PHOTOSYNTHETICA
International Journal for Photosynthesis Research

DOI 10.32615/ps.2025.009 PHOTOSYNTHETICA 63 (1): 73-80, 2025

# Determination of rice (*Oryza sativa* L.) drought stress levels based on chlorophyll *a* fluorescence through independent component analysis

Q. XIA[*], H. TANG[**], J.L. TAN[***], S.I. ALLAKHVERDIEV[#], and Y. GUO[**,+]

*School of Electrical Engineering and Automation, Changshu Institute of Technology, 215500 Changshu, China*[*]
*Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, 214122 Wuxi, China*[**]
*Department of Chemical & Biomedical Engineering, University of Missouri, Columbia, MO 65211, USA*[***]
*Timiryazev Institute of Plant Physiology, RAS, Moscow, Russia*[#]

## Abstract

Sensing rice drought stress is crucial for agriculture, and chlorophyll *a* fluorescence (ChlF) is often used. However, existing techniques usually rely on defined feature points on the OJIP induction curve, which ignores the rich physiological information in the entire curve. Independent Component Analysis (ICA) can effectively preserve independent features, making it suitable for capturing drought-induced physiological changes. This study applies ICA and Support Vector Machine (SVM) to classify drought levels using the entire OJIP curve. The results show that the 20-dimensional ChlF features obtained by ICA provide superior classification performance, with *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient improving by 18.15%, 0.18, 0.17, 0.17, and 0.22, respectively, compared to the entire curve. This work provides a rice drought stress levels determination method and highlights the importance of applying dimension reduction methods for ChlF analysis. This work is expected to enhance stress detection using ChlF.

*Keywords*: chlorophyll *a* fluorescence; dimension reduction; drought; rice.

## Introduction

As the global population rises, the demand for rice increases (van Dijk *et al*. 2021, Mohidem *et al*. 2022), but water scarcity worsened by climate change threatens rice cultivation and food security (López-Pacheco *et al*. 2019). Detecting and mitigating drought stress in rice is essential for sustainable food production. Photosynthesis, crucial for plant survival, is affected by water status, with drought stress reducing photosynthetic efficiency (Xu *et al*. 2021). Real-time monitoring of crop drought stress is challenging due to the lack of universal detection techniques. Chlorophyll *a* fluorescence (ChlF), a noninvasive and reliable tool, offers valuable insights into plant health,

## Highlights

- Dimensionality reduction was applied to extract OJIP induction
- Dimensionality reduction may keep more information than the defined OJIP induction
- Dimensionality reduction improves rice drought stress level determination

*Received* 29 September 2024
*Accepted* 4 March 2025
*Published online* 27 March 2025

[+]Corresponding author
e-mail: guoy@jiangnan.edu.cn

*Abbreviations*: ChlF – chlorophyll *a* fluorescence; D0 – rice drought stress treatment for 0 h; D1 – rice drought stress treatment for 1 h; D2 – rice drought stress treatment for 2 h; D4 – rice drought stress treatment for 4 h; D012 – the rice dataset of D0, D1, and D2; D014 – the rice dataset of D0, D1, and D4; D024 – the rice dataset of D0, D2, and D4; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding; $M_o$ – approximated initial slope (in $ms^{-1}$) of the fluorescence transient; NF1~NF20 – the 1st–20th new feature when the OJIP induction is reduced to the new ChlF feature of 20 dimensions using ICA, respectively; OJIP – chlorophyll *a* fluorescence induction kinetics curve; PCA – Principal Component Analysis; SVD – Singular Value Decomposition; SVM – Support Vector Machine.

*Acknowledgment*: This work was partially supported by the Jiangsu Agricultural Science and Technology Innovation Fund (SCX (22)3115), the National Natural Science Foundation of China (No: 51961125102, 31771680), and the 111 Project (B23008).
*Conflict of interest*: The authors declare that they have no conflict of interest.

photosynthetic efficiency, and stress conditions (Maxwell and Johnson 2000), making it widely used for evaluating plant responses to environmental stresses, including drought (Krause and Weis 1991, Guo and Tan 2015).

Among them, OJIP induction is a commonly used protocol for measuring ChlF. The OJIP induction is very complex and contains much information, which can provide abundant information, especially when exploring the details and nonlinear response of photosynthetic systems. Substantial transient changes in ChlF were observed in plants under different types of environmental stress (Chen *et al*. 2021). Under drought conditions, water supply is limited, leading to the closure of plant stomata and decreasing $CO_2$ absorption, thereby affecting photosynthesis and electron transfer efficiency. It can also cause damage to plants, such as the destruction of PSII structure, which affects the emission of fluorescence. The OJIP curve under drought stress may show a delay between O and J points, indicating hindered electron transfer, while the P point may decrease, indicating a decrease in cooperative ability (Cornic 2000, Colom and Vazzana 2003, Zargar *et al*. 2017). By capturing photosynthesis and physiological responses, the OJIP curve provides a new approach for studying the mechanisms of rice response to drought and identifying resistant varieties.

