 20 (Init_NAG) and 36 (Rain_August) are included from M4 to M8, showing their significance. Importantly, Init_NAG is not typically included in recommendations, yet its consistent inclusion in higher-performing models underscores its potential utility. Moreover, Input 28 (CHU_June) appears from M6 to M8, suggesting that it becomes relevant as the models become more complex.

For ANUE, Inputs 1 (N_Rate), 11 (Init_NO3), and 26 (Init_PRT) are consistently present across all models, indicating their crucial role for ANUE. Moreover, Input 3 (Tillage) is added from M4 to M8, showing its importance. In addition, Inputs 7 (K_M3), 16 (BD_30_45), and 18 (Init_BG) are included in M5 to M8, suggesting their increasing relevance with model complexity. Init_BG, another variable not traditionally considered in recommendations, is consistently associated with superior performance models. Furthermore, Input 33 (Rain_May) is only present in M8, indicating that it might be less critical than others but still relevant for the most complex model.

3.3. AELM-Based Modeling

Figure 5 and Table 4 present the performance evaluation of the AELM-based models in predicting yield, NHI, and ANUE for models using the three to eight features selected. In all figures, the outermost circle represents the statistical metrics for yield, the middle circle for NHI, and the innermost circle for ANUE. The statistical metrics for models M3 to M8, which utilize the AELM approach to estimate yield, NHI, and ANUE, provide an insightful overview of their performance during the training and testing stages. By examining these metrics— R^2 , NRMSE, NSE, and PBIAS—the efficiency of each model can be understood, and the most dependable and accurate models in their estimations can be determined.

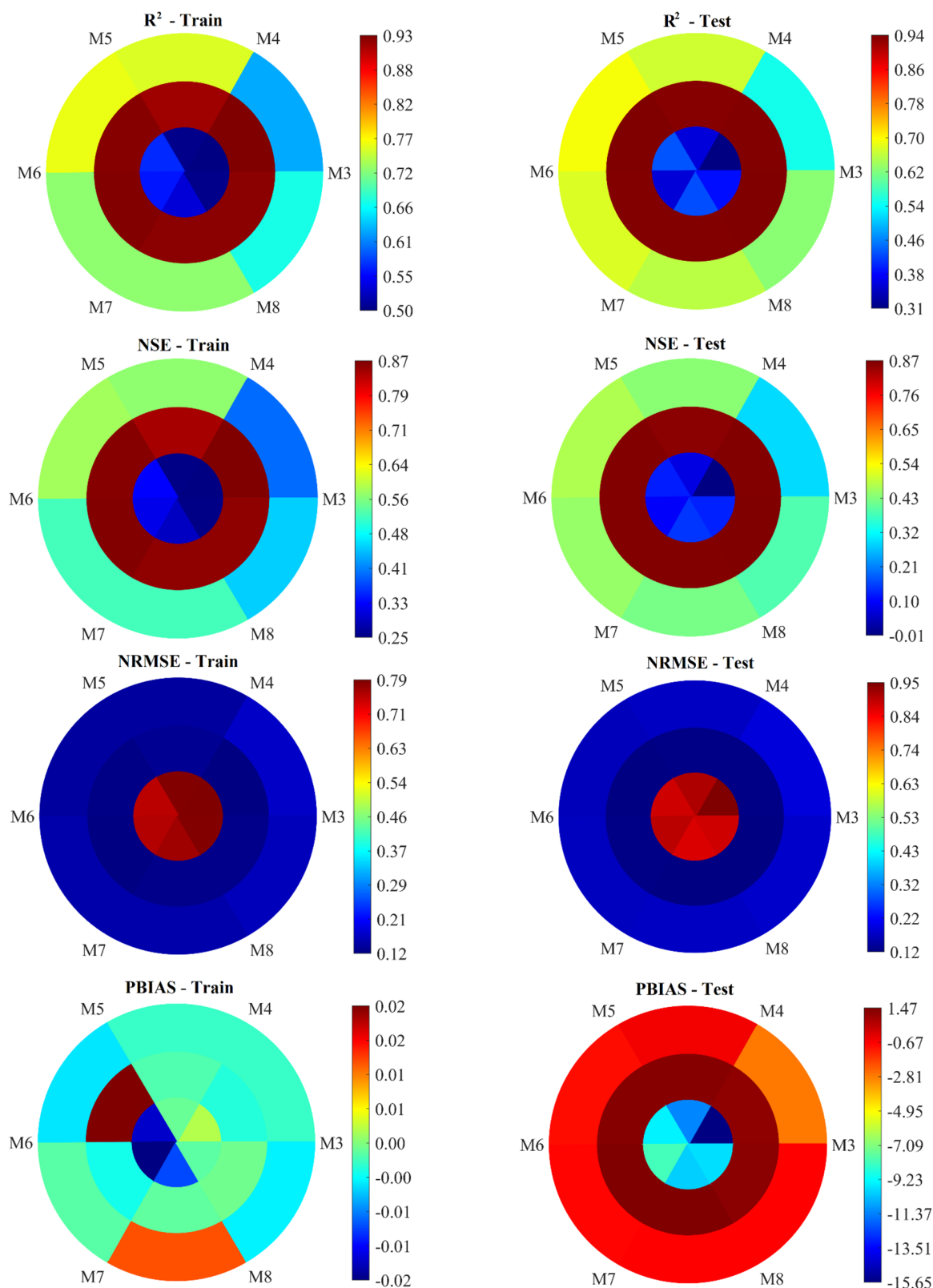


Figure 5. Performance evaluation of the AELM-based models for predicting yield (outermost circle), NHI (middle circle), and ANUE (innermost circle) across models with three to eight features. The charts are read counterclockwise, where the segment between two model labels (e.g., M3 and M4) represents the value associated with the first model (M3). The color scale indicates the performance values varying from low (blue) to high (red) based on the provided gradient legend.

Table 4. Performance Metrics Summary for AELM-based models of yield, NHI, and ANUE.

