



a. Original Image



b. Segmented Image

Figure S1. Pseudo RGB image⁹(R: 650 nm, G:550 nm, B: 450 nm) and corresponding Segmenated image (R: 600 nm, G:500 nm, B: 420 nm)

SUPPLEMENTARY DATA

Supplementary Table S1. Machine learning models for the identification of drought stress in wheat under variable nitrogen levels. This shows the characteristics of each model and why they were chosen for this study.

Model	Characteristics	Reason for choosing the model
Support Vector Machine (SVM)	<ul style="list-style-type: none"> Supervised learning model for classification and regression tasks. Principle is to find the optimal hyperplane that separates distinct sample classes with the highest margin Effective when the number of features is greater than the number of samples 	<ul style="list-style-type: none"> Can handle high-dimensional dataset Good for linear and non-linear data and is robust to overfitting.
Random Forest (RF)	<ul style="list-style-type: none"> A meta estimator that fits a number of decision tree classifiers on various sub-sampling of dataset increase predictive accuracy and control over-fitting. Uses bootstrapped datasets and randomly selects subsets of features for tree construction. Reduces overfitting by averaging the predictions of multiple trees. 	<ul style="list-style-type: none"> RF can identify sensitive/ important features of a dataset which is useful for understanding driving features. RF can handle noisy data and outliers, making it suitable for this work RF approach reduces the likelihood of overfitting compared to single decision trees hence used in this study.
Deep Neural Network (DNN)	<ul style="list-style-type: none"> Has multiple layers (input, hidden, output) with interconnected nodes (neurons). Capable of modelling complex, non-linear relationships between inputs and outputs. 	<ul style="list-style-type: none"> DNNs model is ideal when dealing with large, complex and non-linear datasets that some traditional ML models may be unable to handle. DNN modelling is capable of automatically learning relevant features from raw data, reducing the need for manual feature engineering.

- State-of-the-art performance: Often achieves superior performance in tasks like image and speech recognition, natural language processing, etc

Supplementary Table S2. Correlation coefficients of proposed indices with g_s and P_n (RDI= drought-nitrogen ratio index, NDDI= Normalized difference drought-nitrogen index and DDI= Drought difference index).