The ChlF induction features are considered as a simplified and standardized description of the OJIP induction (Horaczek *et al*. 2020, Zagorchev *et al*. 2021), many studies on drought stress detection based on ChlF rely on a few defined feature values (O, J, I, and P values) on the ChlF induction and the features obtained through algebraic operations based on O, J, I, and P values for statistical analysis (Rastogi *et al*. 2020, Lima-Moro *et al*. 2022, Špánić *et al*. 2023, Chegini *et al*. 2024). However, the issue lies in overlooking the high-dimensional nature of OJIP data, which contains numerous features, and failing to consider the potential physiological information related to rice drought stress that may exist in the residual values of the entire OJIP curve, leading to reduced accuracy in stress identification. There is limited attempt to use the entire OJIP induction for drought stress detection, resulting in a large amount of information on the OJIP induction being abandoned. This may limit the potential of ChlF for rice drought stress detection. Therefore, in-depth research on the whole OJIP induction for rice detection should be paid more attention.

OJIP data is usually high-dimensional data containing a large number of features. With the rapid development of computer technology, artificial intelligence technology can process high-dimensional data and is adept at discovering complex patterns and correlations between data. Dimensionality reduction methods can simplify data structures, reduce redundant information, extract the most representative features, and thus improve the generalization ability and accuracy of models. Artificial intelligence technology plays a significant role in ChlF data analysis (Chen *et al*. 2022, Bartold and Kluczek 2023, Xia *et al*. 2023a,b), and even in making use of the information on the whole OJIP induction. For example, Xia *et al*. (2022) demonstrated that the entire OJIP induction contains more information on plant physiological states. The SVM model can be used as an effective tool for rice drought stress identification. By combining artificial intelligence technology and feature dimensionality reduction methods, OJIP data can be effectively processed and analyzed, more effectively mining information related to rice drought stress, and helping researchers understand the biological significance behind the data. Therefore, there is a need to find suitable feature dimensionality reduction methods to extract the main photosynthetic features from the OJIP curve and capture the key dynamic changes in the photosynthetic process of rice under drought conditions. However, only a limited number of feature extraction methods have been applied to OJIP curve analysis until now (Shomali *et al*. 2023). Existing approaches primarily focus on the application of single methods, these methods vary in their data processing capabilities, noise sensitivity, and biological interpretability, with little systematic research comparing their effectiveness. Thus, conducting comparative studies on the effectiveness of feature extraction techniques is essential for optimizing OJIP curve analysis, enhancing accuracy, and enabling broader applications.

In this paper, various dimensionality reduction methods were applied to the OJIP induction curves of rice under different drought stress levels to preserve ChlF information while reducing data complexity. Independent Component Analysis (ICA) was initially used to analyze the ChlF data due to its ability to preserve data independence, separate non-Gaussian signals, and enhance signal interpretability, which focuses on the independence of data components, enabling the identification of multiple independent physiological processes in rice under drought stress. ICA also reduces redundancy and noise, simplifying the data and improving analysis accuracy. Additionally, it is particularly effective for handling complex temporal data, revealing physiological changes in rice over time and providing deeper insights into stress responses and adaptation mechanisms. Subsequently, a Support Vector Machine (SVM) was applied to classify drought stress levels in rice based on the reduced OJIP data. Through a comparison of different dimensionality reduction methods, this study identifies new ChlF features related to drought stress, offering an efficient and cost-effective approach for assessing drought stress in rice.

## Materials and methods

**Experimental samples**: The experimental samples were rice plants grown for 40 d, and their growth area was Huai'an, Jiangsu, China. To achieve the same level of water status, the rice roots were fully immersed in water for 2 h before ChlF was measured (control group). The rice roots were then placed in a concentration of 20% polyethylene glycol-600 (PEG-600) for different drought stress level treatments (Awan *et al*. 2021, Peršić *et al*. 2022): rice drought stress treatment for 0 h – D0, rice drought stress treatment for 1 h – D1, rice drought stress treatment for 2 h – D2, and rice drought stress treatment for 4 h – D4. The number of rice plant samples treated with drought

stress levels for 0, 1, 2, and 4 h were 1,335, 1,093, 1,322, and 1,146, respectively.

**Fluorescence measurement**: The ChlF of the same rice leaf under different drought stress levels (D0, D1, D2, and D4) after 20 min of dark adaptation was measured by using the OJIP mode in *FluorPen*, *PSI*, *Photon Systems Instruments*, Czech Republic. The light intensity of exciting the ChlF was 2,400 μmol(photon) $m^{-2}s^{-1}$.

**Data analysis**: The involved ChlF datasets include (*1*) the OJIP induction (457 values) and (*2*) the ChlF induction parameters ($F_o$, $F_j$, $F_i$, $F_m$, $F_v$, $V_j$, $V_i$, $F_m/F_o$, $F_v/F_o$, $F_v/F_m$, $M_o$, Area, Fix Area, $S_m$, $S_s$, N, $\varphi P_o$, $\psi_o$, $\varphi E_o$, $\varphi D_o$, φPav, $PI_{ABS}$, ABS/RC, $TR_o$/RC, $ET_o$/RC, and $DI_o$/RC, as shown in Appendix, Stirbet and Govindjee 2011, Rachoski *et al*. 2015). This ChlF data has been analyzed in previous studies, confirming its effectiveness (Xia *et al*. 2022, 2023a,b). ChlF datasets need to be normalized according to Z-score standardization before analysis for all levels of drought stress. The Z-score normalization formula is $Z = (x - \mu)/\sigma$, where $x$ is the value of the ChlF data point, μ is the mean of the ChlF dataset, and σ is the standard deviation of the ChlF dataset. Principal Component Analysis (PCA), Isometric Feature Mapping (Isomap), Singular Value Decomposition (SVD), Local Linear Embedding (LLE), and Independent Component Analysis (ICA) were used to perform feature dimensionality reduction on the OJIP induction and identify the optimal dimensionality reduction method by comparing different approaches within the same dataset. An SVM model was used to classify rice plants under different drought stress levels (Xia *et al*. 2022). 80% of the samples were randomly selected as the training dataset and the remaining 20% as the testing dataset. The kernel function of SVM was the Gaussian kernel function (RBF), the penalty factor (C) was 10, and Gamma was the default value. The data analysis process was conducted on *Python 3.6*.

