Use of machine learning for monitoring the growth stages of an agricultural crop

Shara Ahmed, Nabanita Basu, Catherine E. Nicholson, Simon R. Rutter, John R. Marshall, Justin J. Perry and John R Dean

Department of Applied Sciences, Northumbria University, Ellison Building, Newcastle upon Tyne, NE1 8ST, UK

Supplementary Information

Descriptive details

Copse Agricultural Field

Analysis of NPK and pH by Palintest kit

Analysis of soil organic matter using 'loss on ignition' method

Analysis of oats for quality using a β-glucan extraction method

Hierarchial multinomial logistic regression model: general and final equations.

Tables

Table S5: Statistical comparison of vegetation indices (GRVI, NDVI, GNDVI, SAVI, NDRE and Cl green) between multispectral UAV (MSI-UAV) and the hyperspectral ground reference data (Spectro-1 spectrometer) in (a) June (Stage 3- Flowering), (b) July (Stage 4 – Grain filling), and (c) August (Stage 5 - Ripening)

Table S6: The variables of hierarchical multinomial logistic regression model equation for the CI green VI yield which satisfies all the five yield classes.

Table S7: Model Accuracy and McFadden pseudo R^2.

Table S8: Specificity, sensitivity, and accuracy of model 1, model 2, model 3, model 4, model 5, and model 6 for measured, NDVI and CI green VI yields.

Figure

Figure S1: Pseudo corrected raster NIR band. The NIR spectral band was divided by the maximum DN value of 65535 to convert reflectivity (a) to reflectance by performing to normalise the data and obtain values between 0 to 1.

Hierarchical multinomial logistic regression model

Copse Agricultural Field

The field was subjected to various treatments from seed sowing to harvest. Within a week of the sowing date (early October), a specific herbicide (containing flufenacet and diflufenican) was applied to the soil prior to crop and weed emergence (Herold®, Adama Agricultural Solutions UK Ltd, Reading) alongside Grounded AD (Agrovista UK Ltd., Nottingham, UK), a paraffin oil-based concentrate designed to minimise drift of ground applied pesticides. In spring 2021 (late April), further herbicides were applied simultaneously to control broadleaved weeds i.e. Duplosan KV (containing mecoprop-P), Paramount Max (containing florasulam and tribenuron-methyl) (both from NuFarm UK Ltd., Bradford, Yorkshire) and Gal-Gone (containing fluroxypyr) (Belchim Crop Protection, St Neots, Cambridgeshire). In mid-May the fungicide Protefin (containing prothioconazole and tebuconazole) (Clayton Plant Protection, Dublin, Ireland) was applied. In mid-June, additional fungicides were applied as Aderya® XE (containing fluxapyroxad and mefentrifluconazole) (BASF Agricultural Solutions UK, Littlehampton, Sussex) and Tebucur 250 (containing tebuconazole) (Belchim Crop Protection, St Neots, Cambridgeshire) alongside the growth regulator Canopy (containing mepiquat chloride and prohexadione calcium) (BASF Agricultural Solutions UK, Littlehampton, Sussex). Finally, 21 days prior to harvesting of the crop a pre-harvest desiccate was applied (Roundup Vista plus, Bayer Crop Science, Cambridge, UK). The addition of a desiccate assists in producing an evenly ripe crop (and eliminates perennial weeds) to aid harvesting.

To promote growth of winter oats, fertiliser, NPK(S), 20-8-12($7SO_3$) (CF Fertilisers UK Ltd., Ince, Cheshire) with the composition: total nitrogen (N, 20%), composed of nitric nitrogen (9.2%) and ammoniacal nitrogen (10.8%); total phosphorus (P, 8%) composed of phosphorus pentoxide (P_2O_5); total potassium (K, 12%), composed of potassium oxide (K_2O); and, sulfur trioxide (SO_3, 7%) (CF Fertilisers UK Ltd., Billingham, UK) was applied at the time of sowing (28 September 2020) at a rate of 104 kg / hectare. Additional fertilizer was also applied on 6 March 2021 (rate: 225 kg / hectare), 30 March 2021 (rate: 200 kg / hectare); and 13 April 2021 (rate: 200 kg / hectare). A final addition of muriate of potash (Origin Fertilizers, Royston, Hertfordshire) was applied in July 2021 (rate: 60 kg / hectare). In mid-June, 3 ALO t6p, which contains a precursor of the sugar trehalose-6-phosphate was applied to regulate metabolic and development processes within the crop (Agrovista UK Ltd., Nottingham, UK). In addition,

an adjuvant was applied in mid-May and mid-June, Boost (Dow Agro Sciences, King's Lynn, UK) to assist with ground coverage, soil penetration, crop uptake and rainfastness.