Metric	Model	R^2		NSE		NRMSE		PBIAS	
		Train	Test	Train	Test	Train	Test	Train	Test
Yield	M3	0.627	0.554	0.393	0.289	0.168	0.189	-1.06×10^{-3}	-2.71×10^0
	M4	0.754	0.669	0.569	0.441	0.141	0.168	-1.08×10^{-3}	-4.56×10^{-1}
	M5	0.762	0.694	0.580	0.473	0.139	0.163	-4.02×10^{-3}	-8.87×10^{-1}
	M6	0.723	0.677	0.522	0.453	0.149	0.166	-7.02×10^{-6}	-7.50×10^{-1}
	M7	0.723	0.657	0.522	0.425	0.149	0.170	1.26×10^{-2}	-6.58×10^{-1}
	M8	0.672	0.628	0.452	0.388	0.159	0.175	-3.57×10^{-3}	-7.05×10^{-1}
NHI	M3	0.933	0.933	0.870	0.869	0.122	0.120	-1.82×10^{-3}	1.18×10^0
	M4	0.920	0.931	0.846	0.864	0.132	0.122	-1.70×10^{-4}	1.36×10^0
	M5	0.930	0.935	0.865	0.872	0.124	0.118	2.01×10^{-2}	1.35×10^0
	M6	0.931	0.934	0.867	0.869	0.123	0.119	-2.47×10^{-3}	1.47×10^0
	M7	0.928	0.934	0.861	0.871	0.126	0.119	2.94×10^{-4}	1.46×10^0
	M8	0.927	0.936	0.860	0.874	0.126	0.117	8.88×10^{-4}	1.30×10^0
ANUE	M3	0.501	0.306	0.251	-0.009	0.794	0.945	3.40×10^{-3}	-1.56×10^1
	M4	0.504	0.360	0.254	0.073	0.792	0.906	6.03×10^{-4}	-1.13×10^1
	M5	0.570	0.434	0.325	0.125	0.753	0.881	-1.44×10^{-2}	-9.50×10^0
	M6	0.561	0.358	0.315	0.095	0.759	0.896	-1.69×10^{-2}	-8.16×10^0
	M7	0.537	0.429	0.289	0.144	0.774	0.871	-9.84×10^{-3}	-1.01×10^1
	M8	0.503	0.390	0.253	0.127	0.793	0.879	1.63×10^{-4}	-9.79×10^0

3.3.1. Results Analysis: AELM-Based Yield Estimation

Model M3 (Figure 5) shows a satisfactory performance during both the training and testing phases, albeit leaning toward the lower end of the satisfactory range. The R^2 values of 0.63 during training and 0.55 during testing place it within the satisfactory category, indicating that it explains over half of the variance in yield. However, its NRMSE values (0.17 for training and 0.19 for testing) suggest that the residuals are moderate, falling into the satisfactory range during testing but reaching acceptable limits during training. The NSE values (0.39 for training and 0.29 for testing) indicate unsatisfactory performance, below 0.4, especially during testing. The PBIAS values (-0.00 for training and -2.71 for testing) are close to zero, indicating minimal bias. In summary, M3 demonstrates limited predictive power, particularly in testing, suggesting potential overfitting or insufficient input features.

Model M4 (Figure 5) performs better than M3 across all metrics. Its R^2 values (0.75 for training and 0.67 for testing) fall into the good range, showing a significant ability to explain the variance in yield. The NRMSE values (0.14 for training and 0.17 for testing) are better than those of M3, indicating lower normalized residual errors and better accuracy. The NSE values (0.57 for training and 0.44 for testing) suggest satisfactory performance during testing and good performance during training. With PBIAS values of -0.00 for training and -0.46 for testing, M4 shows minimal bias, further affirming its reliability. Overall, M4 is a strong model with good generalization capability.

Model M5 (Figure 5) exhibits characteristics similar to those of M4, showing slight improvements in some areas. Its R^2 values (0.76 for training and 0.69 for testing) are higher, placing it solidly in the Good category and indicating a robust capacity to capture variance in yield. The NRMSE values (0.14 for training and 0.16 for testing) are marginally better than those of M4, suggesting slightly higher precision. The NSE values (0.58 for training and 0.47 for testing) demonstrate consistent good performance, particularly during training. However, the PBIAS values (-0.00 for training and -0.89 for testing) are slightly worse than M4, indicating a minor increase in bias during testing. Despite this, M5 remains a reliable model with strong generalization, marginally outperforming M4.

Model M6 (Figure 5) maintains high performance, although it shows a slight drop compared to M4 and M5. The R^2 values (0.72 for training and 0.68 for testing) are within the Good range, like the previous models. Its NRMSE values (0.15 for training and 0.17

for testing) are slightly higher than those of M5, suggesting a minor increase in error. The NSE values (0.52 for training and 0.45 for testing) indicate satisfactory performance. The PBIAS values (-0.00 for training and -0.75 for testing) remain low, demonstrating minimal bias. Although slightly less precise than M5, M6 performs reliably with minimal bias and good generalization.

Model M7 (Figure 5) shows a slight decline in performance compared to M5 and M6. The R^2 values (0.72 for training and 0.66 for testing) remain in the good range but are on the lower end. Its NRMSE values (0.15 for training and 0.17 for testing) indicate satisfactory performance but slightly higher errors than M6. The NSE values (0.52 for training and 0.42 for testing) suggest satisfactory performance during testing. Notably, the PBIAS values (0.01 for training and -0.66 for testing) are higher than those of M6, indicating a marginal increase in bias. Despite this, M7 still demonstrates decent performance but with slightly reduced accuracy and increased bias compared to M5 and M6.

Model M8 (Figure 5) shows a notable decline compared to M4, M5, and M6. The R^2 values (0.67 for training and 0.63 for testing) are still in the good range but lower than the previous models. The NRMSE values (0.16 for training and 0.18 for testing) are higher, indicating increased errors. The NSE values (0.45 for training and 0.39 for testing) fall into the satisfactory range but are the lowest among the good models. The PBIAS values (-0.00 for training and -0.71 for testing) indicate minimal bias. While the M8 performs adequately, it is the least precise among the models with good ratings.

To sum up, M4 and M5 stand out as the most promising. M6 and M7 also perform well but show marginally higher errors and biases. M3, while satisfactory, lags in predictive power, particularly in testing. M8 is the least precise among the good models. Each model's input features contribute differently to its performance, indicating that a careful selection of inputs is crucial for optimizing model accuracy and reliability.