One-way analysis of variance (*ANOVA*) was performed on the ChlF data under different drought stress levels, and *Tukey*'s Honestly Significant Difference test (*Tukey* HSD) was performed with a *p*-value of 0.05 or 0.01 to test for significant differences in the data between different drought stress levels in *SPSS* (*IBM SPSS Amos 21*, Armonk, NY).

**Evaluating indicators**: *Accuracy* is a measure of the correct classification ratio of the SVM model on rice drought stress levels, usually used to evaluate the overall performance of the SVM model. The calculation formula for *Accuracy* is:

$$Accuracy = Num_{ture}/Num_{all} \quad (1)$$

The $Num_{ture}$ represents the number of rice samples under drought stress level correctly classified by the SVM model, and $Num_{all}$ represents the total number of rice samples.

*Precision* represents the proportion of rice samples correctly classified by the SVM model under a specific drought stress level, which is used to measure the correctness of the model for a group. The calculation formula for *Precision* is:

$$Precision = TP/(TP + FP) \quad (2)$$

The *TP* represents the number of rice samples under a specific drought stress level that the SVM model correctly classified, and the *FP* represents the number of rice samples under a specific drought stress level that the SVM model incorrectly classified.

*Recall* refers to the ratio of the number of rice samples successfully classified by the SVM model for a certain drought stress level to the actual total number of samples in that category, which is used to measure whether the model can capture all instances of a certain drought stress level in rice. The calculation formula for *Recall* is:

$$Recall = TP/(TP + FN) \quad (3)$$

The *TP* represents the number of rice samples correctly identified by the SVM model for a certain drought stress level, and the *FN* represents the number of rice samples incorrectly identified by the SVM model for other drought stress levels, but belonging to that drought stress level. The range of *Recall* is between 0 and 1, and the higher the value, the better the model performs in capturing samples of that drought stress level. The SVM model can better avoid misclassifying the real rice samples of that drought stress level as other drought stress level.

The *F1-score* comprehensively considers the *Accuracy* and *Recall* of the SVM model, and is the harmonic average of *Accuracy* and *Recall*. The calculation formula for *F1-score* is:

$$F1\text{-}score = 2 \times Precision \times Recall/(Precision + Recall) \quad (4)$$

The *F1-score* ranges from 0 to 1, and the higher the value, the better the classification performance of the SVM model at a certain drought stress level.

In this work, after the *Precision*, *Recall*, and *F1-score* were calculated for each drought stress level, the macro average method was used to integrate the *Precision*, *Recall*, and *F1-score*. Firstly, for each rice drought stress level, the *Recall* and *F1-score* for that drought stress level were calculated separately. Then, the *Precision*, *Recall*, and *F1-score* for all drought stress categories were averaged, without considering the number of rice samples for each drought stress category.

*Cohen*'s *Kappa* (abbreviated as *Kappa*) coefficient is a statistical indicator that can be used to measure the consistency between model identification and actual observation.

## Results

**OJIP induction under different drought levels**: The mean of rice OJIP induction under different drought stress levels is shown in Fig. 1S (*supplement*). Although Fig. 1 shows that ChlF intensity (D1, D2, and D4) was higher under drought stress compared to non-drought conditions (D0), no significant difference in ChlF intensity (D1, D2, and D4) was observed across different drought durations.

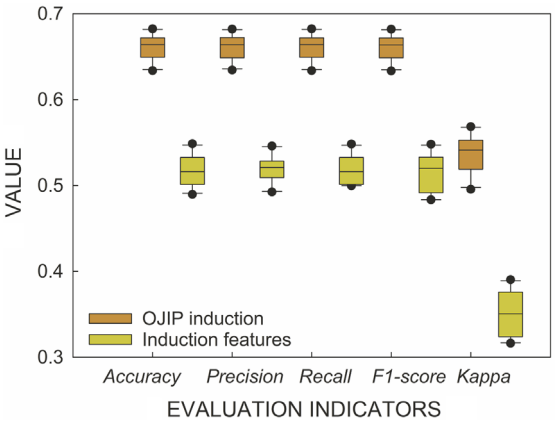


Fig. 1. Distribution of the evaluation indicators for different levels of rice drought stress classification based on different ChlF datasets. $n = 10$.

**Drought stress classification from different ChlF datasets without dimensionality reduction**: To compare the ability of two ChlF datasets (OJIP induction and induction features) to distinguish drought stress levels of rice, the computation was carried out through ten independent runs using the SVM model. Tables 1S (*supplement*) and 2S (*supplement*) show the maximum (Max), minimum (Min), and average evaluation indicators (*Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa*) about different ChlF datasets in the training dataset and test dataset. There is not much difference in accuracy between the model on the training and testing sets in Tables 1S and 2S, which indicates that the model has good generalization ability and can effectively maintain similar performance to the training dataset on the test dataset. It can be seen that the *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* of drought stress level classification using the OJIP induction as the SVM model input were higher, *i.e*, 65.50%, 0.65, 0.65, 0.65 and 0.54, respectively, and 12.86%, 0.12, 0.13, 0.13, and 0.17 higher than the induction features as the input of the SVM model (Table 2S).