Analysis of NPK and pH by Palintest kit

The Soiltest 10 Bluetooth photometer in Palintest kit (SKW500) was used to analyse the chemical soil health indicators of NPK. At the beginning of analysis, the photometer was calibrated using certified Palintest check standards.

Soil nitrogen analysis: Firstly, to a sample container filled with 50 ml of deionised water, 2.5 ml of 1 M ammonium chloride powder (Extract N) was added and shaken to dissolve the contents. Afterwards, 2 ml soil sample was added to the container and shaken for 1 minute to extract the nitrate and nitrate from soil samples. Then, one level spoonful of Nitratest™ powder was added to the container and shaken for 1 minute to reduce nitrate to nitrate. Afterwards, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. Then, 10 ml filtrate was then transferred to a clean cuvette and used to blank the Soiltest 10 photometer. After blanking the photometer, one Nitricol tablet was added to the same cuvette and crushed to dissolve, and the cuvette was left at room temperature for 10 minutes to allow full colour development. Nitricol helps to produce a range of colours from pale pink to deep violet depending on the nitrogen concentration in soil sample. After 10 minutes the cuvette was inserted into the photometer and the result was recorded as nitrate in mg/l.

Soil phosphorus analysis: Firstly, to a sample container filled with 50 ml of deionised water, 5 tablets of 0.5 M sodium bicarbonate (Extract P) was added and shaken gently to dissolve the contents. Afterwards, 2 ml soil sample was added to the container and shaken for 1 minute to extract the phosphorus from soil samples. Then, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. Then, 2 ml filtrate was then transferred to a clean cuvette and filled until 10 ml using deionised water and one acidifying S tablet was crushed and mixed gently to dissolve the contents. This solution was used to blank the photometer and afterwards one Phosphate P tablet was crushed and mixed to

dissolve. Then the cuvette was left at room temperature for 10 minutes to allow full colour development. Phosphate P tablet helps to form a bright blue complex solution depending on the phosphorus concentration in soil sample. After 10 minutes the cuvette was inserted into the photometer and the result was recorded as phosphorus (mg/l).

Soil Potassium analysis: Firstly, to a sample container filled with 50 ml of deionised water, 2.5 ml of 0.1 M magnesium acetate (Extract K) was added and shaken gently to dissolve the contents. Afterwards, 2 ml soil sample was added to the container and shaken for 1 minute to extract the potassium from soil samples. Then, a filter paper was folded into quarters and inserted into a funnel and placed in a clean container where the extract solution was poured to the funnel to collect the extraction filtrate. Then, 10 ml filtrate was then transferred to a clean cuvette and used to blank the Soiltest 10 photometer. After blanking the photometer, one potassium K tablet was added to the same cuvette and crushed to dissolve, and the cuvette was left at room temperature for 2 minutes to allow full colour development. The potassium K tablet consists of sodium tetraphenylboron to form an insoluble white complex depending on the potassium concentration in soil sample. After 2 minutes the cuvette was inserted into the photometer and the result was recorded as potassium (mg/l).

Soil pH analysis using the multiparameter pocket sensor: Prior to analysis the pH probe was calibrated using the pH 4, 7 and 10 buffers provided by the Palintest kit. Afterwards, two levels of 2 ml soil sample and 10 ml deionised water was added to a sample container and shaken gently for 1 minute. Then, 10 ml sample was transferred to the multiparameter pocket sensor cap. The pH measurement mode was selected on the multiparameter pocket sensor and then inserted into the cap allowing the reading to stabilise and the soil pH was recorded.