No.	RDI Equation	Correlation coefficient (r)		No.	NDDI Equation	Correlation (r)		No.	DDI Equation	Correlation coefficient (r)	
		P_n	g_s			P_n	g_s			P_n	g_s
RDI1	$\lambda 557 / \lambda 669$	-0.34	-0.54	NDDI 1	$(\lambda 557 - \lambda 669) / (\lambda 557 + \lambda 669)$	-0.65	-0.6	DDI 1	$\lambda 557 - \lambda 669$	-0.28	-0.44
RDI 2	$\lambda 557 / \lambda 674$	-0.09	-0.04	NDDI 2	$(\lambda 557 - \lambda 674) / (\lambda 557 + \lambda 674)$	-0.68	-0.63	DDI 2	$\lambda 557 - \lambda 674$	0.22	-0.41
RDI 3	$\lambda 557 / \lambda 722$	-0.58	-0.50	NDDI 3	$(\lambda 557 - \lambda 722) / (\lambda 557 + \lambda 722)$	-0.72	-0.66	DDI 3	$\lambda 557 - \lambda 722$	0.51	-0.12
RDI 4	$\lambda 557 / \lambda 940$	-0.21	-0.54	NDDI 4	$(\lambda 557 - \lambda 940) / (\lambda 557 + \lambda 940)$	-0.70	-0.67	DDI 4	$\lambda 557 - \lambda 940$	0.09	0.04
RDI 5	$\lambda 557 / \lambda 957$	-0.55	-0.55	NDDI 5	$(\lambda 557 - \lambda 957) / (\lambda 557 + \lambda 957)$	-0.44	-0.23	DDI 5	$\lambda 557 - \lambda 957$	-0.15	-0.01
RDI 6	$\lambda 557 / \lambda 636$	-0.63	-0.55	NDDI 6	$(\lambda 557 - \lambda 636) / (\lambda 557 + \lambda 636)$	-0.28	-0.20	DDI 6	$\lambda 557 - \lambda 636$	-0.41	-0.39
RDI 7	$\lambda 557 / \lambda 683$	-0.45	-0.55	NDDI 7	$(\lambda 557 - \lambda 683) / (\lambda 557 + \lambda 683)$	-0.65	-0.60	DDI 7	$\lambda 557 - \lambda 683$	-0.33	-0.42
RDI 8	$\lambda 557 / \lambda 542$	-0.69	-0.55	NDDI 8	$(\lambda 557 - \lambda 542) / (\lambda 557 + \lambda 542)$	-0.69	-0.60	DDI 8	$\lambda 557 - \lambda 542$	-0.28	-0.42
RDI 9	$\lambda 669 / \lambda 674$	-0.69	-0.55	NDDI 9	$(\lambda 669 - \lambda 674) / (\lambda 669 + \lambda 674)$	-0.61	-0.60	DDI 9	$\lambda 669 - \lambda 674$	0.45	-0.40
RDI 10	$\lambda 669 / \lambda 722$	-0.70	-0.55	NDDI 10	$(\lambda 669 - \lambda 722) / (\lambda 669 + \lambda 722)$	-0.68	-0.59	DDI 10	$\lambda 669 - \lambda 722$	0.45	-0.38
RDI 11	$\lambda 669 / \lambda 940$	-0.72	-0.67	NDDI 11	$(\lambda 669 - \lambda 940) / (\lambda 669 + \lambda 940)$	-0.47	-0.58	DDI 11	$\lambda 669 - \lambda 940$	0.05	-0.09
RDI 12	$\lambda 669 / \lambda 957$	-0.71	-0.63	NDDI 12	$(\lambda 669 - \lambda 957) / (\lambda 669 + \lambda 957)$	-0.07	-0.58	DDI 12	$\lambda 669 - \lambda 957$	0.20	0.03
RDI 13	$\lambda 669 / \lambda 636$	0.20	0.20	NDDI 13	$(\lambda 669 - \lambda 636) / (\lambda 669 + \lambda 636)$	-0.78	-0.63	DDI 13	$\lambda 669 - \lambda 636$	-0.05	-0.01
RDI 14	$\lambda 669 / \lambda 683$	-0.24	-0.08	NDDI 14	$(\lambda 669 - \lambda 683) / (\lambda 669 + \lambda 683)$	-0.49	-0.56	DDI 14	$\lambda 669 - \lambda 683$	-0.48	-0.45
RDI 15	$\lambda 669 / \lambda 542$	0.61	0.58	NDDI 15	$(\lambda 669 - \lambda 542) / (\lambda 669 + \lambda 542)$	-0.41	-0.30	DDI 15	$\lambda 669 - \lambda 542$	-0.48	-0.43
RDI 16	$\lambda 674 / \lambda 722$	0.08	0.45	NDDI 16	$(\lambda 674 - \lambda 722) / (\lambda 674 + \lambda 722)$	0.22	0.16	DDI 16	$\lambda 674 - \lambda 722$	-0.35	-0.40