The box plot in Fig. 1 shows the distribution of classification evaluation indicators for drought stress levels after ten independent runs of the SVM model on the two different ChlF datasets. It can be seen that the distribution of classification evaluation indicators of the SVM model was relatively concentrated under different ChlF datasets, indicating that the SVM model has good stability in the identification of rice drought stress levels. The *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* of rice drought stress level classification using the OJIP induction as input of the SVM model was higher than that using the 26 ChlF induction features as input dataset of the SVM model.

The confusion matrix when using the SVM model to classify drought stress levels for the two different ChlF datasets (the OJIP induction and the induction features) is shown in Fig. 2S (*supplement*). Fig. 2S*A* and Fig. 2S*B* are based on the OJIP induction and the induction features, respectively. The values in the confusion matrix represent the probability of classifying the true drought stress levels in rice represented by the vertical axis into the identified drought stress level of rice represented by the horizontal axis using the SVM model. The larger the diagonal value in the confusion matrix, the higher the classification accuracy. Comparing Fig. 2S*A* and Fig. 2S*B*, we can see that the classification accuracy of each drought stress level based on the OJIP induction was higher than that based on the induction features.

**Drought classification based on the OJIP induction with dimensionality reduction**: To map the whole OJIP induction to a new low dimensional feature space to reduce the dimensionality of the OJIP induction, create a new feature space, and retain the information of the original OJIP induction as much as possible, the original OJIP induction was processed linearly or nonlinearly to obtain new ChlF features. PCA, SVD, ICA, Isomap, and LLE were used to reduce the dimensionality of the original OJIP induction to obtain new ChlF features. Based on the new ChlF features obtained after the dimensionality reduction methods, SVM was used to identify different drought stress levels. The evaluation indicators (*Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa*) of the classification of different drought stress levels using SVM based on the new ChlF features obtained after feature dimensionality reduction are shown in Tables 1–5.

It can be seen that the new ChlF features obtained after dimensionality reduction of the original OJIP induction based on ICA has higher classification *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient than PCA, SVD, IOPMAP, and LLE (Tables 1–5). The SVM model can achieve optimal evaluation indicators for identifying drought stress levels based on the new ChlF features with 20 dimensions obtained by ICA; the *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient were 83.68%, 0.83, 0.82, 0.82, and 0.77, respectively (Tables 1–5).

The evaluation indicators of different drought stress levels using SVM under different dimensionality reduction methods for the OJIP induction are shown in Fig. 3S (*supplement*). The *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient were normalized to the maximum values in Fig. 3S. We can see from Fig. 3S that the classification *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient of ICA dimensionality reduction were the highest, significantly better than PCA, SVD, Isomap, and LLE. When using ICA to reduce the dimensionality of the OJIP induction to 20 dimensions of the new ChlF features, the evaluation indicators were the highest (Fig. 3S).

To evaluate the performance of the new ChlF features derived from the OJIP induction after reducing the ChlF data to 20 dimensions using the ICA method, the heatmap in Fig. 4S (*supplement*) illustrates the contribution of the OJIP induction (457 features) to the new ChlF features. The results reveal that OJIP induction contributed a specific value to the new ChlF features.

The distribution of evaluation indicators for ten independent runs of rice drought stress level classification using SVM based on new ChlF features with 20 dimensions after dimensionality reduction of the original

Table 1. *Accuracy* based on the new ChlF features. PCA – Principal Component Analysis; SVD – Singular Value Decomposition; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding. The bold font highlights the best value among all comparison indicators.

| Dimension | PCA [%] | SVD [%] | ICA [%] | Isomap [%] | LLE [%] |
|---|---|---|---|---|---|
| 5 | 48.67 | 45.05 | 56.70 | 37.87 | 31.70 |
| 10 | 60.39 | 48.10 | 75.81 | 39.24 | 32.08 |
| 15 | 62.76 | 48.43 | 79.78 | 39.66 | 33.09 |
| 20 | 62.86 | 48.41 | **83.68** | 39.95 | 33.53 |
| 25 | 62.95 | 48.38 | 82.22 | 40.45 | 34.47 |
| 30 | 62.95 | 48.37 | 80.49 | 40.36 | 34.76 |

Table 2. *Precision* based on the new ChlF features. PCA – Principal Component Analysis; SVD – Singular Value Decomposition; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding. The bold font highlights the best value among all comparison indicators.

| Dimension | PCA | SVD | ICA | Isomap | LLE |
|---|---|---|---|---|---|
| 5 | 0.48 | 0.45 | 0.57 | 0.38 | 0.29 |
| 10 | 0.60 | 0.48 | 0.76 | 0.39 | 0.32 |
| 15 | 0.62 | 0.48 | 0.80 | 0.40 | 0.33 |
| 20 | 0.63 | 0.48 | **0.83** | 0.40 | 0.34 |
| 25 | 0.63 | 0.48 | 0.82 | 0.40 | 0.35 |
| 30 | 0.63 | 0.48 | 0.80 | 0.40 | 0.36 |