Analysis of soil organic matter using 'loss on ignition' method

Firstly, 5 g of soil sample in a crucible was oven dried overnight at 105°C and weighed afterwards. The samples were then re-heated to 800°C and weighed afterwards. The percentage organic matter was calculated as below.

$$\%\ organic\ matter\ = \left(\frac{Weight\ 105°C - Weight\ 800°C}{Weight\ 105°C} \right) * 100 \qquad \text{(SI1)}$$

Analysis of oats for quality using a β-glucan extraction method

The gathered oat grain was extracted for its β-glucan content (48 samples) using a modified Wrobel at al., 2012 methods. Initially, 0.2 g of weighed oats separated from the husk were crushed using a mortar and pestle to make a fine powder. The β-glucan from the samples were extracted in three stages. Stage 1 was based on an alkali extraction where the oats samples were dissolved in 15 ml of 3% NaOH in a water bath at 75°C for 1 hour. The samples were then centrifuge at 8,000 rpm for 30 minutes. The alkali extraction was performed three times for each sample to obtain the β-glucan containing supernatant. Then, in stage 2, acid neutralization was performed where the supernatant obtained from stage 1 was mixed with 20 ml of 0.1 M Tris-HCl (pH = 7.4) and centrifuged at 8,000 rpm for 30 minutes. Furthermore, 20 ml of 0.01 M Tris-HCl (pH = 7.4) was added and centrifuged at 8,000 rpm for 30 minutes. The pellet obtained after the acid neutralisation in stage 2, consists of the β-glucan extracted as a gel deposit. Then, in stage 3, excess solvents from the β-glucan gel deposit were removed by freeze drying where the samples were left in a freeze dryer overnight.

After extraction of the β-glucan, acid hydrolysis was done to convert it to glucose ready for analysis. This was done, in the final stage 4, as follows: 2 ml of de-ionised water was added to the extracted β-glucan sample and vortexed. Acid hydrolysis was performed by converting β-glucan to glucose monomers by adding 66 µL of the β-glucan sample to 33 µL of trifluoracetic acid followed by heating on a heating block for 2 h at 90 °C. In addition, a control was prepared consisting of, 66 µL of the β-glucan sample and 33 µL of de-ionised water which was heated for 2 h at 90 °C. The control was used to check for contamination of samples. The acid hydrolysed glucose samples were stored in a freezer at -4 °C until further analysis.

Analysis of glucose was performed by using a glucose colorimetric assay kit (Randox Laboratories Ltd, County Antrim, UK). The kit employs a GOD-PAP reagent and a buffer that performs enzymatic oxidation in the presence of glucose oxidase to determine glucose. Under the catalysis of peroxidase, the formed hydrogen peroxide (equation 1) reacts with phenol and 4-aminiphenazone to form a red - violet quinoneimine dye (equation 2) as an indicator. The absorbance of the sample can be measured using a spectrophotometer (Randox ref)

$$\text{Glucose} + O_2 + H_2O_2 \xrightarrow{\text{Glucose oxidase}} \text{gluconic acid} + H_2O_2 \qquad \text{(SI2)}$$

$$2H_2O_2 + \text{4-aminophenazone} + \text{phenol} \xrightarrow{\text{Peroxidase}} \text{quinoneimine} + 4H_2O_2 \qquad \text{(SI3)}$$

A 10 µL of sample and 1000 µL of reagent (GOD-PAP reagent + Buffer) was mixed in a cuvette and incubated for 25 min at room temperature. As a standard, a glucose standard solution from the Randox kit was used, and deionised water was used as a reagent blank. Afterwards, the absorbance of the sample, standard and the reagent blank was measured at 500 nm using a spectrophotometer (Jenway 7300 and 7305 spectrometer).

Six standard glucose curves were prepared over a concentration range from 0.1 mM to 40 mM. The absorbance of the glucose standards was measured after the colorimetric assay by the Randox kit. The glucose concentration was calculated by interpolation of the generated linear regression equation. The free glucose (Mr = 180) determined was adjusted by multiplying by a conversion factor of 0.9, to account for the difference in molecular weight of glucose monomers bound in a polysaccharide of β-glucan (Mr = 162). Finally, the percentage β-glucan in oat grain was determined by dividing the amount of glucose measured in each sample by the fresh weight of oat sample as follows: Initially, the mass of glucose in 66 µL of sample was calculated using the measured glucose concentration.