3.3.2. Results Analysis: AELM-Based NHI Estimation

Model M3 (Figure 5) exhibits outstanding performance during the training and testing. The R^2 values of 0.93 during training and testing place it within the very good category, indicating that it explains a substantial portion of the variance in NHI. Its NRMSE values (0.12 for training and 0.12 for testing) are low, suggesting high accuracy. The NSE values (0.87 for training and 0.87 for testing) further reinforce the model's robustness, categorizing it as very good. The PBIAS values (-0.00 for training and 1.18 for testing) indicate minimal bias. In summary, M3 demonstrates excellent predictive power and accuracy, making it a highly reliable model for estimating NHI.

Model M4 (Figure 5) also shows excellent performance, with slightly lower accuracy than M3. Its R^2 values (0.92 for training and 0.93 for testing) fall within the very good range, showing that it effectively explains the variance in NHI. The NRMSE values (0.13 for training and 0.12 for testing) are slightly higher than those of M3 but still indicate high precision. The NSE values (0.85 for training and 0.86 for testing) suggest good performance. The PBIAS values (-0.00 for training and 1.36 for testing) are minimal, indicating low bias. Overall, M4 is a strong model with high generalization capability, slightly less precise than M3, but still highly reliable.

Model M5 (Figure 5) performs similarly to M3, with slight improvements in some areas. Its R^2 values (0.93 for training and 0.93 for testing) are high, indicating a robust variance explanation. The NRMSE values (0.12 for training and 0.12 for testing) are comparable to those of M3, suggesting high accuracy. The NSE values (0.86 for training and 0.87 for testing) consistently perform well. The PBIAS values (0.02 for training and 1.35 for testing) are slightly higher, indicating a minor increase in bias during testing. Despite this, M5 remains a highly reliable model with strong generalization, marginally outperforming M4.

Model M6 (Figure 5) maintains high performance, closely following M5. The R^2 values (0.93 for training and 0.93 for testing) are within the very good range. Its NRMSE values (0.12 for training and 0.12 for testing) suggest high precision, similar to the other top-performing models. The NSE values (0.87 for training and 0.87 for testing) indicate consistently excellent performance. The PBIAS values (-0.00 for training and 1.47 for testing) are slightly higher, demonstrating minimal bias. Although somewhat less precise than M5, M6 performs reliably with minimal bias and strong generalization.

Model M7 (Figure 5) shows a slight decline in performance compared to M5 and M6. The R^2 values (0.93 for training and 0.93 for testing) remain in the very good range. Its NRMSE values (0.13 for training and 0.12 for testing) indicate high precision but slightly higher errors than M6. The NSE values (0.86 for training and 0.87 for testing) suggest excellent performance during testing. The PBIAS values (0.00 for training and 1.46 for testing) are minimal, indicating low bias. Despite this, M7 demonstrates excellent performance but with slightly reduced accuracy and increased bias compared to M5 and M6.

Model M8 (Figure 5) shows a notable performance that is comparable to that of the best models. The R^2 values (0.93 for training and 0.94 for testing) are high, indicating excellent variance explanation. The NRMSE values (0.13 for training and 0.12 for testing) are slightly higher than those of M5, suggesting high accuracy. The NSE values (0.86 for training and 0.87 for testing) fall into the very good range, indicating strong performance. The PBIAS values (0.00 for training and 1.30 for testing) are minimal, indicating low bias. While M8 performs excellently, it is on par with the best models but with a slightly higher bias.

To wrap up, Models M3, M5, and M8 stand out as the most reliable and precise, with M5 having a slight edge over M4 and M6. M7 also performs well but shows marginally higher errors and biases. Each model demonstrates excellent predictive power, indicating that the selection of inputs significantly contributes to optimizing model accuracy and reliability.

3.3.3. Results Analysis: AELM-Based ANUE Estimation

Model M3 (Figure 5) shows a mixed performance during the training and testing. The R^2 values of 0.50 during training and 0.31 during testing indicate that it explains around half of the variance in ANUE during training but much less during testing. Its NRMSE values (0.79 for training and 0.95 for testing) suggest that the model errors are relatively high, especially during testing, reflecting unsatisfactory accuracy. The NSE values (0.25 for training and -0.01 for testing) indicate unsatisfactory performance, particularly during testing, where the negative NSE value suggests that the model's predictions are worse than the mean of the observed data. The PBIAS values (0.00 for training and 0.34 for testing) show minimal bias during training but are slightly underestimated during testing. M3 demonstrates limited predictive capability, especially in testing, indicating potential model complexity or input sufficiency issues.

Model M4 (Figure 5) exhibits slightly better performance compared to M3. Its R^2 values (0.50 for training and 0.36 for testing) show a marginal improvement in the explained variance during testing. The NRMSE values (0.79 for training and 0.91 for testing) indicate a slight reduction in error compared to M3. The NSE values (0.25 for training and 0.07 for testing) suggest some improvement, moving into a slightly positive range during testing, which is better than M3's negative NSE. The PBIAS values (0.00 for training and 0.07 for testing) indicate minimal bias, showing a slight underestimation during testing but better control over bias than M3. Overall, M4 slightly improves predictive accuracy and bias control compared to M3.

Model M5 (Figure 5) shows noticeable improvement over M3 and M4. Its R^2 values (0.57 for training and 0.43 for testing) indicate better variance explanation during both phases. The NRMSE values (0.75 for training and 0.88 for testing) are lower than those of M3 and M4, suggesting better accuracy. The NSE values (0.33 for training and 0.12 for testing) show improvement, particularly during testing. However, PBIAS values (-0.01 for training and -0.14 for testing) indicate a slight overestimation bias during both phases. Despite the bias, M5 demonstrates better overall performance, making it more reliable than M3 and M4.

Model M6 (Figure 5) maintains high performance, like M5, but with a slight drop in some areas. The R^2 values (0.56 for training and 0.36 for testing) are slightly lower than M5, indicating less variance explanation during testing. The NRMSE values (0.76 for training and 0.90 for testing) suggest a slight increase in error. The NSE values (0.32 for training and 0.09 for testing) indicate a slight drop in predictive accuracy during testing. The PBIAS values (-0.02 for training and -0.09 for testing) show a minimal overestimation bias. Although slightly less precise than M5, M6 still provides promising performance with good generalization.

Model M7 (Figure 5) shows mixed performance compared to M5 and M6. The R^2 values (0.54 for training and 0.43 for testing) are comparable to M5 but slightly lower than M6 during training. Its NRMSE values (0.77 for training and 0.87 for testing) indicate slightly higher errors during training but comparable performance during testing. The NSE values (0.29 for training and 0.14 for testing) are somewhat lower during training but higher during testing than M6. The PBIAS values (-0.01 for training and -0.14 for testing) indicate a slight overestimation bias. Despite this, M7 performs decently, with accuracy and bias control comparable to M5.