Table 3. *Recall* based on the new ChlF features. PCA – Principal Component Analysis; SVD – Singular Value Decomposition; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding. The bold font highlights the best value among all comparison indicators.

| Dimension | PCA | SVD | ICA | Isomap | LLE |
|---|---|---|---|---|---|
| 5 | 0.48 | 0.44 | 0.57 | 0.37 | 0.30 |
| 10 | 0.60 | 0.47 | 0.76 | 0.38 | 0.31 |
| 15 | 0.62 | 0.47 | 0.79 | 0.39 | 0.32 |
| 20 | 0.62 | 0.47 | **0.82** | 0.39 | 0.33 |
| 25 | 0.62 | 0.47 | 0.82 | 0.40 | 0.34 |
| 30 | 0.62 | 0.47 | 0.80 | 0.39 | 0.34 |

OJIP induction by ICA is shown in Fig. 2. It shows that the classification *Accuracy*, *Precision*, *Recall*, and *F1-score* were all above 0.80, and the *Kappa* coefficient was above 0.74.

The confusion matrix when the SVM model was used to classify the drought stress levels from the new ChlF features obtained by ICA dimension reduction is shown in Fig. 3. We found that the classification *Accuracy* of each drought stress level was above 81%, the overall classification *Accuracy* was 83.56% (the average classification *Accuracy* of each drought stress level). In addition, 7.73% of the rice samples under drought stress treatment for 1 h (D1) were incorrectly classified into 4 h

Table 4. *F1-score* based on the new ChlF features. PCA – Principal Component Analysis; SVD – Singular Value Decomposition; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding. The bold font highlights the best value among all comparison indicators.

| Dimension | PCA | SVD | ICA | Isomap | LLE |
|---|---|---|---|---|---|
| 5 | 0.48 | 0.42 | 0.56 | 0.36 | 0.20 |
| 10 | 0.60 | 0.46 | 0.76 | 0.38 | 0.29 |
| 15 | 0.62 | 0.46 | 0.79 | 0.38 | 0.30 |
| 20 | 0.62 | 0.46 | **0.82** | 0.39 | 0.32 |
| 25 | 0.62 | 0.46 | 0.82 | 0.39 | 0.33 |
| 30 | 0.62 | 0.46 | 0.80 | 0.39 | 0.34 |

Table 5. *Kappa* based on the new ChlF features. PCA – Principal Component Analysis; SVD – Singular Value Decomposition; ICA – Independent Component Analysis; Isomap – Isometric Feature Mapping; LLE – Local Linear Embedding. The bold font highlights the best value among all comparison indicators.

| Dimension | PCA | SVD | ICA | Isomap | LLE |
|---|---|---|---|---|---|
| 5 | 0.31 | 0.26 | 0.42 | 0.16 | 0.07 |
| 10 | 0.47 | 0.30 | 0.68 | 0.18 | 0.08 |
| 15 | 0.50 | 0.31 | 0.73 | 0.19 | 0.10 |
| 20 | 0.50 | 0.31 | **0.77** | 0.19 | 0.10 |
| 25 | 0.50 | 0.30 | 0.76 | 0.20 | 0.12 |
| 30 | 0.50 | 0.30 | 0.74 | 0.20 | 0.12 |

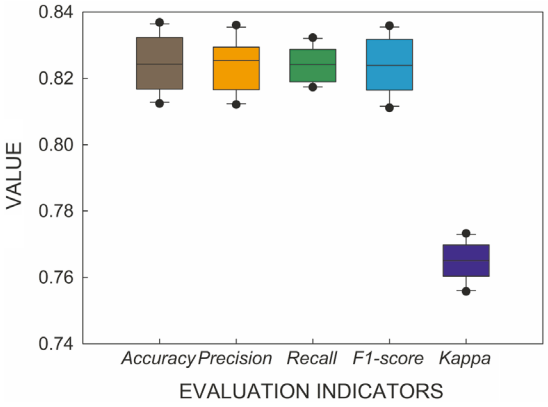


Fig. 2. Distribution of classification evaluation indicators for different levels of rice drought stress classification based on new ChlF features with 20 dimensions. $n = 10$.

of drought stress level treatment (D2), while 10.55% of the rice under drought stress treatment for 2 h (D2) was incorrectly classified into D4 (Fig. 3).

As shown in Fig. 5S (*supplement*), the new ChlF features obtained after ICA dimensionality reduction showed different changes with the extension of drought duration (the values in Fig. 5S were normalized to the maximum value). Table 3S (*supplement*) shows the *p*-value of the new ChlF features with 20 dimensions obtained after ICA dimension reduction for the OJIP induction between

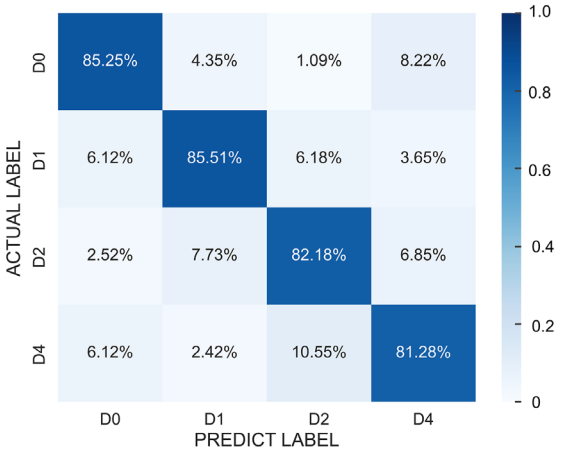


Fig. 3. The confusion matrix based on the new ChlF features with 20 dimensions using ICA. D0, D1, D2, and D4 represent rice drought stress treatment for 0 h, rice drought stress treatment for 1 h, rice drought stress treatment for 2 h, and rice drought stress treatment for 4 h, respectively.

different drought stress levels. The results indicate that most of the new ChlF features obtained by ICA were statistically different between different drought stress levels (Table 3S).