RDI 17	$\lambda 674 / \lambda 940$	0.66	0.06	NDDI 17	$(\lambda 674 - \lambda 940) / (\lambda 674 + \lambda 940)$	-0.06	-0.15	DDI 17	$\lambda 674 - \lambda 940$	-0.62	-0.36
RDI 18	$\lambda 674 / \lambda 957$	-0.21	-0.01	NDDI 18	$(\lambda 674 - \lambda 957) / (\lambda 674 + \lambda 957)$	-0.75	-0.52	DDI 18	$\lambda 674 - \lambda 957$	-0.01	-0.07
RDI 19	$\lambda 674 / \lambda 636$	-0.55	-0.07	NDDI 19	$(\lambda 674 - \lambda 636) / (\lambda 674 + \lambda 636)$	-0.47	-0.53	DDI 19	$\lambda 674 - \lambda 636$	0.08	0.03
RDI 20	$\lambda 674 / \lambda 683$	0.58	-0.11	NDDI 20	$(\lambda 674 - \lambda 683) / (\lambda 674 + \lambda 683)$	-0.59	-0.49	DDI 20	$\lambda 674 - \lambda 683$	-0.05	-0.01
RDI 21	$\lambda 674 / \lambda 542$	-0.53	-0.14	NDDI 21	$(\lambda 674 - \lambda 542) / (\lambda 674 + \lambda 542)$	-0.62	-0.48	DDI 21	$\lambda 674 - \lambda 542$	-0.44	-0.38
RDI 22	$\lambda 722 / \lambda 940$	-0.07	-0.17	NDDI 22	$(\lambda 722 - \lambda 940) / (\lambda 722 + \lambda 940)$	-0.74	-0.65	DDI 22	$\lambda 722 - \lambda 940$	-0.33	-0.35
RDI 23	$\lambda 722 / \lambda 957$	-0.22	-0.19	NDDI 23	$(\lambda 722 - \lambda 957) / (9\lambda 722 + \lambda 957)$	-0.67	0.67	DDI 23	$\lambda 722 - \lambda 957$	-0.52	-0.30
RDI 24	$\lambda 722 / \lambda 636$	-0.62	-0.20	NDDI 24	$(\lambda 722 - \lambda 636) / (\lambda 722 + \lambda 636)$	-0.53	-0.63	DDI 24	$\lambda 722 - \lambda 636$	-0.01	-0.05
RDI 25	$\lambda 722 / \lambda 683$	-0.15	-0.10	NDDI 25	$(\lambda 722 - \lambda 683) / (\lambda 722 + \lambda 683)$	-0.23	-0.63	DDI 25	$\lambda 722 - \lambda 683$	0.01	0.03
RDI 26	$\lambda 722 / \lambda 542$	0.08	0.06	NDDI 26	$(\lambda 722 - \lambda 542) / (\lambda 722 + \lambda 542)$	-0.03	-0.63	DDI 26	$\lambda 722 - \lambda 542$	-0.19	-0.01
RDI 27	$\lambda 940 / \lambda 957$	-0.45	-0.28	NDDI 27	$(\lambda 940 - \lambda 957) / (\lambda 940 + \lambda 957)$	-0.78	-0.63	DDI 27	$\lambda 940 - \lambda 957$	-0.41	-0.31
RDI 28	$\lambda 940 / \lambda 636$	-0.37	-0.57	NDDI 28	$(\lambda 940 - \lambda 636) / (\lambda 940 + \lambda 636)$	-0.74	-0.61	DDI 28	$\lambda 940 - \lambda 636$	-0.26	-0.26
RDI 29	$\lambda 940 / \lambda 683$	-0.29	-0.22	NDDI 29	$(\lambda 940 - \lambda 683) / (\lambda 940 + \lambda 683)$	-0.7	-0.63	DDI 29	$\lambda 940 - \lambda 683$	-0.22	-0.03
RDI 30	$\lambda 940 / \lambda 542$	-0.68	-0.62	NDDI 30	$(\lambda 940 - \lambda 542) / (\lambda 940 + \lambda 542)$	-0.56	-0.63	DDI 30	$\lambda 940 - \lambda 542$	0.23	0.03
RDI 31	$\lambda 957 / \lambda 636$	-0.72	-0.65	NDDI 31	$(\lambda 957 - \lambda 636) / (\lambda 957 + \lambda 636)$	-0.72	-0.67	DDI 31	$\lambda 957 - \lambda 636$	-0.16	-0.01
RDI 32	$\lambda 957 / \lambda 683$	-0.33	-0.61	NDDI 32	$(\lambda 957 - \lambda 683) / (\lambda 957 + \lambda 683)$	0.22	0.35	DDI 32	$\lambda 957 - \lambda 683$	-0.22	-0.19
RDI 33	$\lambda 957 / \lambda 542$	-0.23	-0.23	NDDI 33	$(\lambda 957 - \lambda 542) / (\lambda 957 + \lambda 542)$	-0.51	0.01	DDI 33	$\lambda 957 - \lambda 542$	-0.53	-0.02
RDI 34	$\lambda 636 / \lambda 683$	-0.25	-0.20	NDDI 34	$(\lambda 636 - \lambda 683) / (\lambda 636 + \lambda 683)$	-0.75	-0.67	DDI 34	$\lambda 636 - \lambda 683$	0.22	0.03
RDI 35	$\lambda 636 / \lambda 542$	-0.74	-0.67	NDDI 35	$(\lambda 636 - \lambda 542) / (\lambda 636 + \lambda 542)$	-0.5	-0.6	DDI 35	$\lambda 636 - \lambda 542$	-0.25	-0.01
RDI 36	$\lambda 683 / \lambda 542$	-0.68	-0.60	NDDI 36	$(\lambda 683 - \lambda 542) / (\lambda 683 + \lambda 542)$	-0.55	-0.6	DDI 36	$\lambda 683 - \lambda 542$	-0.21	-0.01



Supplementary Table S3. A one-way ANOVA test on the F1-score for all the models.
The dissimilar lower-case group (a, b, c) represents a significant difference with $p < 0.05$.

Training dataset	Models		
	RF	SVM	DNN
Known VIs	0.925 ^a	0.881 ^b	0.933 ^a
Proposed VIs	0.911 ^a	0.93 ^b	0.949 ^c
Combined VIs	0.984 ^a	0.982 ^a	0.979 ^a
PCA Features	0.962 ^a	0.94 ^b	0.900 ^c