Model M8 (Figure 5) shows performance like M5 and M7. The R^2 values (0.50 for training and 0.39 for testing) are slightly lower than M7, indicating less variance explanation. The NRMSE values (0.79 for training and 0.88 for testing) suggest higher errors than M5 and M7. The NSE values (0.25 for training and 0.13 for testing) show similar predictive accuracy to M5. The PBIAS values (0.00 for training and 0.13 for testing) indicate minimal bias. While M8 performs adequately, it is less precise than M5 and M7.

In conclusion, Models M5 and M7 stand out as the most reliable and accurate, with M5 having a slight edge. M6 and M8 also perform well but show marginally higher errors and biases. M3, while satisfactory, shows limited predictive power, particularly in testing. Each model input feature contributes differently to its performance, indicating that a careful selection of inputs is crucial for optimizing model accuracy and reliability.

3.3.4. Relative Error Analysis in Yield Estimation

Figure 6 shows the relative error across all models in estimating yield. All models demonstrate their strongest performance in the 0–5% relative error range. The analysis indicates that Model M8 generally has a higher percentage of samples in the lower error ranges (0–5%), indicating good accuracy for a significant portion of predictions. However, it also has higher percentages in the higher error ranges (>40%), suggesting some variability in performance. Models M4 and M5 also demonstrate strong performance in lower error ranges, with relatively fewer samples in higher error ranges, indicating more consistent accuracy.

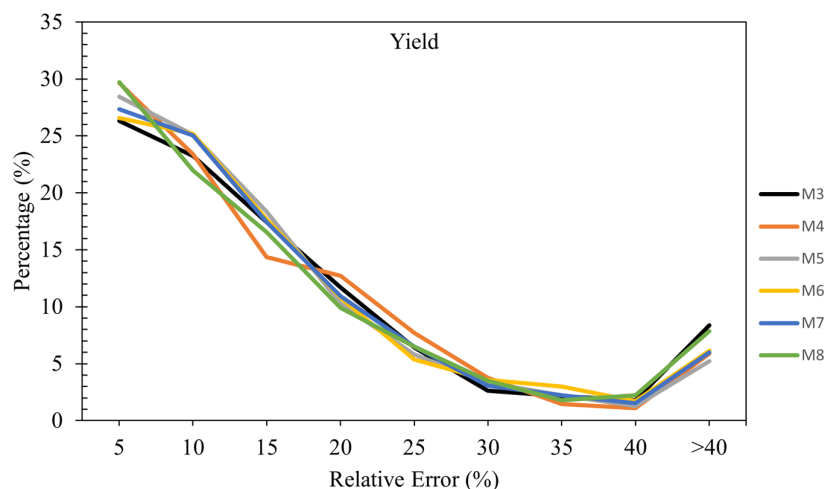


Figure 6. Relative error across all models in estimating yield.

3.3.5. Relative Error Analysis in NHI Estimation

Figure 7 shows the relative error across all models in estimating NHI. In the 0–5% relative error range, Models M3, M4, M5, and M8 have similar performance, each with around 33% of their samples falling into this category. The analysis shows that models M3, M4, M5, and M8 generally have a high percentage of samples in the lower error ranges (0–5%), indicating good accuracy for a significant portion of predictions. However, model M8 shows some variability with higher percentages in the higher error ranges (>40%). Models M4 and M5 also demonstrate strong performance in the lower error ranges, with relatively fewer samples in higher error ranges, indicating more consistent accuracy. Models M6 and M7 exhibit balanced performance across different error ranges, with fewer extreme inaccuracies.

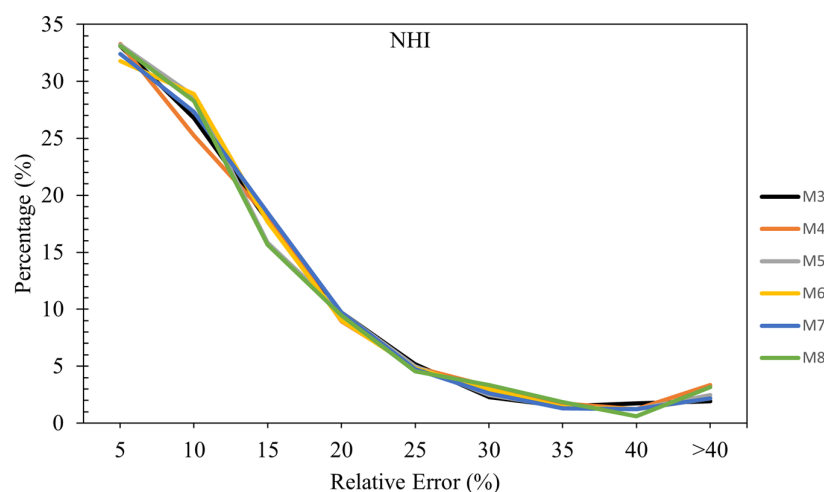


Figure 7. Relative error across all models in estimating NHI.

3.3.6. Relative Error Analysis in ANUE Estimation

Figure 8 compares six models (M3 to M8) based on the percentage of samples falling within various relative error ranges for ANUE. All six models (M3 to M8) exhibit a high percentage of samples with relative errors greater than 40%, ranging from 52.57% to 55.14%. This suggests that these models struggle with accuracy for a significant portion of their predictions. The consistently high error rates across all models indicate a need for further investigation and refinement to improve their predictive performance and reduce the occurrence of significant errors.

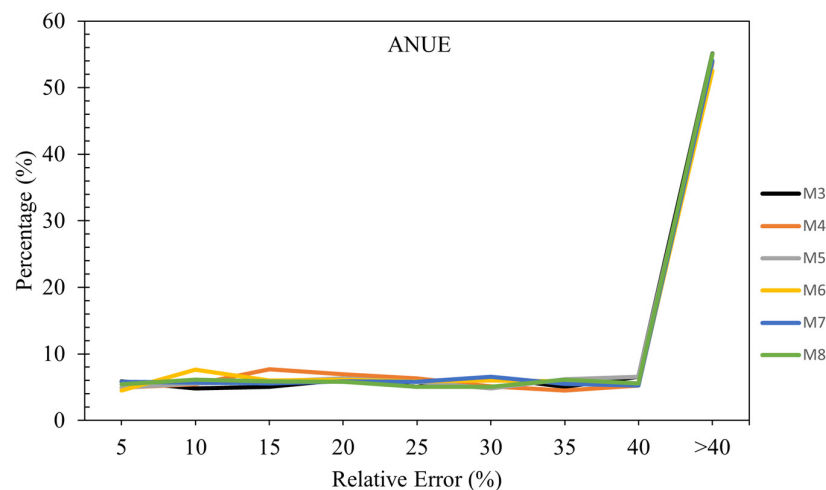


Figure 8. Relative error across all models in estimating ANUE.

