agriculture

MDPI

Article

Rice (*Oryza sativa* L.) Growth Modeling Based on Growth Degree Day (GDD) and Artificial Intelligence Algorithms

Li-Wei Liu [1,2], Chun-Tang Lu [3], Yu-Min Wang [4], Kuan-Hui Lin [5], Xing-Mao Ma [2,*] and Wen-Shin Lin [5,*]

1 Department of Civil Engineering, National Pingtung University of Science and Technology, Pingtung County 91201, Taiwan; p10333003@g4e.npust.edu.tw
2 Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX 77840, USA
3 Crop Science Division, Taiwan Agricultural Research Institute, Council of Agriculture, Executive Yuan, Taichung City 413008, Taiwan; tang@tari.gov.tw
4 General Research Service Center, National Pingtung University of Science and Technology, Pingtung County 91201, Taiwan; wangym@mail.npust.edu.tw
5 Department of Plant Industry, National Pingtung University of Science and Technology, Pingtung County 91201, Taiwan; m10611010@mail.npust.edu.tw
* Correspondence: xma@civil.tamu.edu (X.-M.M.); wslin@mail.npust.edu.tw (W.-S.L.)

Abstract: Rice (*Oryza sativa* L.) growth prediction is key for precise rice production. However, the traditional linear rice growth forecasting model is ineffective under rapidly changing climate conditions. Here we show that growth rate (Gr) can be well-predicted by artificial intelligence (AI)-based artificial neural networks (ANN) and gene-expression programming (GEP), with accumulated air temperatures based on growth degree day (GDD). In total, 10,246 Gr from 95 cultivations were obtained with three cultivars, TK9, TNG71, and KH147, in Central and Southern Taiwan. The model performance was evaluated by the Pearson correlation coefficient (r), root mean square error (RMSE), and relative RMSE (r-RMSE) in the whole growth period (lifecycle), as well as the average and specific key stages (transplanting, 50% initial tillering, panicle initiation, 50% heading, and physiological maturity). The results in lifecycle Gr modeling showed that ANN and GEP models had comparable r (0.9893), but the GEP model had the lowest RMSE (3.83 days) and r-RMSE (7.24%). In stage average and specific key stages, each model has its own best-fit growth period. Overall, GEP model is recommended for rice growth prediction considering the model performance, applicability, and routine farming work. This study may lead to smart rice production due to the enhanced capacity to predict rice growth in the field.

Keywords: agricultural innovation; agricultural management precision agriculture; thermal time; rice growth prediction; artificial neural networks (ANN); gene-expression programming (GEP)

check for updates

Citation: Liu, L.-W.; Lu, C.-T.; Wang, Y.-M.; Lin, K.-H.; Ma, X.-M.; Lin, W.-S. Rice (*Oryza sativa* L.) Growth Modeling Based on Growth Degree Day (GDD) and Artificial Intelligence Algorithms. *Agriculture* **2022**, *12*, 59. https://doi.org/10.3390/agriculture12010059

Academic Editors: Nen-Fu Huang and Ho-Hsien Chen

Received: 15 November 2021
Accepted: 31 December 2021
Published: 3 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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

Rice is the second-highest produced cereal in terms of yield and is a staple food for approximately four billion people globally [1]; therefore, knowing the critical requirements for rice growth and the best timing for rice planting and harvest are crucial for understanding the effects of policy, and optimizing agricultural practice to achieve higher food security [2]. A precise method is needed so that the rice growth stages may be accurately predicted at varying environmental conditions [3], so as to effectively implement field cultivation management [4], achieve rationalization of irrigation and fertilization [5] and increase yield and profits for farmers [6]. Due to the development of digital agriculture modeling, the automatic field operation may be achieved by accurate growth period prediction, leading to precision agriculture [6]. Conversely, mismatched fertilization timing may cause lower fertilizer utilization efficiency [7,8], especially in traditional rice production under the critical threat of climate change. For example, the raised carbon dioxide concentration

causes existing rice growth patterns to change in Taiwan [9]. The increased extreme rainfall events cause variation in rice growth patterns [10] and yield loss [11].

Typical rice growth is modeled based on the accumulated thermal time, called growth degree days (GDD), to reach a specific growth stage regardless of the year or location [12,13]. The first rice GDD model, initially referred to as the Degree Day 50, was established in the 1970s and has been widely used in the Southern USA [14]. Many modern rice growth prediction models have been developed based on the GDD or similar linear-based principles, such as ORYZA [15] and DSSAT [16]. The accuracy of GDD model has been found to be much higher in fieldwork timing than the traditional calendar day method [17,18]. However, several difficulties hinder the application of the GDD model. Firstly, the basic growth temperature (Tb) must be calibrated by a long-term experiment before application [19–21]. Secondly, the GDD model is limited due to the nonlinear relationship between the rice growth rate, rice's cultivars and the growth conditions such as the temperature, humidity, and soil composition. The traditional GDD correlates developmental rate linearly to temperatures in some growth periods; however, linearization is often criticized for its oversimplification despite being widely used [22].

Due to the limitations of the above-mentioned linear-based GDD model, artificial intelligence (AI) based nonlinear algorithms have been developed and applied in rice research in recent years. Examples include yield prediction by artificial neural networks (ANN) [23] and gene-expression programming (GEP) [24], and short-term rice blast forecasting by ANN [25]. These AI-based studies showcase the ability of AI algorithms for complex problem-solving. However, although there are many separate studies on developing and using traditional GDD and AI algorithms in rice studies, no research has aimed to integrate GDD in AI-based modeling to predict rice growth. Due to the importance of accurate rice growth prediction for smart agriculture and increasing demand on rice production [26,27], this study aimed to apply AI algorithms to develop nonlinear rice growth models by GDD. The model performances were assessed by Pearson correlation coefficient (r), root mean squared error (RMSE), and relative RMSE (r-RMSE) in whole rice growth period (lifecycle), stage average, and specific key stages, to evaluate the applicability of AI-enabled model by comparing with traditional GDD model in different rice growth stages.