To investigate the correlation between known drought-stress biomarkers (such as $F_v/F_m$, $F_o/F_m$, $PI_{ABS}$) and the new ChlF features derived from ICA dimensionality reduction, Fig. 6S (*supplement*) presents the correlation heatmap between these new features and OJIP induction parameters. Fig. 6S demonstrates that some new features obtained through dimensionality reduction show significant correlations with the OJIP induction parameters.

To further evaluate the drought classification performance of the new ChlF features derived from the dimensionality reduction of OJIP induction using ICA across three drought stress levels, Table 4S (*supplement*) presents the maximum (Max), minimum (Min), and average values of the evaluation indicators (*Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa*) for classifying the three drought stress levels based on the new 20-dimensional ChlF features obtained after ICA dimensionality reduction. From Table 4S, we see that using the new ChlF features with a dimension of 20 obtained after ICA dimensionality reduction, the mean values of classification *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* can reach over 88%, 0.89, 0.87, 0.88, and 0.81, respectively.

The confusion matrix of three different drought stress levels based on the new ChlF features is presented in Fig. 7S (*supplement*). The classification *Accuracy* of each drought stress level was above 87%, and the overall classification *Accuracy* for D012, D014, and D024 were 90.52, 88.37, and 83.56%, respectively (Fig. 7S).

## Discussion

In plant physiology, the OJIP induction contains a large amount of plant physiological information, but high dimensionality leads to complex and time-consuming calculations. Machine learning methods have great advantages in processing large data and discovering key information, which helps reveal important features in plant physiological processes from ChlF OJIP induction.

The classification *Accuracy* of drought stress level using all data on the OJIP induction as the input dataset for the SVM model was higher than that only using the ChlF induction features as the input dataset for the SVM model (Table 2S). This result was expected, as OJIP measures dynamic changes in plants at various stages of photosynthesis, providing more comprehensive and integrated data information. In contrast, using only the O, J, I, and P feature points on the OJIP curve and the ChlF parameters obtained through mathematical operations based on these feature points may lose some key information, limiting the model's understanding of complex physiological states. The result further confirms that richer and more useful plant physiological information is hidden in the OJIP induction, which can help construct a more discriminative classification model for rice drought stress, making SVM models have stronger classification ability and robustness.

Applying the dimensionality reduction method in OJIP induction analysis is beneficial for better extracting ChlF information. This can be proven from the analysis results. Compared with the whole OJIP induction as the input dataset of the SVM model (Table 2S), the classification *Accuracy*, *Precision*, *Recall*, *F1-score*, and *Kappa* coefficient of the new ChlF feature projections with 20 dimensions obtained by ICA dimensionality reduction were improved by 18.15%, 0.18, 0.17, 0.17, and 0.22 (Table 2S *vs.* Tables 1–5), respectively. Dimensionality reduction reduces redundant information in the data, highlights key features of rice drought on the OJIP curve, and improves the accuracy of the SVM model in drought classification. This enables SVM classification to capture the characteristics of rice drought stress, reduces the risk of overfitting, improves the generalization ability of the model, and achieves more significant performance in drought classification. Compared to PCA, SVD, Isomap, LLE, and other methods, Tables 1–5 and Fig. 3S show that the best performance was achieved when classifying rice drought stress using the new ChlF features obtained through ICA dimensionality reduction. These feature dimensionality reduction methods (PCA, Isomap, SVD, LLE, ICA) differ fundamentally in their principles. PCA and SVD focus on linear dimensionality reduction by maximizing data variance, which is effective for capturing global variability in the data. In contrast, LLE and Isomap emphasize preserving the local structure of data, making them more suitable for nonlinear data. While PCA and SVD are well-suited for Gaussian data, they struggle to capture nonlinear and complex patterns. ICA, however, decomposes data into independent components, which is ideal for separating non-Gaussian signals and identifying independent physiological processes. This makes ICA particularly effective for analyzing complex biological signals. In the case of OJIP data, ICA enhances the identification of physiological changes in rice under

drought stress by removing redundancy and noise, preserving key features, and improving classification accuracy. As a result, ICA typically outperforms other dimensionality reduction methods in OJIP data classification. In the classification of three different levels of rice drought stress (D0, D1, and D4), the *Accuracy* and *Kappa* coefficient of using the dimensionality reduction method in this paper is better than that of Xia *et al*. (2022) directly using the entire OJIP curve as the input of SVM model for drought level classification, with improvement of 6.29% and 0.08%, respectively. In addition, after ICA performs feature dimensionality reduction on the OJIP data, the new ChlF features obtained are statistically different between different drought levels (Table 3S). This shows that the features processed by ICA dimensionality reduction can better reflect the changes in the physiological state of vegetation under different drought levels, providing more distinguishing and significant information for the classification of drought levels, and helping improve drought monitoring and classification. In addition, the OJIP induction parameters serve as established biomarkers for drought stress, and the new chlorophyll fluorescence features, obtained after dimensionality reduction, are correlated with these parameters (Fig. 6S). Therefore, the new ChlF features can provide additional information for drought stress.