4. Advantages, Limitations, and Future Improvements

4.1. Advantages

The methods applied in this study, particularly the AELM and the PSO, offer several notable advantages. Firstly, the AELM method is highly efficient, significantly reducing the training time compared to traditional gradient-based learning algorithms. This efficiency is achieved by randomly initializing the input weights and biases and then solving the output weights analytically, bypassing the need for iterative tuning. This aspect makes AELM particularly suitable for real-time applications where quick training is essential. Moreover, incorporating an iterative approach, the AELM method overcomes the limitation of random initialization of input weights and biases. This reduces the variability in performance and ensures more consistent and robust models, fully leveraging the data's potential.

Additionally, the AELM demonstrates excellent generalization capabilities, which helps maintain high training and unseen test data accuracy. This reduces the risk of overfitting, ensuring that the models are robust and reliable across various datasets. Thanks to its swarm intelligence-based approach, the PSO algorithm further complements this by effectively navigating large solution spaces and avoiding local minima. This results in a more global optimum solution, enhancing the model's overall performance.

Integrating AELM with PSO for feature selection (AELMPSO) provides another layer of optimization, ensuring that the most relevant features are selected, simplifying the models, and improving their interpretability and performance. This method helps reduce the curse of dimensionality and minimizes computational requirements, making it a practical approach for handling large datasets.

Overall, the best models incorporated soil parameters such as nitrogen application rate (N_Rate [In1]), soil nitrate levels at harvest (Harvest_NO3 [In13]), and enzyme activities (e.g., Init_NAG [In20]) alongside climate variables such as cumulative heat units (CHU_July [In29], CHU_August [In30]) and rainfall during key periods (Rain_June [In34], Rain_October [In38]), to predict yield, nitrogen harvest index (NHI), and agronomic nitrogen use efficiency (ANUE) effectively. Furthermore, our results indicate that soil health indicators, such as enzyme activities at side dress (β -glucosidase [BG-In18] and N-acetyl- β -D-glucosaminidase [NAG-In20]) and soil proteins, hold significant potential for enhancing nitrogen response predictions. Currently excluded from nitrogen recommendations models, these soil health indicators could provide valuable insights to improve yield outcomes and environmental sustainability in nitrogen management.

4.2. Limitations

Despite these advantages, the proposed methods do have certain limitations. One significant limitation is the AELM's performance in predicting ANUE. The analysis showed that the models struggled with accuracy for ANUE, with high relative errors and low NSE values, indicating unsatisfactory performance. Additionally, while the AELM performed reasonably well in predicting yield, there is still room for improvement to enhance its predictive accuracy and reliability.

Another limitation is the computational complexity associated with the PSO algorithm in the introduced AELMPSO feature selection method, especially with large populations and high-dimensional search spaces. The need for tuning various parameters, such as inertia weight, population size, and learning coefficients, adds to the complexity and requires careful balancing to ensure optimal performance.

4.3. Future Improvements

Several future improvements can be considered to address these limitations and further enhance the performance and reliability of the applied methods. One potential improvement is incorporating more sophisticated initialization techniques for the AELM's input weights and biases. Methods such as Autoencoder as a deep learning technique and metaheuristic algorithms initialization can help reduce the dependency on random initialization, leading to more consistent and robust model performance.

Additionally, hybrid optimization approaches that combine PSO with other optimization techniques could be explored. For instance, integrating PSO with Genetic Algorithms (GA) [63] could provide a more comprehensive search strategy, balancing exploration and exploitation more effectively and potentially improving convergence rates.

Last, expanding the scope of feature selection methods by incorporating advanced techniques such as mutual information-based selection, regularization methods, or embedded methods within the AELMPSO framework could further refine the feature selection process. This would improve model accuracy and interpretability and reduce computational overhead.

By addressing these areas, the robustness, efficiency, and applicability of the AELM and PSO methods can be significantly enhanced, paving the way for their broader adoption in various complex machine-learning tasks.

5. Conclusions

This study investigated feature selection using the adaptive extreme learning machine particle swarm optimization (AELMPSO) and modeled three corn nitrogen response metrics (yield, NHI, and ANUE) using AELM. Six models with three to eight features (M3 to M8) were chosen based on AELMPSO for all outputs, and their accuracy and reliability were evaluated.

Feature selection through AELMPSO identified different sets of inputs for each output. For yield, the inputs varied across models, with Model M8 utilizing the maximum features and Model M3 the minimum. A similar pattern was observed for NHI, while all models consistently selected the first few inputs for ANUE.

In AELM-based modeling for yield, Models M4 and M5 demonstrated the highest reliability and precision, consistently performing well during both training and testing phases. Models M6 and M7 also performed well but showed slightly higher errors and biases. Model M3 had limited predictive power, indicating potential overfitting or insufficient input features. Model M8, while adequate, was the least precise among the good models.

Models M3, M5, and M8 stood out as the most reliable and precise for NHI, with M5 having a slight edge. These models explained a substantial portion of the variance in

NHI and demonstrated high accuracy. Models M4 and M6 maintained high performance, closely following the top models, while Model M7 exhibited excellent performance with low errors and biases.

Models M5 and M7 were the most reliable and accurate in modeling ANUE, showing better variance explanation and lower errors than others. Models M6 and M8 also performed well but had marginally higher errors and biases. Models M3 and M4 demonstrated limited predictive capability, especially in testing, indicating potential issues with model complexity or input sufficiency.

This study highlights the importance of careful feature selection and model evaluation in achieving accurate and reliable predictions. Models M4 and M5 consistently showed strong performance across different outputs, making them the most dependable choices. However, the high error rates observed in ANUE across all models suggest that further investigation and refinement are needed to enhance predictive capabilities. Future research could optimize the feature selection process and explore advanced modeling techniques to improve accuracy and reliability.

