The predictive power of regression models to determine grass weed infestations in cereals based on drone imagery – statistical and practical aspects
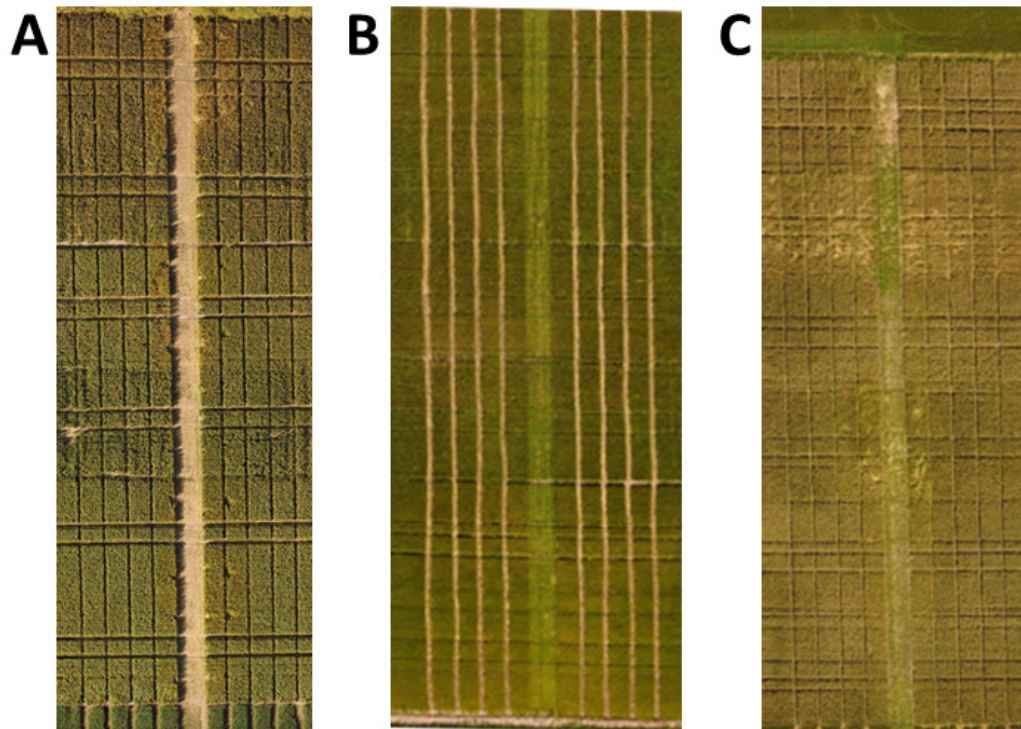

Figure S1. Orthomosaic from UAV imaging of selected fields to illustrate the visible plot variability due to weed infestations. A: Field 2. B: Field 3. C: Field 4.

Table S1: AIC differences (AIC – min AIC) for the prediction models based on all data for each field for different vegetation indices and different functional form of the vegetation index. Min AIC for each field is marked as bold.

Vegetation index	Functional form	Field 1	Field 2	Field 3	Field 4	Field 5	Field 6
Blue	Lin	2.55	39.82	1.20	35.29	9.34	**0.00**
	Exp	**0.00**	35.59	28.86	34.40	75.93	42.72
	Log	2.93	39.59	1.55	35.29	23.95	4.35
ExG	Lin	3.52	39.02	23.10	35.05	40.56	5.80
	Exp	3.55	39.07	22.97	35.05	42.16	6.49
	Log	3.44	38.60	23.53	35.02	26.33	2.70
Red index	Lin	6.19	9.68	16.72	**0.00**	26.34	18.04
	Exp	2.88	34.65	30.62	30.21	78.52	45.86
	Log	6.22	14.18	16.02	1.28	34.70	-
R/B	Lin	3.20	5.14	4.85	30.54	63.59	46.80
	Exp	3.40	**0.00**	28.90	31.70	76.49	47.23
	Log	3.11	7.17	**0.00**	30.30	57.58	46.58
NDI	Lin	5.61	20.43	29.56	31.81	72.43	22.30
	Exp	5.63	20.64	29.54	31.84	72.47	22.98
	Log	-	-	-	-	-	-
ExG - ExR	Lin	5.08	30.84	29.75	33.90	27.09	20.94
	Exp	7.31	39.16	29.72	31.84	86.33	52.76
	Log	-	-	-	-	-	-
R+G+B	Lin	4.32	33.54	28.87	26.27	**0.00**	0.79
	Exp	6.30	34.71	30.62	30.90	75.94	48.69
	Log	4.42	33.90	29.00	26.37	8.10	5.43

Table S2: RMSE for the prediction models based on all data for each field for different vegetation indices and different functional form of the vegetation index. Min RMSE for each field is marked as bold.

Vegetation index	Functional form	Field 1	Field 2	Field 3	Field 4	Field 5	Field 6
Blue	Lin	1.68	143.92	**95.21**	439.22	449.09	80.10
	Exp	1.69	156.49	149.99	431.32	>10000	120.85
	Log	1.69	142.87	96.49	439.43	535.10	84.76
ExG	Lin	1.66	146.42	121.63	434.83	624.91	85.16
	Exp	1.66	146.43	121.22	434.88	633.12	85.99
	Log	**1.65**	146.28	123.00	433.71	545.25	81.25
Red index	Lin	1.76	122.12	111.77	**311.60**	551.02	96.45
	Exp	1.89	3204.22	>10000	538.85	2001.59	185.15
	Log	-	-	-	-	-	-
R/B	Lin	1.71	102.74	101.10	398.01	782.01	112.07
	Exp	1.72	**91.56**	130.35	411.61	807.40	112.04
	Log	1.71	106.36	96.70	396.05	751.20	112.46
NDI	Lin	1.73	123.29	128.15	424.09	704.33	99.68
	Exp	1.73	123.70	128.11	424.25	704.81	100.35
	Log	-	-	-	-	-	-
ExG - ExR	Lin	1.72	134.05	132.98	442.18	569.47	95.03
	Exp	3.88	143.07	176.26	433.96	769.93	135.54
	Log	-	-	-	-	-	-
R+G+B	Lin	1.70	137.89	125.01	409.21	**400.11**	**79.91**
	Exp	1.98	>10000	>10000	529.38	>10000	114.15
	Log	1.71	138.84	125.37	409.74	444.02	84.68

Table S3: Average R^2 for training sets of different strategies for selecting a subset of data to train the linear prediction model. Results are given for different percentages (number of plots) used as training sets.