2. Materials and Methods

Three rice (*O. Sativa* subsp. *Japonica*) cultivars: Taiken 9 (TK9), Tainung 71 (TNG71) and Kaohsiung147 (KH147) were used in this study and were grown in 2006–2008, 2006–2009, and 2019 in two crop seasons, respectively. The growing days for TK9, TNK71, and KH147 in the first and second crop season were approximately 123 and 114 days (mid to late maturing), 118 and 104 days (early to mid-maturing), and 128 and 115 days (mid to late maturing), respectively. The cultivars TK9 and TNG71 had the same transplanting dates from 2006–2008. In 2009, TNG71 was planted in two different lots in the same crop season with a 14-day difference in the transplanting dates. Similar cultivation was performed for KH147 in 2019, with the plant grown in four separate lots with the transplanting dates separated by 14 days in the same crop season. In total, 95 cultivations were conducted by completely randomized design.

In each cultivation, growth data were obtained by observing the key growth stages of rice classified by Counce, et al. [28], including the transplanting stage (Stage 0, V3), 50% initial tillering stage (Stage 1, V5, the date when the tillers number exceeds 50%), panicle initiation stage (Stage 2, R0, the date when more than five random rice with panicle lengths observed in 2 mm), 50% heading stage (Stage 3, R3, the date when 50 plants' heading number reach 50%), and physiological maturity stage (Stage 4, R7, the date when most of the grains in the panicle are golden yellow, and 2 to 3 grains at the base of the panicle are still yellow-green). Based on key stages, the rice growth rate (Gr) was calculated by day^{-1}, and continuously accumulated from Gr = 0 to Gr = 4 (i.e., V3 to R7). For example, in 10 days in Stage 1, the Gr of the first day is 1/10 = 0.10, the second day is 2/10 = 0.20, etc., until the end of this stage (day 10, Gr = 1.00). Then, Gr for the next stages was calculated similarly

by dividing the specific date by the duration of the stage plus the Gr values at the end of the previous stage (e.g., 1.00 for Stage 1 and 2.00 for Stage 2).

2.1. Experiment Sites

The phenological data of rice was collected by the Taiwan Agriculture Research Institute (TARI) in Taichung (TK9 and TNG 71) and National Pingtung University of Science and Technology (NPUST) in Pingtung (KH147), respectively. The experimental sites of this study are marked in Figure 1. Two crop seasons were cultivated in both places following the typical continuous flood irrigation (Figure 2). The soil texture in TARI is alluvium of Holocene with fine-textured and well-drained red soil; and in NPUST, it is terrace deposits of Quaternary Period with non-calcareous well-drained shallow alluvial soil [29].

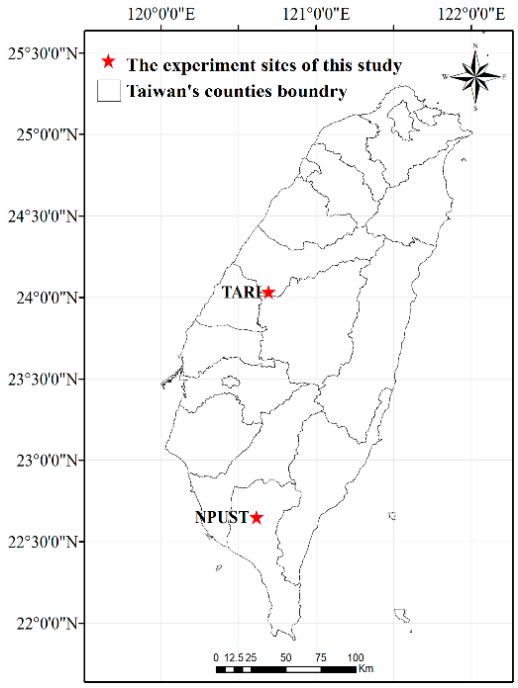

Figure 1. Experimental sites of this study.

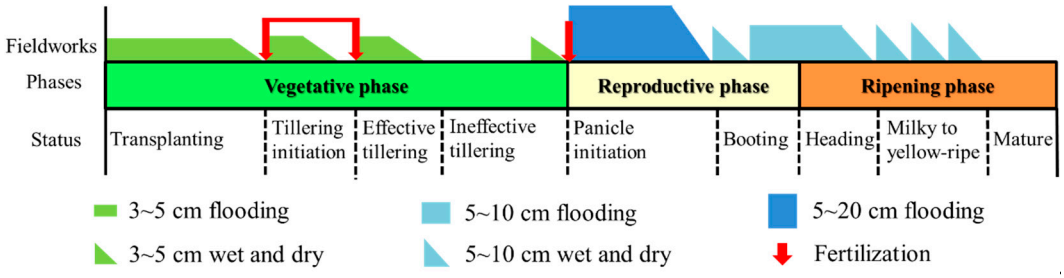

Figure 2. Common practices of rice cultivation in Taiwan.

2.2. Growth Degree Day

For temperature-based phenology models, the growth rate is generally modeled as a function of the effective temperature accumulation [30]. The prediction results can be used in field operations to guide agricultural managements, such as the timing for fertilization and irrigation. Several studies have been performed to improve the procedures for calculating the GDD [3]. In this study, a conventional GDD was calculated. It was assumed that a certain amount of effective temperature (°C d^{-1}) is needed to complete a given developmental stage [31]. GDD is generally computed based on a daily averaged

temperature minus base temperature (Tb). Traditionally, 10 °C is used as the Tb of rice cultivars. However, it had been found that the base temperatures of TK9, TNG71, and KH147 were 6.4, 6.6, and 7.6 °C, respectively [32]. A negative GDD indicates that the crop growth has stopped [33]. Finally, the daily GDD is calculated and summed, which constitutes the cumulative heating time required for the rice to reach the corresponding Gr [34]. The GDD in this study was calculated using Equation (1) below.

$$GDD = \frac{T_{min} + T_{max}}{2} - Tb \quad (1)$$

where, T_{min} is daily minimum temperature (°C), T_{max} is daily maximum temperature (°C), and Tb is base temperature (°C).