Our findings highlight that the innovative approach of integrating soil health indicators with ML algorithms, such as AELM and PSO, facilitates accurate nitrogen efficiency predictions even with large and complex datasets, thereby validating the hypothesis of enhanced NUE modeling through advanced ML approaches. These findings are consistent with previous studies showing that innovative machine learning techniques can significantly improve agricultural nutrient management and fertilization practices [10,11]. This study underscores the potential of soil health indicators, including enzyme activities (BG and NAG) and soil proteins, as valuable tools for predicting crop responses to nitrogen. These findings suggest that integrating such biological indicators into nitrogen models could open new pathways to improve the sustainability and accuracy of fertilization practices in corn.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/agronomy15010244/s1>, Table S1: Descriptive statistics of the inputs and outputs; Figure S1: Heatmap of Correlation Between Inputs and Outputs.

Author Contributions: Conceptualization, G.D. and J.D.-R.; methodology, J.B.; software, I.E.; validation, J.D.-R., H.B., G.D. and A.N.R.; formal analysis, J.B. and I.E.; investigation, J.B.; resources, G.D. and J.D.-R.; data curation, J.B.; writing—original draft preparation, J.B. and I.E.; writing—review and editing, J.B., I.E., G.D., A.N.R., H.B. and J.D.-R.; supervision, J.D.-R.; project administration, J.D.-R. and G.D.; funding acquisition, J.D.-R. and G.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministère de l’Agriculture, des Pêcheries et de l’Alimentation du Québec (MAPAQ) through the InnovAction Program, project number IA121702.

Data Availability Statement: The dataset is available upon request from the authors.

Acknowledgments: The authors thank the administrative and laboratory teams at Université Laval for their invaluable support throughout this study. We also extend our thanks to the technical field teams at PleineTerre Inc. for their assistance in data collection and site management, which significantly contributed to the success of this research.

Conflicts of Interest: Author Gabriel Deslauriers was employed by the company PleineTerre Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. CRAAQ. *Guide de Référence en Fertilisation*, 2nd ed.; Centre de Référence en Agriculture et Agroalimentaire du Québec: Québec, QC, Canada, 2010.
2. Hergoualc'h, K.; Akiyama, H.; Bernoux, M.; Chirinda, N.; Prado, A.d.; Kasimir, Å.; MacDonald, J.D.; Ogle, S.M.; Regina, K.; Weerden, T.J.v.d. N₂O emissions from managed soils, and CO₂ emissions from lime and urea application. In *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*; IPCC: Geneva, Switzerland, 2019.
3. Shcherbak, I.; Millar, N.; Robertson, G.P. Global metaanalysis of the nonlinear response of soil nitrous oxide (N₂O) emissions to fertilizer nitrogen. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9199–9204. [[CrossRef](#)] [[PubMed](#)]
4. Philibert, A.; Loyce, C.; Makowski, D. Quantifying Uncertainties in N₂O Emission Due to N Fertilizer Application in Cultivated Areas. *PLoS ONE* **2012**, *7*, e50950. [[CrossRef](#)] [[PubMed](#)]
5. Ju, X.-T.; Xing, G.-X.; Chen, X.-P.; Zhang, S.-L.; Zhang, L.-J.; Liu, X.-J.; Cui, Z.-L.; Yin, B.; Christie, P.; Zhu, Z.-L.; et al. Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 3041–3046. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, X.; Wang, Q.; Xu, J.; Gilliam, F.S.; Tremblay, N.; Li, C. In Situ Nitrogen Mineralization, Nitrification, and Ammonia Volatilization in Maize Field Fertilized with Urea in Huanghuaihai Region of Northern China. *PLoS ONE* **2015**, *10*, e0115649. [[CrossRef](#)]
7. Sen, S.; Setter, T.L.; Smith, M.E. Maize root morphology and nitrogen use efficiency—A Review. *Agric. Rev.* **2012**, *33*, 16–26.
8. Zhang, J.; He, P. Field-Specific Fertilizer Recommendations for Better Nitrogen Use in Maize. *Better Crops* **2018**, *102*, 12–14. [[CrossRef](#)]
9. Kafle, J.; Bhandari, L.; Neupane, S.; Aryal, S. A Review on Impact of Different Nitrogen Management Techniques on Maize (*Zea mays* L.) Crop Performance. *AgroEnviron. Sustain.* **2023**, *1*, 192–198. [[CrossRef](#)]
10. Ennaji, O.; Vergütz, L.; El Allali, A. Machine learning in nutrient management: A review. *Artif. Intell. Agric.* **2023**, *9*, 1–11. [[CrossRef](#)]
11. Qin, Z.; Myers, D.B.; Ransom, C.J.; Kitchen, N.R.; Liang, S.Z.; Camberato, J.J.; Carter, P.R.; Ferguson, R.B.; Fernandez, F.G.; Franzen, D.W. Application of machine learning methodologies for predicting corn economic optimal nitrogen rate. *Agron. J.* **2018**, *110*, 2596–2607. [[CrossRef](#)]
12. Cardoso, E.J.B.N.; Vasconcellos, R.L.F.; Bini, D.; Miyauchi, M.Y.H.; Santos, C.A.d.; Alves, P.R.L.; Paula, A.M.d.; Nakatani, A.S.; Pereira, J.d.M.; Nogueira, M.A. Soil health: Looking for suitable indicators. What should be considered to assess the effects of use and management on soil health? *Sci. Agric.* **2013**, *70*, 274–289. [[CrossRef](#)]
13. Griffiths, B.S.; Faber, J.; Bloem, J. Applying Soil Health Indicators to Encourage Sustainable Soil Use: The Transition from Scientific Study to Practical Application. *Sustainability* **2018**, *10*, 3021. [[CrossRef](#)]
14. Dahal, S.; Franklin, D.H.; Subedi, A.; Cabrera, M.L.; Ney, L.; Fatzinger, B.; Mahmud, K. Interrelationships of Chemical, Physical and Biological Soil Health Indicators in Beef-Pastures of Southern Piedmont, Georgia. *Sustainability* **2021**, *13*, 4844. [[CrossRef](#)]
15. Wade, J.; Culman, S.W.; Logan, J.A.R.; Poffenbarger, H.; Demyan, M.S.; Grove, J.H.; Mallarino, A.P.; McGrath, J.M.; Ruark, M.; West, J.R. Improved soil biological health increases corn grain yield in N fertilized systems across the Corn Belt. *Sci. Rep.* **2020**, *10*, 3917. [[CrossRef](#)] [[PubMed](#)]
16. Cardoso, E.G.; Sá, J.C.D.M.; Briedis, C.; Ferreira, A.D.O.; Borszowskei, P.R.; Santos, J.B.; Massinham, A.; Ferreira, C.F.; Siuta Júnior, D.; Baranek, E.J. Nitrogen dynamics in soil management systems: II—Mineralization and nitrification rates. *Rev. Bras. Ciência Solo* **2011**, *35*, 1651–1660. [[CrossRef](#)]
17. Laine, M.; Rütting, T.; Alakukku, L.; Palojarvi, A.; Strömmer, R. Process rates of nitrogen cycle in uppermost topsoil after harvesting in no-tilled and ploughed agricultural clay soil. *Nutr. Cycl. Agroecosystems* **2018**, *110*, 39–49. [[CrossRef](#)]
18. Adetunji, A.T.; Lewu, F.B.; Mulidzi, R.; Ncube, B. The biological activities of β -glucosidase, phosphatase and urease as soil quality indicators: A review. *J. Soil Sci. Plant Nutr.* **2017**, *17*, 794–807. [[CrossRef](#)]
19. Pandey, D.; Agrawal, M.; Bohra, J.S. Assessment of soil quality under different tillage practices during wheat cultivation: Soil enzymes and microbial biomass. *Chem. Ecol.* **2015**, *31*, 510–523. [[CrossRef](#)]
20. Stott, D.; Andrews, S.; Liebig, M.; Wienhold, B.J.; Karlen, D. Evaluation of β -glucosidase activity as a soil quality indicator for the soil management assessment framework. *Soil Sci. Soc. Am. J.* **2010**, *74*, 107–119. [[CrossRef](#)]
21. Congreves, K.A.; Otchere, O.; Ferland, D.; Farzadfar, S.; Williams, S.; Arcand, M.M. Nitrogen Use Efficiency Definitions of Today and Tomorrow. *Front. Plant Sci.* **2021**, *12*, 637108. [[CrossRef](#)]
22. Craswell, E.T.; Godwin, D.C. The efficiency of nitrogen fertilizers applied to cereals grown in different climates. *Adv. Plant Nutr.* **1984**, *1*, 1–55.
23. Fageria, N. Nitrogen harvest index and its association with crop yields. *J. Plant Nutr.* **2014**, *37*, 795–810. [[CrossRef](#)]
24. Good, A.G.; Shrawat, A.K.; Muench, D.G. Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? *Trends Plant Sci.* **2004**, *9*, 597–605. [[CrossRef](#)] [[PubMed](#)]
25. Moll, R.; Kamprath, E.; Jackson, W. Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *Agron. J.* **1982**, *74*, 562–564. [[CrossRef](#)]