These methods provide powerful tools to extract valuable information from large-scale and complex data. Artificial intelligence analysis methods are highly effective in handling complex nonlinear data, with strong adaptive capabilities that enable them to manage ChlF data under various environmental stresses, such as drought, temperature fluctuations, and light variations. These merits enhance the accuracy of OJIP curve analysis but also ensure its broad applicability across diverse environmental conditions. The ICA method in the field of OJIP curve analysis demonstrates its potential, but it also faces some challenges. As the size of the dataset increases, the scalability of the computational method becomes a key issue. In addition, it is necessary to conduct a thorough evaluation of the biological interpretation and effectiveness of the new ChlF feature projections obtained through ICA methods.

Dimensionality reduction techniques can help simplify large data, improve efficiency, and enhance model accuracy and generalization. We expect further development of dimensionality reduction techniques for OJIP curve analysis. This will address the computational challenges of larger datasets and facilitate the broader application of research findings in plant science. In future research, samples subjected to drought stress under natural field conditions could be used to validate the performance of the proposed method, while plants from diverse environmental and physiological conditions may provide insights into its generalizability. Ultimately, such advancements will impact the study and practical applications of plant growth, resistance, and adaptability.

**Conclusion**: In this work, feature dimensionality reduction methods were used to reduce the complexity of the OJIP induction, resulting in ChlF feature projection in the independent component spaces related to drought stress. These features outperformed traditional ChlF values and the entire OJIP induction in identifying rice under drought stress. This work presents a rice-physiological-based method for determining drought stress levels but also emphasizes the importance of dimension reduction in ChlF induction analysis. By improving the ability to detect drought stress, this approach is expected to support real-time crop assessment in breeding and agricultural monitoring systems.

## References


Awan S.A., Khan I., Rizwan M. *et al*.: Exogenous abscisic acid and jasmonic acid restrain polyethylene glycol-induced drought by improving the growth and antioxidative enzyme activities in pearl millet. – Physiol. Plantarum **172**: 809-819, 2021.

Bartold M., Kluczek M.: A machine learning approach for mapping chlorophyll fluorescence at inland wetlands. – Remote Sens.-Basel **15**: 2392, 2023.

Chegini S.N., Jafarinia M., Ghotbi-Ravandi A.A.: Unraveling the impacts of progressive drought stress on the photosynthetic light reaction of tomato: assessed by chlorophyll-*a* fluorescence and gene expression analysis. – Cell. Mol. Biol. **70**: 176-184, 2024.

Chen D., Zhang J., Zhang Z. *et al*.: Analyzing the effect of light on lettuce $F_v/F_m$ and growth by machine learning. – Sci. Hortic.-Amsterdam **306**: 111444, 2022.

Chen X., Zhou Y., Cong Y. *et al*.: Ascorbic acid-induced photosynthetic adaptability of processing tomatoes to salt stress probed by fast OJIP fluorescence rise. – Front. Plant Sci. **12**: 594400, 2021.

Colom M.R., Vazzana C.: Photosynthesis and PSII functionality of drought-resistant and drought-sensitive weeping lovegrass plants. – Environ. Exp. Bot. **49**: 135-144, 2003.

Cornic G.: Drought stress inhibits photosynthesis by decreasing stomatal aperture – not by affecting ATP synthesis. – Trends Plant Sci. **5**: 187-188, 2000.

Guo Y., Tan J.: Recent advances in the application of chlorophyll *a* fluorescence from photosystem II. – Photochem. Photobiol. **91**: 1-14, 2015.

Horaczek T., Dąbrowski P., Kalaji H.M. *et al*.: JIP-test as a tool for early detection of the macronutrients deficiency in *Miscanthus* plants. – Photosynthetica **58**: 507-517, 2020.

Krause A.G., Weis E.: Chlorophyll fluorescence and photosynthesis: the basics. – Annu. Rev. Plant Biol. **42**: 313-349, 1991.

Lima-Moro A., Bertoli S.C., Braga-Reis I. *et al*.: Photosynthetic activity and OJIP fluorescence with the application of a nutritional solution. – Acta Physiol. Plant. **44**: 67, 2022.

López-Pacheco I.Y., Silva-Núñez A., Salinas-Salazar C. *et al*.: Anthropogenic contaminants of high concern: Existence in water resources and their adverse effects. – Sci. Total Environ. **690**: 1068-1088, 2019.

Maxwell K., Johnson G.N.: Chlorophyll fluorescence – a practical guide. – J. Exp. Bot. **51**: 659-668, 2000.

Mohidem N.A., Hashim N., Shamsudin R., Man H.C.: Rice for food security: Revisiting its production, diversity, rice milling process and nutrient content. – Agriculture **12**: 741, 2022.

Peršić V., Ament A., Antunović Dunić J. *et al*.: PEG-induced physiological drought for screening winter wheat genotypes sensitivity – integrated biochemical and chlorophyll *a* fluorescence analysis. – Front. Plant Sci. **13**: 987702, 2022.

Rachoski M., Gazquez A., Calzadilla P. *et al.*: Chlorophyll fluorescence and lipid peroxidation changes in rice somaclonal lines subjected to salt stress. – Acta. Physiol. Plant. **37**: 117, 2015.