2.3. Model Development

2.3.1. Artificial Neural Network (ANN)

An ANN is a network of processing elements (PEs) assembled in layers and connected through several links or weights. ANN calculates its output from the dynamic input at epochs and compares it with the expected output from each input vector to compute the error [35]. The classic backpropagation neural network (BPNN) was used for Gr modeling in this study. NeuroSolution 7.1 from NeuroDimension, Inc. (Gainesville, FL, USA) was used for model establishment in this study. The model framework is shown in Figure 3. The total data, including input and output, was pre-processed by data min-max normalization from 0 to 1, and the data queue were re-sorted by random order to avoid errors caused by outliners and systemic cumulative errors. After the data pre-processing, the dataset was divided into three datasets: training, cross-validation (CV), and testing at a ratio of 70%, 20% and 10% following the typical practice of AI-based modeling. In general, the training dataset was used for initial modeling. However, the initially developed model may be overfitted by model optimization, i.e., model's hyper-parameters adjustment. The potential hyper-parameters, such as the number of iterations during the training, learning rate, layer parameters, or the number of layers, were adjusted by the CV dataset to increase the modeling accuracy. Finally, the testing dataset was applied to the adjusted model the for model performance evaluation. The initial setting of the model development included a 1000 maximum epoch with a 0.01 learning rate. These parameters were selected based on a previous study [25] because the goal of this study was to establish rice growth prediction models instead of comparing the model's performance by varying the model's parameters. The Levenberg–Marquardt gradient search method was applied with an early stopping callback to prevent overfitting. A single hidden layer was conducted with 10 PEs.

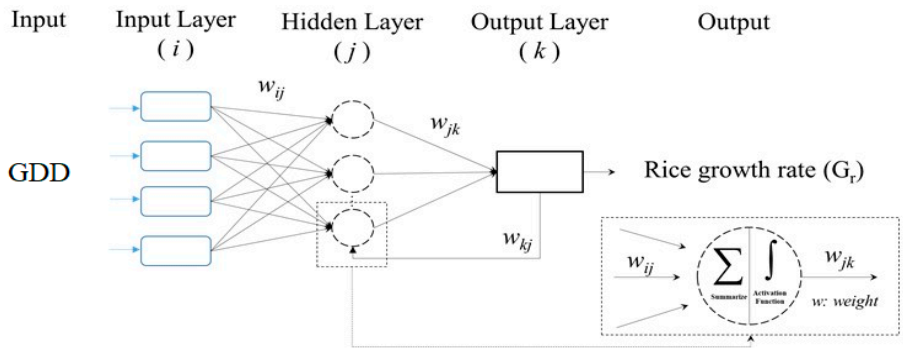

Figure 3. The working architecture of the backpropagation neural network (BPNN) model.

2.3.2. Gene-Expression Programming (GEP)

A classic GEP begins with a major contest and undergoes a continuous evolutionary process, such as selection, replication, mating, mutation, adaptation, reversal, and transformation to evolve toward a predetermined objective [36]. It overcomes genetic algorithm's (GA) chief shortcomings, such as the premature convergence and combined

explosion, and its evolution is significantly faster than GA and genetic programming (GP) [37]. The modeling flowchart and the structure are shown in Figure 4. The GEP model was trained, cross-validated and tested by the same dataset as the ANN model by GeneXproTools 5.0 program (Gepsoft Ltd., Bristol, UK), following the same ratio of data splitting (70%/20%/10%). The model adopted the same model's calculation elements, fitness function and parameter setting [38], namely five genes with fifty chromosomes and evolved to ten thousand generations. The genetic operators or set of functions were determined. Calculation elements of +, −, *, /, ln, ceiling, floor, absolute, tangent, exponential, 10^x, x^2, $x^{1/3}$, min, and avg. Afterward, a tree structure was established and linked by each between sub-expression trees (ET_{sub}). More details on the GEP theory can be found in [36,37].

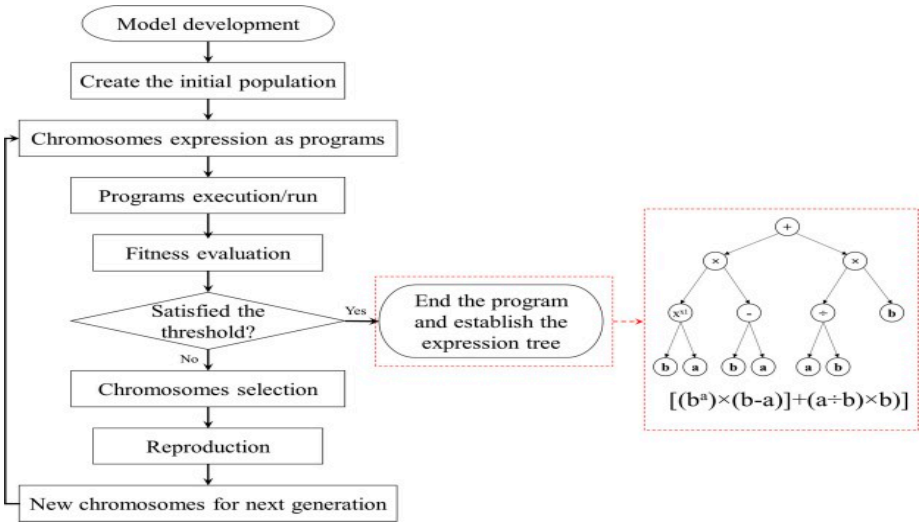

Figure 4. The modeling flowchart of GEP model [38].