26. Parent, L.E.; Deslauriers, G. Simulating Maize Response to Split-Nitrogen Fertilization Using Easy-to-Collect Local Features. *Nitrogen* **2023**, *4*, 331–349. [[CrossRef](#)]
27. Ransom, C.J.; Kitchen, N.R.; Camberato, J.J.; Carter, P.R.; Ferguson, R.B.; Fernández, F.G.; Franzen, D.W.; Laboski, C.A.M.; Myers, D.B.; Nafziger, E.D.; et al. Combining corn N recommendation tools for an improved economical optimal nitrogen rate estimation. *Soil Sci. Soc. Am. J.* **2023**, *87*, 902–917. [[CrossRef](#)]
28. Siddiqi, M.Y.; Glass, A.D. Utilization index: A modified approach to the estimation and comparison of nutrient utilization efficiency in plants. *J. Plant Nutr.* **1981**, *4*, 289–302. [[CrossRef](#)]
29. Steenbjerg, F.; Jakobsen, S.T. Plant nutrition and yield curves. *Soil Sci.* **1963**, *95*, 69–88. [[CrossRef](#)]
30. Xu, X.; He, P.; Pampolino, M.F.; Qiu, S.; Zhao, S.; Zhou, W. Spatial Variation of Yield Response and Fertilizer Requirements on Regional Scale for Irrigated Rice in China. *Sci. Rep.* **2019**, *9*, 3589. [[CrossRef](#)]
31. Mahal, N.K.; Castellano, M.J.; Miguez, F.E. Conservation Agriculture Practices Increase Potentially Mineralizable Nitrogen: A Meta-Analysis. *Soil Sci. Soc. Am. J.* **2018**, *82*, 1270–1278. [[CrossRef](#)]
32. Franzluebbers, A.J.; Shoemaker, R.; Cline, J.; Lipscomb, B.; Stafford, C.; Farmaha, B.S.; Waring, R.; Lowder, N.; Thomason, W.E.; Poore, M.H. Adjusting the N fertilizer factor based on soil health as indicated by soil-test biological activity. *Agric. Environ. Lett.* **2022**, *7*, e20091. [[CrossRef](#)]
33. Kelleher, J.D.; Mac Namee, B.; D’arcy, A. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*; MIT Press: Cambridge, MA, USA, 2020.
34. Theng, D.; Bhoyar, K.K. Feature selection techniques for machine learning: A survey of more than two decades of research. *Knowl. Inf. Syst.* **2024**, *66*, 1575–1637. [[CrossRef](#)]
35. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
36. Bonakdari, H.; Ebtehaj, I.; Ladouceur, J.D. *Machine Learning in Earth, Environmental and Planetary Sciences: Theoretical and Practical Applications*; Elsevier: Amsterdam, The Netherlands, 2023.
37. Huang, G.-B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2011**, *42*, 513–529. [[CrossRef](#)] [[PubMed](#)]
38. Safari, M.J.S.; Ebtehaj, I.; Bonakdari, H.; Es-haghi, M.S. Sediment transport modeling in rigid boundary open channels using generalize structure of group method of data handling. *J. Hydrol.* **2019**, *577*, 123951. [[CrossRef](#)]
39. Ebtehaj, I.; Bonakdari, H. A reliable hybrid outlier robust non-tuned rapid machine learning model for multi-step ahead flood forecasting in Quebec, Canada. *J. Hydrol.* **2022**, *614*, 128592. [[CrossRef](#)]
40. Ebtehaj, I.; Bonakdari, H.; Safari, M.J.S.; Gharabaghi, B.; Zaji, A.H.; Madavar, H.R.; Khozani, Z.S.; Es-haghi, M.S.; Shishegaran, A.; Mehr, A.D. Combination of sensitivity and uncertainty analyses for sediment transport modeling in sewer pipes. *Int. J. Sediment Res.* **2020**, *35*, 157–170. [[CrossRef](#)]
41. Ebtehaj, I.; Soltani, K.; Amiri, A.; Faramarzi, M.; Madramootoo, C.A.; Bonakdari, H. Prognostication of shortwave radiation using an improved No-Tuned fast machine learning. *Sustainability* **2021**, *13*, 8009. [[CrossRef](#)]
42. Razmi, M.; Saneie, M.; Basirat, S. Estimating discharge coefficient of side weirs in trapezoidal and rectangular flumes using outlier robust extreme learning machine. *Appl. Water Sci.* **2022**, *12*, 176. [[CrossRef](#)]
43. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN’95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948.
44. Mahdavi-Meymand, A.; Sulisz, W. Development of particle swarm clustered optimization method for applications in applied sciences. *Prog. Earth Planet. Sci.* **2023**, *10*, 17. [[CrossRef](#)]
45. Alaliyat, S.; Yndestad, H.; Sanfilippo, F. Optimisation Of Boids Swarm Model Based On Genetic Algorithm and Particle Swarm Optimisation Algorithm (Comparative Study). In Proceedings of the ECMS, Brescia, Italy, 27–30 May 2014; pp. 643–650.
46. Singh, A.K.; Kumar, A. An improved dynamic weighted particle swarm optimization (IDW-PSO) for continuous optimization problem. *Int. J. Syst. Assur. Eng. Manag.* **2023**, *14*, 404–418. [[CrossRef](#)]
47. Guler, E. CITE-PSO: Cross-ISP Traffic Engineering Enhanced by Particle Swarm Optimization in Blockchain Enabled SDONs. *IEEE Access* **2024**, *12*, 27611–27632. [[CrossRef](#)]
48. Bohre, A.K.; Agnihotri, G.; Dubey, M. Hybrid butterfly based particle swarm optimization for optimization problems. In Proceedings of the 2014 First International Conference on Networks & Soft Computing (ICNSC2014), Guntur, India, 19–20 August 2014; pp. 172–177.
49. Tang, H.; Sun, W.; Yu, H.; Lin, A.; Xue, M.; Song, Y. A novel hybrid algorithm based on PSO and FOA for target searching in unknown environments. *Appl. Intell.* **2019**, *49*, 2603–2622. [[CrossRef](#)]
50. Ebtehaj, I.; Bonakdari, H.; Es-haghi, M.S. Design of a hybrid ANFIS–PSO model to estimate sediment transport in open channels. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2019**, *43*, 851–857. [[CrossRef](#)]
51. Mirjalili, S.; Song Dong, J.; Lewis, A.; Sadiq, A.S. Particle swarm optimization: Theory, literature review, and application in airfoil design. *Nat.-Inspired Optim. Theor. Lit. Rev. Appl.* **2020**, *811*, 167–184.