Number of moles = glucose concentration x volume in 66 µL (SI4)

Mass of glucose in 66 µL (g) = moles x molecular weight of glucose (Mr = 180) (SI5)

$$\%\ Beta-glucan = \frac{Mass\ of\ glucose\ in\ 66\ \mu L\ (g) * DF * CF}{Fresh\ weight\ of\ oats\ sample\ (g)} * 100$$

(SI6)

$$\text{Dilution factor (DF)} = \frac{2000\ \mu L}{66\ \mu L} = 30 \quad \text{(SI7)}$$

$$\text{Conversion factor (CF)} = \frac{162\ (Beta-glucan\ polysaccaride)}{180\ (Glucose)} = 0.9$$

(SI8)

Hierarchical Multinomial Logistic Model

The general form of the equations of the hierarchical multinomial logistic model in view of the five yield classes identified are:

For the very low to low yield:

$$\log\left(\frac{p(0 < y \leq 4.97)}{p(y > 4.97)}\right) = \beta_0 + \beta_{01}X_1 + \beta_{02}X_2 \ldots\ldots \beta_{016}X_{16} \quad \text{(SI9)}$$

For the low to medium yield:

$$\log\left(\frac{p(4.97 < y \leq 6.18)}{p(y > 6.18)}\right) = \beta_1 + \beta_{11}X_1 1 + \beta_{12}X_2 \ldots\ldots \beta_{116}X_{16} \quad \text{(SI10)}$$

For the medium to high yield:

$$\log\left(\frac{p(6.18 < y \leq 7.11)}{p(y > 7.11)}\right) = \beta_2 + \beta_{21}X_1 + \beta_{22}X_2 \ldots\ldots \beta_{216}X_{16} \quad \text{(SI11)}$$

And, for the high to very high yield:

$$\log\left(\frac{p(7.11 < y \leq 7.62)}{p(y > 7.62)}\right) = \beta_3 + \beta_{31}X_1 + \beta_{32}X_2 \ldots\ldots \beta_{316}X_{16} \quad \text{(SI12)}$$

The trained hierarchical multinomial logistic regression model equations for the CI green VI estimated yield are:

For the very low to low yield:

$$\log\left(\frac{p(0 < y \leq 4.97)}{p(y > 4.97)}\right) = -4.54 + 0.22X_1 + 1.44X_2 \ldots\ldots\ldots - 0.19X_{16} \quad \text{(SI13)}$$

For the low to medium yield:

$$\log\left(\frac{p(4.97 < y \leq 6.18)}{p(y > 6.18)}\right) = -5.79 - 0.5X_1 + 2.77X_2 \ldots\ldots\ldots + 3.6X_{16} \quad \text{(SI14)}$$

For the medium to high yield:

$$\log\left(\frac{p(6.18 < y \leq 7.11)}{p(y > 7.11)}\right) = -4.38 - 0.63X_1 - 3.15X_2 \ldots\ldots\ldots + 8.32X_{16} \quad \text{(SI15)}$$

For the high to very high yield:

$$\log\left(\frac{p(7.11 < y \leq 7.62)}{p(y > 7.62)}\right) = -146.66 - 37.93X_1 + 12.69X_2 \ldots\ldots\ldots - 60.42X_{16} \quad \text{(SI16)}$$

The coefficients β_0, β_{01}, β_{02} and β_{016} (equation SI9), β_1, β_{11}, β_{12} and β_{116} (equation SI10), β_2, β_{21}, β_{22} and β_{216} (equation SI11) and β_3, β_{31}, β_{32} and β_{316} (equation SI12) were replaced with

their values on the trained CI green VI data model (equations SI13-SI16). All the coefficients of β for the 16 features in the CI green VI combination model, outlined in equations SI13-SI16, for each of the five yield classes are shown in Table S6.

Table S1: Oats data to estimate oats grain yield.