2.3.3. Simple Regression Model (REG)

The simple regression model is a conventional GDD modeling algorithm based on a linear relationship. The dataset used for REG model was the same as that for ANN and GEP models. However, the REG algorithm did not allow the hyper-parameter adjustment by CV process; therefore, the 20% CV data were added to the training dataset, which means 90% data were used for training and 10% were used for testing in REG. The model can be described as Equation (2).

$$Gr = \alpha x_i + \beta + \varepsilon \quad (2)$$

where, x_i is the ith GDD; α is the linear equation slope; β is an intercept of the regression, and ε is the error term.

2.4. Model Assessment

The performances of ANN, GEP, and REG models were assessed by r, RMSE and relative root mean squared error (r-RMSE, Equation (3)). The significance of the Pearson correlation between observed Gr and predicted Gr of each model was calculated. The whole growth period data, including model training, CV, and testing datasets, were used to retrieve the entire rice lifecycle prediction results by ANN, GEP, and REG models, and the errors were calculated in days (1/Gr) by 0.1 intervals to compute RMSE based on the average Gr increment of each stage, i.e., Stage 1 is 0.0632, Stage 2 is 0.0334, Stage 3 is 0.0409, Stage 4 is 0.0269. A comparison by lifecycle, stage average, i.e., Stage 1~4, and specific key stage, i.e., Gr = 1 (V3), Gr = 2 (V5), Gr = 3 (R0), and Gr = 4 (R3), was conducted to evaluate the model performance.

$$r\text{-}RMSE(\%) = \frac{\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-1}}}{\overline{y}} \times 100 \quad (3)$$

where $\hat{y}_i$ is the predicted value, y_i is the observed value, $\overline{y}$ is the average of the observations, and n is the number of actual observations.

3. Results

3.1. Rice Growth Predicted by Growth Degree Day

The GDD-Gr relation was established for the whole dataset (Figure 5). The RMSE of Gr ranged from 0.1234 to 0.3199, as shown in Table 1. In Stage 1, the GDD model has an unignorable prediction bias result from the offset of Gr = 0, which means the initial modeling rice growth stage may start from Gr ≈ 0.45, possibly leading to a wrong fieldwork decision. Higher prediction errors also can be found in Stage 3 and Stage 4.

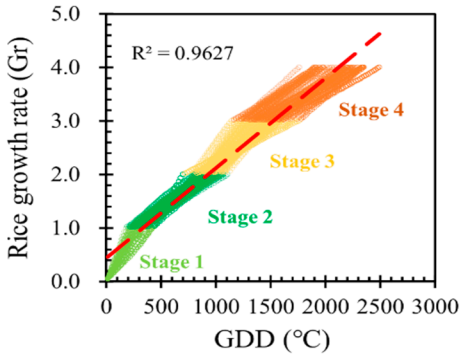

Figure 5. The relationship between accumulated GDD and Gr. Each data point represents the observed rice growth rate (Gr) data and its related cumulated GDD. The red dash-line is the least-squares regression line.

Table 1. GDD in each specific key stage.

Stages	GDD	RMSE	
	°C	Gr	1/Gr (Days)
Stage 1	308.4	0.1234	1.9531
Stage 2	915.7	0.1469	4.3970
Stage 3	1442.3	0.2791	6.8260
Stage 4	2192.5	0.3199	11.8907

3.2. Model Results

The r, RMSE, and r-RMSE of ANN, GEP, and REG model accuracy assessment results in model testing are shown in Table 2. ANN has the highest r, lowest RMSE and r-RMSE in the model testing stage. The established REG and GEP models presented in equations are shown in Equations (4)–(10).

$$Gr_{REG} = \left(1.6807 \times 10^{-3}\right) \times GDD + 4.3594 \times 10^{-1} \quad (4)$$

$$Gr_{GEP} = min[ET_{sub1}, ET_{sub2}, ET_{sub3}, ET_{sub4}, ET_{sub5}] \quad (5)$$

$$ET_{sub1} = \frac{\left(0.5 \times \left[\sqrt[3]{GDD}, -0.9385\right]\right)^2}{Ce(min[GDD,\ 7.373]) + 10^{\frac{-17.271}{GDD}}} \quad (6)$$

$$ET_{sub2} = \frac{Ce(GDD) \times (Fl(GDD))^2}{3.932 \times GDD^3} \quad (7)$$

$$ET_{sub3} = \left(0.5\langle 0.5\left\{0.5\left[1.186, \frac{143.1823}{GDD}\right], -8.1702\right\}, ln(GDD)\rangle\right)^2 \quad (8)$$

$$ET_{sub4} = abs(0.5[-8.949\sqrt[3]{0.5[GDD, -10.3011]}\ ,\ tanavg\ [-2.923, \frac{GDD}{6.2407}]]) \quad (9)$$

$$ET_{sub5} = 0.5[4.0326(Ce(10.3537) + Fl(GDD)), \frac{0.3054}{GDD}] \quad (10)$$

where, Gr_{REG} is the Gr predicted by REG model, Gr_{GEP} is the Gr predicted by GEP model, GDD is the accumulated growth degree day (°C), $ET_{sub1\sim5}$ is the sub-expression tree from 1 to 5, *min* is the *minimum* function, *Ce* is *ceiling* function, *Fl* is *floor* function, *tan* is the *tangent* function, and *abs* is the *absolute* function.

Table 2. Model performances on model testing phase show the similarity between two AI-based ANN and GEP models.