52. Freitas, D.; Lopes, L.G.; Morgado-Dias, F. Particle swarm optimisation: A historical review up to the current developments. *Entropy* **2020**, *22*, 362. [[CrossRef](#)]
53. Hastie, T.; James, G.; Witten, D.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013.
54. Bermingham, M.L.; Pong-Wong, R.; Spiliopoulou, A.; Hayward, C.; Rudan, I.; Campbell, H.; Wright, A.F.; Wilson, J.F.; Agakov, F.; Navarro, P.; et al. Application of high-dimensional feature selection: Evaluation for genomic prediction in man. *Sci. Rep.* **2015**, *5*, 10312. [[CrossRef](#)]
55. Aqel, D.; Al-Zubi, S.; Mughaid, A.; Jararweh, Y. Extreme learning machine for plant diseases classification: A sustainable approach for smart agriculture. *Clust. Comput.* **2022**, *25*, 2007–2020. [[CrossRef](#)]
56. Karamvand, A.; Hosseini, S.A.; Sharafati, A. SMAP products for prediction of surface soil moisture by ELM network model and agricultural drought index. *Acta Geophys.* **2023**, *71*, 1845–1856. [[CrossRef](#)]
57. Amiri, A.; Soltani, K.; Ebtehaj, I.; Bonakdari, H. A novel machine learning tool for current and future flood susceptibility mapping by integrating remote sensing and geographic information systems. *J. Hydrol.* **2024**, *632*, 130936. [[CrossRef](#)]
58. Salimi, A.; Noori, A.; Ebtehaj, I.; Ghobrial, T.; Bonakdari, H. Advancing Spatial Drought Forecasts by Integrating an Improved Outlier Robust Extreme Learning Machine with Gridded Data: A Case Study of the Lower Mainland Basin, British Columbia, Canada. *Sustainability* **2024**, *16*, 3461. [[CrossRef](#)]
59. Akbarian, M.; Saghafian, B.; Golian, S. Monthly streamflow forecasting by machine learning methods using dynamic weather prediction model outputs over Iran. *J. Hydrol.* **2023**, *620*, 129480. [[CrossRef](#)]
60. Okhravi, S.; Alemi, M.; Afzalimehr, H.; Schügerl, R.; Velísková, Y. Flow resistance at lowland and mountainous rivers. *J. Hydrol. Hydromech.* **2023**, *71*, 464–474. [[CrossRef](#)]
61. Paca, V.H.d.M.; de Souza, E.B.; Queiroz, J.C.B.; Espinoza-Dávalos, G.E. Assessment of precipitation and evapotranspiration in an urban area using remote sensing products (CHIRP, CMORPH, and SSEBop): The case of the metropolitan region of Belem, Amazon. *Water* **2023**, *15*, 3498. [[CrossRef](#)]
62. Grégoire, G.; Fortin, J.; Ebtehaj, I.; Bonakdari, H. Forecasting pesticide use on golf courses by integration of deep learning and decision tree techniques. *Agriculture* **2023**, *13*, 1163. [[CrossRef](#)]
63. Deng, Y.; Zhang, D.; Zhang, D.; Wu, J.; Liu, Y. A hybrid ensemble machine learning model for discharge coefficient prediction of side orifices with different shapes. *Flow Meas. Instrum.* **2023**, *91*, 102372. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.