Rastogi A., Kovar M., He X. *et al.*: JIP-test as a tool to identify salinity tolerance in sweet sorghum genotypes. – Photosynthetica **58**: 518-528, 2020.

Shomali A., Aliniaeifard S., Bakhtiarizadeh M.R. *et al.*: Artificial neural network (ANN)-based algorithms for high light stress phenotyping of tomato genotypes using chlorophyll fluorescence features. – Plant Physiol. Biochem. **201**: 107893, 2023.

Špánić V., Šunić K., Duvnjak J. *et al.*: Chlorophyll *a* fluorescence during flag leaf senescence of field-grown winter wheat plants under drought conditions. – Ann. Appl. Biol. **183**: 80-92, 2023.

Stirbet A., Govindjee: On the relation between the Kautsky effect (chlorophyll *a* fluorescence induction) and photosystem II: basics and applications of the OJIP fluorescence transient. – J. Photoch. Photobio. B **104**: 236-257, 2011.

van Dijk M., Morley T., Rau M.L., Saghai Y.: A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. – Nat. Food **2**: 494-501, 2021.

Xia Q., Fu L.J., Tang H. *et al.*: Sensing and classification of rice (*Oryza sativa* L.) drought stress levels based on chlorophyll fluorescence. – Photosynthetica **60**: 102-109, 2022.

Xia Q., Tang H., Fu L. *et al.*: Determination of $F_v/F_m$ from chlorophyll *a* fluorescence without dark adaptation by an LSSVM model. – Plant Phenomics **5**: 0034, 2023a.

Xia Q., Tang H., Fu L. *et al.*: A drought stress-sensing technique based on wavelet entropy of chlorophyll fluorescence excited with pseudo-random binary sequence. – Comput. Electron. Agr. **210**: 107933, 2023b.

Xu H.-J., Wang X.-P., Zhao C.-Y.: Drought sensitivity of vegetation photosynthesis along the aridity gradient in northern China. – Int. J. Appl. Earth. Obs. Geoinf. **102**: 102418, 2021.

Zagorchev L., Atanasova A., Albanova I. *et al.*: Functional characterization of the photosynthetic machinery in *Smicronix* galls on the parasitic plant *Cuscuta campestris* by JIP-test. – Cells **10**: 1399, 2021.

Zargar S.M., Gupta N., Nazir M. *et al.*: Impact of drought on photosynthesis: Molecular perspective. – Plant Gene **11**: 154-159, 2017.

Appendix. ChlF induction parameters and definitions.

| OJIP parameters | Definition |
|---|---|
| $F_o$ | Minimum ChlF under dark adaptation |
| $F_j$ | ChlF at J point |
| $F_i$ | ChlF at I point |
| $F_m$ | Maximum ChlF intensity |
| $F_v = F_m - F_o$ | Variable ChlF under dark adaptation |
| $V_i = (F_i - F_o)/(F_m - F_o)$ | Relative variable fluorescence intensity at the I step |
| $V_j = (F_j - F_o)/(F_m - F_o)$ | Relative variable fluorescence intensity at the J step |
| $F_v/F_o = (F_m - F_o)/F_o$ | Quantum efficiency of photosystem II |
| $F_m/F_o$ | Electron transport through photosystem II |
| $F_v/F_m = (F_m - F_o)/F_m$ | Maximum photochemical quantum yield of photosystem II in the dark |
| $M_o = 4 \times (F_{300} - F_o)/(F_m - F_o)$ | Approximated initial slope (in $ms^{-1}$) of the fluorescence transient |
| Area | The area between the ChlF curve and $F_m$ (minus background) |
| Fix Area | Fluorescence curve area between $F_{40\mu s}$ and $F_{1s}$ |
| $S_s$ | ChlF enhancement complementary area standardized by O–J phase |
| $S_m = Area/F_v$ | The area between the OJIP induction curve and $F_m$ after standardization |
| $N = (M_o \times S_m)/V_j$ | The number of times $Q_A$ is restored during the time period from the start of illumination to reaching $F_m$ |
| $\varphi E_o = 1 - (F_o/F_m) \times (F_v/F_m)$ | Quantum yield of electron transport |
| $\varphi P_o = 1 - (F_o/F_m)$(or $F_v/F_m$) | Maximum quantum yield of PSII |
| $\varphi D_o = 1 - \varphi P_o - (F_o/F_m)$ | Quantum yield of energy dissipation |
| $\psi_o = 1 - V_j$ | The probability that a trapped exciton moves an electron further than $Q_A^-$ |
| $\varphi Pav = P_o \times (1 - V_j)$ | The average quantum yield of primary photochemistry (from $t_0$ to $t_{Fm}$) |
| $ABS/RC = (M_o/V_j)(1/\varphi P_o)$ | Absorption per reaction center |
| $PI_{ABS} = 4[(F_{300} - F_o) \times (F_m - F_j) \times F_m]/[(F_m - F_o)(F_j - F_o) \times F_o]$ | Performance index for energy conservation from exciton to the reduction of intersystem electron acceptors |
| $ET_o/RC = (M_o/V_j)(1 - V_j)$ | Electron transport per reaction center |
| $TR_o/RC = M_o \times (1/V_j)$ | Trapped energy flux per reaction center |
| $DI_o/RC = ABS/RC - TR_o/RC$ | Dissipation per reaction center (at t = 0) |