Indices	ANN	GEP	REG
r	0.9901 ***	0.9896 ***	0.9829 ***
RMSE	0.1573	0.1591	0.2043
r-RMSE	7.21%	7.30%	9.37%

***: The symbol denotes significant correlation at $p < 0.0001$.

3.3. Model Performance Evaluation

The model testing was conducted in the entire growth period for each model. The correlation between observed and predicted Gr is shown in Figure 6. The retrieved prediction results can be divided into three periods: the whole lifecycle, stage average, and the key stage, which is shown in Table 3. The ANN and GEP has an equivalent r value (0.9893) in the whole lifecycle prediction, but the GEP model has a relative lower RMSE and r-RMSE, followed by ANN and REG models. Based on the performance of r, the best performing model at stage 1 to 4 is GEP, REG, ANN, and REG, respectively. Furthermore, the performance of the AI-based ANN and GEP models can be determined by the RMSE (days) ratio between AI and REG models (Table 4). The result shows that the performance of ANN in the lifecycle, stage average from 1 to 4, and key stage Gr = 2, 3, 4, is better than the REG model. Only on key stage Gr = 1, the ANN performed slightly worse than the REG model. For the GEP model, its performance in the whole lifecycle, stage average 1, 3, 4, and key stage Gr = 3, 4, was better than the REG model, but the prediction accuracy at stage average 2, and key stages Gr = 1, 2, was lower than the REG model.

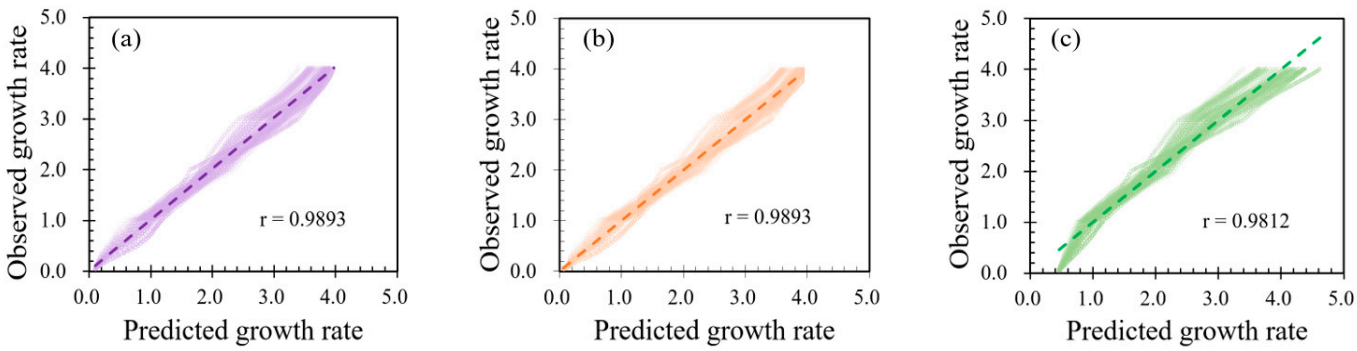

Figure 6. Modeling results for the whole growth period by three models: (**a**) ANN; (**b**) GEP; (**c**) REG. Each data point represents the observed rice growth rate (Gr) data and its related prediction of Gr. The dash-line is the least-squares regression line.

Table 3. Modeling results for the whole growth period by three models in the lifecycle, stage average, and specific key stages.

Period	Gr									1/Gr		
	r			RMSE			r-RMSE			RMSE (Days)		
	ANN	GEP	REG	ANN	GEP	REG	ANN	GEP	REG	ANN	GEP	REG
Lifecycle	**0.9893** ‡	**0.9893**	0.9812	0.1614	**0.1597**	0.2117	7.32%	**7.24%**	9.60%	3.8709	**3.8300**	5.0776
Stage 1 average	0.9119	**0.9131**	0.8969	0.1274	**0.1271**	0.2464	24.06%	**24.00%**	46.52%	2.0144	**2.0091**	3.8951
Stage 2 average	0.9430	0.9380	**0.9441**	**0.1116**	0.1165	0.1154	**7.36%**	7.69%	7.61%	**3.3434**	3.4914	3.4572
Stage 3 average	**0.8683**	0.8678	0.8624	**0.1747**	0.1777	0.1784	**6.93%**	7.05%	7.08%	**4.2685**	4.3426	4.3604
Stage 4 average	0.8045	0.8064	**0.8163**	0.2025	**0.1929**	0.2803	5.76%	**5.49%**	7.98%	7.5393	**7.1816**	10.4375
Gr = 1	-	-	-	0.1546	0.1685	**0.1237**	15.47%	16.85%	**12.37%**	2.4451	2.6644	**1.9561**
Gr = 2	-	-	-	0.1478	0.1618	**0.1475**	7.39%	8.09%	**7.38%**	4.4288	4.8465	**4.4199**
Gr = 3	-	-	-	0.2377	**0.2293**	0.2807	7.92%	**7.64%**	9.36%	5.8098	**5.6030**	6.8610
Gr = 4	-	-	-	0.2462	**0.2266**	0.3212	6.16%	**5.66%**	8.03%	9.1691	**8.3846**	11.9623

‡: All of the correlation coefficients (r) were significant at $p < 0.0001$. The boldface numbers represent the best performance of the model in each period comparison.

Table 4. RMSE comparison between AI-based ANN, GEP models to the REG model.

Model	Lifecycle	Stage Average				Specific Key Stage			
		1	2	3	4	Gr = 1	Gr = 2	Gr = 3	Gr = 4
ANN	23.77%	48.28%	**3.29%**	**2.11%**	27.77%	**−25.00%**	**1.20%**	15.32%	23.35%
GEP	**24.57%**	**48.42%**	−0.99%	0.41%	**31.19%**	−36.22%	−9.65%	**18.34%**	**29.91%**

The boldface numbers represent the best performance of the model in each period comparison.

4. Discussion

This study used rice growth data from three rice cultivars grown in tropical and subtropical climates and calculated the most common rice growth prediction parameter, GDD, by an adjunction weather station for each cultivation. The developed models were applied to retrieve the whole growth period dataset for model performance and applicability evaluation in the lifecycle, stage average, and specific key stages.