Field	VI	Random selection			Selection from extremes		Selection from the whole range		
		10% (6)	20% (12)	30% (18)	10% (6)	20% (12)	8% (5)	16% (10)	25% (15)
1	Blue	0.24	0.14	0.11	0.29	0.16	0.32	0.19	0.17
2	R/B	0.47	0.45	0.45	0.56	0.50	0.55	0.51	0.51
3	Blue	0.46	0.41	0.40	0.82	0.80	0.53	0.50	0.49
4	Red index	0.49	0.46	0.46	0.64	0.59	0.57	0.50	0.47
5	R+G+B	0.62	0.69	0.74	0.84	0.82	0.80	0.81	0.80
6	Blue	0.68	0.62	0.59	0.91	0.90	0.63	0.60	0.59

Table S4: Average RMSE for the training sets of different strategies for selecting a subset of data to train the prediction model. Results are given for different percentages (number of plots) used as training sets.

Field	VI	Random selection			Selection from extremes		Selection from the whole range		
		10% (6)	20% (12)	30% (18)	10% (6)	20% (12)	8% (5)	16% (10)	25% (15)
1	Blue	1.61	2.05	2.21	3.90	3.13	2.76	3.11	3.17
2	R/B	96.78	120.68	129.48	178.69	162.69	129.62	144.95	149.01
3	Blue	94.88	107.49	112.45	75.99	86.04	94.88	102.24	105.52
4	Red index	291.63	345.64	359.27	372.94	414.03	301.63	342.12	359.35
5	R+G+B	335.40	453.15	482.34	573.16	667.39	350.41	462.41	494.99
6	Blue	68.15	83.61	90.55	55.19	61.24	86.75	92.38	94.89

Table S5: Sensitivity, specificity, and balanced accuracy in the training sets for different strategies for selecting a subset of data to train the linear prediction model to describe the association between vegetation index and grass weed density. Results are given for different percentages (number of plots) used as training sets.

	Field	VI	Random selection			Selection from extremes		Selection from the whole range		
			10% (6)	20% (12)	30% (18)	10% (6)	20% (12)	8% (5)	16% (10)	25% (15)
Sensitivity										
	1	Blue	0.61	0.62	0.63	0.59	0.56	0.66	0.64	0.64
	2	R/B	0.64	0.78	0.83	0.69	0.80	0.80	0.87	0.90
	3	Blue	1	1	1	0.99	1	0.99	1	1
	4	Red	0.89	0.88	0.86	0.86	0.90	0.74	0.79	0.80
	5	R+G+B	0.69	0.74	0.78	0.81	0.92	0.75	0.80	0.81
	6	Blue	0.88	0.90	0.92	0.89	0.95	0.86	0.89	0.91
Specificity										
	1	Blue	0.75	0.67	0.66	0.90	0.74	0.84	0.70	0.67
	2	R/B	0.82	0.77	0.76	0.91	0.84	0.85	0.78	0.76
	3	Blue	-	-	-	-	-	-	-	-
	4	Red	0.84	0.75	0.73	0.92	0.87	0.88	0.81	0.77
	5	R+B+G	0.84	0.78	0.80	0.96	0.97	0.91	0.80	0.78
	6	Blue	0.79	0.77	0.76	0.90	0.82	0.79	0.76	0.76
Accuracy										
	All		0.81	0.81	0.81	0.87	0.87	0.84	0.82	0.82

Table S6: Sensitivity, specificity, and balanced accuracy in the training sets for different strategies for selecting a subset of data to train the prediction model using a logistic regression model to describe the association between vegetation index and probability of grass weed density. Results are given for different percentages (number of plots) used as training sets.

	Field	VI	Random selection			Selection from extremes		Selection from the whole range		
			10% (6)	20% (12)	30% (18)	10% (6)	20% (12)	8% (5)	16% (10)	25% (15)
Sensitivity										
	1	Blue	0.47	0.48	0.46	0.45	0.46	0.48	0.51	0.48
	2	R/B	0.49	0.58	0.61	0.61	0.71	0.56	0.62	0.63
	3	Blue	1	1	1	1	1	1	1	1
	4	Red	0.81	0.70	0.63	0.75	0.65	0.45	0.47	0.48
	5	R+G+B	0.51	0.49	0.49	0.57	0.63	0.45	0.50	0.49
	6	Blue	0.81	0.86	0.88	0.77	0.83	0.73	0.84	0.87
Specificity										
	1	Blue	0.83	0.75	0.72	0.86	0.79	1	0.85	0.81
	2	R/B	0.86	0.80	0.78	0.93	0.93	1	0.90	0.87
	3	Blue	-	-	-	-	-	-	-	-
	4	Red	0.98	0.95	0.94	1	1	1	0.90	0.85
	5	R+B+G	0.88	0.79	0.77	1	1	1	0.84	0.78
	6	Blue	1	1	1	1	1	1	1	1
Accuracy										
	All		0.80	0.78	0.77	0.83	0.83	0.81	0.79	0.77