We compared the modeled results with a previous study [39] that predicted rice growth date by five conventional models. The models' input factors included critical day length, the minimum number of days required for heading, photoperiod sensitivity parameter, etc., and in-situ experimental parameters. The RMSEs of those models ranged from 6.09 to 8.60 days in heading and 4.23 to 9.12 days in maturity [39]. Those results are similar to the REG model in this study, i.e., 6.86 days in heading and 11.96 days. In comparison, the AI-based ANN and GEP models, showed slightly better predicting abilities at heading stage (Gr = 3). However, in the maturity stage (Gr = 4), ANN and GEP models do not present a significant improvement in modeling ability. It should be noticed that the modeling result of rice growth is influenced by many factors, such as the environment, cultivar, tillage method, etc. In general, the comparison results indicated that the AI-based ANN and GEP rice growth models have higher accuracy than the REG model.

4.1. Lifecycle vs. Stage Average vs. Specific Key Stage

These three different scenarios can be used for different purposes. For example, the lifecycle scenario was developed by a continuous dataset, which has integral prior probability and does not affect the model transition situation. Therefore, the lifecycle scenario is more suitable for smart agriculture, which operates routine fieldwork entirely based on modeling results by considering the prior probability, and guides farmers lacking in rice farming experience. The stage average scenario is more suitable for semi-automatic rice production, i.e., modern paddy rice farming. In the stage average scenario, farmers will need to modify the model based on their experience and have some background knowledge to select a suitable algorithm. Farmers can also decide the timing to start Stage 2 and use ANN model according to the rice seedling development. In specific key stage prediction, the physiological transformation and growth need to be adjusted at the

stage. Specific key stage scenario may be more beneficial for researchers. For farmers, the management measures may not be implemented in those specific key stages because agricultural management may require "early" or "late" applications to supply the necessary nutrition and water for the rice. Therefore, the specific key stage scenario results are more relevant to research. For farmers, the lifecycle or stage average scenario are more useful to conduct in their fieldwork.

4.2. Model Applicability

Irrigation is one of the most critical factors in rice production, even more so than fertilization. Therefore, the essential stages of rice production are the ones with a variant water control strategy, i.e., panicle initiation (Gr ≈ 1.9 to 2.1), booting to late heading (Gr ≈ 2.7 to 3.3), and around 10 days before harvest (Gr ≈ 3.8). Moreover, the dry land method can reduce non-effective tillering in the late-Vegetative phase since Gr starts around 1.5. In fertilization, the critical timing is panicle initiation, tillering initiation to 10~15 days after tillering initiation (Gr ≈ 1.0 to 1.3), leading to significant yield promotion when an appropriate application is made. Based on the comparison of important fieldwork conditions, the model performance indicated by RMSE (days) of Gr in 1.9 to 2.1 is REG > ANN > GEP. REG seems to have a better modeling ability. RMSE (days) of predicted Gr = 2 is 4.4199, followed by ANN (4.4288 days) and GEP (4.8465 days). The RMSE for all three models is comparable. Therefore, these three models have similar modeling accuracy on Gr prediction.

The second critical stage is booting to late-heading (Gr ≈ 2.7 to 3.3), and 10 days before harvest (Gr ≈ 3.8), in which appropriate water control is required to increase the rice quality and yield [40,41]. The model performance assessed by RMSE (days) in Gr = 2.7 to 3.3 and Gr ≈ 3.8 is GEP > ANN > REG. The error was higher when Gr was greater than 3, and it became even large in Gr near to 4. This is likely because that rice grain is not sensitive to temperature in the last growth stage [8,42]. Therefore, a traditional linear algorithm to model this stage with GDD cannot generate good predictions.

The third critical stage is the tillering initiation, which increases panicles per plant by appropriate fertilization [43]. The fertilization timing is relatively flexible depending on rice cultivars and crop season—only one application is needed before effective tillering (Gr ≈ 1.0 to 1.3). However, it should be applied for some cultivars twice, for which the Gr is around 1.0 and 1.3. In this case, the REG model has the lowest RMSE in Gr = 1 (1.9651 days) followed by ANN and GEP (2.4451 and 2.6644 days). The differences error between REG and AI-based ANN and GEP are 0.48 days (≈11.5 h) and 0.6993 days (≈16.8 h), respectively. Considering the relatively flexible schedule for rice fertilization, all three models can be used for Gr modeling at this stage.

5. Conclusions

GDD is the most convenient index used to construct the rice growth prediction model for agronomists. However, the model displayed limited modeling accuracy for rice growth in the late-Reproductive phase and Ripening phase resulted from notable organ developments in these stages. Because of the importance of rice production and smart agriculture, GDD was conducted in AI algorithms-ANN and GEP, to predict rice growth variations. The results indicate that each algorithm can be adapted for specific purposes using the whole dataset combined with different rice cultivars and crop seasons. Based on RMSE, both AI-based ANN and GEP models are suitable for lifecycle, stage average prediction, and specific key stage prediction at the Gr of 3 and 4. Traditional REG methods have a good modeling ability when the specific key stage Gr is 1 and 2. Overall, the GEP model is recommended to farmers for conducting routine fieldwork, considering that the GEP model can be encrypted to Python-based code for further field environment monitoring system development, which may have lower hardware (e.g., CPU, RAM, etc.) requirements compared to ANN model. The accurate rice growth model indicates that it is possible to provide guidance for machines to operate fieldwork automatically based on the prediction

result, achieving the goal of smart and precise rice production. Onto further enhance the modeling ability on rice growth prediction, we suggest that future research focus on optimizing the model's parameters (e.g., the number of neurons, the number of hidden layers, etc.), transfer functions and fitness functions selection, and developing more sophisticated AI algorithms (e.g., generative adversarial network, etc.).

Author Contributions: Conceptualization, L.-W.L. and W.-S.L.; methodology, L.-W.L.; software, L.-W.L.; validation, X.-M.M. and W.-S.L.; formal analysis, L.-W.L. and K.-H.L.; investigation, K.-H.L. and C.-T.L.; resources, X.-M.M., C.-T.L., Y.-M.W. and W.-S.L.; writing—original draft preparation, L.-W.L.; writing—review and editing, X.-M.M., Y.-M.W. and W.-S.L.; visualization, L.-W.L.; supervision, X.-M.M., Y.-M.W. and W.-S.L.; project administration, Y.-M.W. and W.-S.L.; funding acquisition, L.-W.L. and W.-S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Ministry of Science and Technology (MOST), Taiwan (grant number: 109-2917-I-020-001 and 110-2637-B-020-005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors gratefully acknowledge the team of Hydraulic Laboratory and International Irrigation Research and Development Service Centre of NPUST for data collection and analysis, and the Department of Civil and Environmental Engineering, Texas A&M University for providing computing equipment. The authors also appreciate Ho-Hsien Chen from the College of Agriculture, NPUST for his comprehensive assistance and encouragement, and Joshua DiCaglio from the Department of English, Texas A&M University, and his working group, including Jacie Salas, Reagan McDonald, and Laura Gunn, for the English editing assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Awika, J.M. Major cereal grains production and use around the world. In *Advances in Cereal Science: Implications to Food Processing and Health Promotion*; American Chemical Society: Washington, DC, USA, 2011; Volume 1089, pp. 1–13.
2. Orlando, F.; Alali, S.; Vaglia, V.; Pagliarino, E.; Bacenetti, J.; Bocchi, S. Participatory approach for developing knowledge on organic rice farming: Management strategies and productive performance. *Agric. Syst.* **2020**, *178*, 102739. [CrossRef]
3. Zhou, T.; Wang, Y.; Huang, S.; Zhao, Y. Synthesis composite hydrogels from inorganic-organic hybrids based on leftover rice for environment-friendly controlled-release urea fertilizers. *Sci. Total Environ.* **2018**, *615*, 422–430. [CrossRef] [PubMed]
4. Assefa, Y.; Yadav, S.; Mondal, M.K.; Bhattacharya, J.; Parvin, R.; Sarker, S.R.; Rahman, M.; Sutradhar, A.; Prasad, P.V.; Bhandari, H. Crop diversification in rice-based systems in the polders of Bangladesh: Yield stability, profitability, and associated risk. *Agric. Syst.* **2021**, *187*, 102986. [CrossRef]
5. Perea, R.G.; Daccache, A.; Díaz, J.R.; Poyato, E.C.; Knox, J.W. Modelling impacts of precision irrigation on crop yield and in-field water management. *Precis. Agric.* **2018**, *19*, 497–512. [CrossRef]
6. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.; Jha, S.K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.* **2020**, *21*, 1–35. [CrossRef]
7. Fageria, N.; Baligar, V. Ameliorating soil acidity of tropical Oxisols by liming for sustainable crop production. *Adv. Agron.* **2008**, *99*, 345–399.
8. Yoshida, S. *Fundamentals of Rice Crop Science*; International Rice Research Institute: Manila, Philippines, 1981.
9. Wang, W.G.; Peng, S.Z.; Sun, F.C.; Xing, W.Q.; Luo, Y.F.; Xu, J.Z. Spatiotemporal variations of rice irrigation water requirements in the mid-lower reaches of Yangtze River under changing climate. *Adv. Water Sci.* **2012**, *23*, 656–664.
10. Yao, M.-H.; Chen, S.-H. The impact evaluation of climate change on the growth and yield of rice. *Crop Environ. Bioinform.* **2009**, *6*, 141–156.
11. Hsu, H.-H.; Chen, C.-T. Observed and projected climate change in Taiwan. *Meteorol. Atmos. Phys.* **2002**, *79*, 87–104. [CrossRef]
12. Sharifi, P.; Aminpanah, H.; Erfani, R.; Mohaddesi, A.; Abbasian, A. Evaluation of genotype× environment interaction in rice based on AMMI model in Iran. *Rice Sci.* **2017**, *24*, 173–180. [CrossRef]
13. Frizzell, D.; Branson, J.; Wilson, C., Jr.; Norman, R.; Moldenhauer, K.; Gibbons, J. Development of Degree-Day 50 thermal unit thresholds for new rice cultivars. *BR Wells Rice Res. Ser.-Ark. Agric. Exp. Stn. Univ. Ark.* **2011**, *591*, 187–193.
14. Counce, P.; Siebenmorgen, T.; Ambardekar, A. Rice reproductive development stage thermal time and calendar day intervals for six US rice cultivars in the Grand Prairie, Arkansas, over 4 years. *Ann. Appl. Biol.* **2015**, *167*, 262–276. [CrossRef]

15. Bouman, B.A.M.; Kropff, M.J.; Tuong, T.P.; Wopereis, M.; ten Berge, H.F.M.; van Laar, H.H. *ORYZA2000: Modeling Lowland Rice*; International Rice Research Institute and Wageningen University and Research Centre: Manila, Philippines, 2001.
16. Hoogenboom, G.; Porter, C.; Shelia, V.; Boote, K.; Singh, U.; White, J.; Hunt, L.; Ogoshi, R.; Lizaso, J.; Koo, J. *Decision Support System for Agrotechnology Transfer (DSSAT)*; Version 4.7; DSSAT Foundation: Gainesville, FL, USA, 2017. Available online: https://DSSAT.net (accessed on 14 May 2021).
17. Bonhomme, R. Bases and limits to using 'degree. day'units. *Eur. J. Agron.* **2000**, *13*, 1–10. [CrossRef]
18. Shaykewich, C. An appraisal of cereal crop phenology modelling. *Can. J. Plant Sci.* **1995**, *75*, 329–341. [CrossRef]
19. Stuerz, S.; Shrestha, S.P.; Schmierer, M.; Vu, D.H.; Hartmann, J.; Sow, A.; Razafindrazaka, A.; Abera, B.B.; Chuma, B.A.; Asch, F. Climatic determinants of lowland rice development. *J. Agron. Crop Sci.* **2020**, *206*, 466–477. [CrossRef]
20. Van Oort, P.; Saito, K.; Tanaka, A.; Amovin-Assagba, E.; Van Bussel, L.; Van Wart, J.; De Groot, H.; Van Ittersum, M.; Cassman, K.; Wopereis, M. Assessment of rice self-sufficiency in 2025 in eight African countries. *Glob. Food Secur.* **2015**, *5*, 39–49. [CrossRef]
21. Zhang, S.; Tao, F. Improving rice development and phenology prediction across contrasting climate zones of China. *Agric. For. Meteorol.* **2019**, *268*, 224–233. [CrossRef]
22. Ruml, M.; Vuković, A.; Milatović, D. Evaluation of different methods for determining growing degree-day thresholds in apricot cultivars. *Int. J. Biometeorol.* **2010**, *54*, 411–422. [CrossRef]
23. Gandhi, N.; Petkar, O.; Armstrong, L.J. Rice crop yield prediction using artificial neural networks. In Proceedings of the 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India, 15–16 July 2016; pp. 105–110.
24. Traore, S.; Zhang, L.; Guven, A.; Fipps, G. Rice yield response forecasting tool (YIELDCAST) for supporting climate change adaptation decision in Sahel. *Agric. Water Manag.* **2020**, *239*, 106242. [CrossRef]
25. Liu, L.-W.; Hsieh, S.-H.; Lin, S.-J.; Wang, Y.-M.; Lin, W.-S. Rice blast (magnaporthe oryzae) occurrence prediction and the key factor sensitivity analysis by machine learning. *Agronomy* **2021**, *11*, 771. [CrossRef]
26. Killenga, S.; Chuwa, C.; Mvukiye, N.; Zakayo, J.; Paul, I.; Kimaro, D. *Rice Production Manual*; International Institute of Tropical Agriculture: Ibadan, Nigeria, 2020.
27. Liu, Y.; Zhu, X.; He, X.; Li, C.; Chang, T.; Chang, S.; Zhang, H.; Zhang, Y. Scheduling of nitrogen fertilizer topdressing during panicle differentiation to improve grain yield of rice with a long growth duration. *Sci. Rep.* **2020**, *10*, 1–10. [CrossRef] [PubMed]
28. Counce, P.A.; Keisling, T.C.; Mitchell, A.J. A uniform, objective, and adaptive system for expressing rice development. *Crop Sci.* **2000**, *40*, 436–443. [CrossRef]
29. Taiwan Central Geological Survey. *Soil Profile Data, V.3.0*; National Geological Data Warehouse; Ministry of Economic Affairs: New Taipei City, Taiwan, 2011. Available online: https://gis3.moeacgs.gov.tw (accessed on 14 February 2020).
30. Gao, L.; Jin, Z.; Li, L. Photo-thermal models of rice growth duration for various varietal types in China. *Agric. For. Meteorol.* **1987**, *39*, 205–213. [CrossRef]
31. Gao, L.; Jin, Z.; Huang, Y.; Chen, H.; Li, B. *Rice Cultivational Simulation-Optimization-Decision Making System (RCSODS)*; China Agricultural Scientech Press: Beijing, China, 1992.
32. Lin, K.H. *Study on Using Heat Unit to Construct the Cultivation Calendar of Main Growth Stages in Rice (Oryza sativa L.)*; National Pingtung University of Science and Technology: Pingtung, Taiwan, 2019.
33. Zhang, T.; Zhu, J.; Yang, X. Non-stationary thermal time accumulation reduces the predictability of climate change effects on agriculture. *Agric. For. Meteorol.* **2008**, *148*, 1412–1418. [CrossRef]
34. McMaster, G.S.; Wilhelm, W. Growing degree-days: One equation, two interpretations. *Agric. For. Meteorol.* **1997**, *87*, 291–300. [CrossRef]
35. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
36. Ferreira, C. Gene expression programming: A new adaptive algorithm for solving problems. *Complex Syst.* **2001**, *13*, 87–129.
37. Ferreira, C. Gene expression programming in problem solving. In *Soft Computing and Industry*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 635–653.
38. Liu, L.-W.; Ma, X.; Wang, Y.-M.; Lu, C.-T.; Lin, W.-S. Using artificial intelligence algorithms to predict rice (*Oryza sativa* L.) growth rate for precision agriculture. *Comput. Electron. Agric.* **2021**, *187*, 106286. [CrossRef]
39. Zhang, S.; Tao, F.; Zhang, Z. Uncertainty from model structure is larger than that from model parameters in simulating rice phenology in China. *Eur. J. Agron.* **2017**, *87*, 30–39. [CrossRef]
40. Datta, A.; Ullah, H.; Ferdous, Z. Water management in rice. In *Rice Production Worldwide*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 255–277.
41. Fukao, T.; Harris, T.; Bailey-Serres, J. Evolutionary analysis of the Sub1 gene cluster that confers submergence tolerance to domesticated rice. *Ann. Bot.* **2009**, *103*, 143–150. [CrossRef]
42. Wu, G.; Wilson, L. Growth and yield response of rice to rice water weevil injury. *Environ. Entomol.* **1997**, *26*, 1191–1201. [CrossRef]
43. Wang, Y.; Lu, J.; Ren, T.; Hussain, S.; Guo, C.; Wang, S.; Cong, R.; Li, X. Effects of nitrogen and tiller type on grain yield and physiological responses in rice. *AoB Plants* **2017**, *9*, plx012. [CrossRef] [PubMed]