Sample	Recovered oats + husk from sampled 1m² (g)	Average oats + husk (20 g)	± SD (n = 3)	Average oats (20 g)	± SD (n = 3)	Calculated oats 1 m² (g)	Unit conversion for oats 1 m² (t)	Final Oats yield in 1m² (t/ha)
A-1	896.4	20.6272	0.0948	18.9745	0.2653	824.6	0.0008246	8.25
A-2	939.4	20.5003	0.3909	19.4250	0.2287	890.1	0.0008901	8.90
A-3	703.1	20.7281	0.2285	18.8708	0.2900	640.1	0.0006401	6.40
A-4	630.6	20.6553	0.2455	19.3659	0.2080	591.2	0.0005912	5.91
A-5	971.1	20.3844	0.4785	18.7863	0.5741	8954.0	0.0008950	8.95
A-6	564.9	20.7074	0.2475	19.8530	0.1209	541.6	0.0005416	5.42
A-7	527.2	20.6318	0.2494	16.3601	5.2574	418.0	0.0004180	4.18
A-8	201.6	20.7656	0.1505	19.6161	0.2395	190.4	0.0001904	1.90
B-1	800.6	20.1838	0.1120	19.3314	0.1387	766.8	0.0007668	7.67
B-2	802.4	20.4741	0.3359	19.0529	0.0275	746.7	0.0007467	7.47
B-3	843.1	20.4068	0.0508	19.1494	0.1256	791.2	0.0007912	7.91
B-4	781.7	20.6551	0.3221	19.3389	0.3419	731.9	0.0007319	7.32
B-5	810.2	20.4540	0.0976	19.6350	0.2051	777.8	0.0007778	7.78
B-6	788.0	20.6843	0.2568	19.4324	0.3017	740.3	0.0007403	7.40
B-7	664.2	20.5087	0.0773	19.3786	0.0886	627.6	0.0006276	6.28
B-8	346.8	20.3913	0.2377	19.4864	0.3055	331.4	0.0003314	3.31
C-1	1016.3	20.3864	0.2306	18.6343	0.2807	929.0	0.0009290	9.29
C-2	724.8	20.3786	0.3946	19.1671	0.4289	681.7	0.0006817	6.82
C-3	863.5	20.5385	0.1612	19.3801	0.2140	814.8	0.0008148	8.15
C-4	891.3	20.5954	0.2449	18.9850	0.5568	821.6	0.0008216	8.22
C-5	843.1	20.5335	0.2343	19.3484	0.2024	794.4	0.0007944	7.94
C-6	728.4	20.4974	0.3359	19.4950	0.3146	692.8	0.0006928	6.93
C-7	597.6	20.5716	0.1293	19.5135	0.2610	566.9	0.0005669	5.67
C-8	517.9	20.5719	0.0557	19.1451	0.5683	482.0	0.0004820	4.82
D-1	885.0	20.7355	0.2281	19.4518	0.1913	830.2	0.0008302	8.30
D-2	961.3	20.6760	0.3833	19.4563	0.5084	904.6	0.0009046	9.05
D-3	875.4	20.6096	0.2941	19.3750	0.4638	823.0	0.000823	8.23
D-4	929.0	20.7926	0.0870	19.4672	0.0425	869.8	0.0008698	8.70
D-5	955.9	20.4049	0.4865	19.1066	0.5327	895.1	0.0008951	8.95
D-6	284.9	20.3851	0.3311	19.2472	0.7382	269.0	0.0002690	2.69
D-7	339.6	20.4036	0.2748	19.3010	0.6562	321.2	0.0003212	3.21

D-8	370.8	20.5214	0.2770	19.3437	0.2983	349.5	0.0003495	3.50
E-1	658.3	20.2823	0.1332	18.8869	0.1570	613.0	0.0006130	6.13
E-2	996.4	20.4655	0.1798	19.1481	0.0491	932.3	0.0009323	9.32
E-3	748.1	20.5399	0.0204	19.2648	0.0901	701.7	0.0007017	7.02
E-4	1034.3	20.1673	0.1497	18.8590	0.3007	967.2	0.0009672	9.67
E-5	881.5	20.4418	0.3609	18.8695	0.3651	813.7	0.0008137	8.14
E-6	935.6	20.3488	0.2582	19.0134	0.2760	874.2	0.0008742	8.74
E-7	767.5	20.3836	0.1661	18.8579	0.1733	710.1	0.0007101	7.10
E-8	1219.4	20.3724	0.1201	19.3083	0.0960	1155.7	0.0011557	11.56
F-1	622.3	20.7099	0.2865	19.5649	0.2903	587.9	0.0005879	5.88
F-2	765.8	20.4470	0.2105	16.8513	4.0714	631.1	0.0006311	6.31
F-3	718.8	20.5860	0.2804	18.9438	0.1124	661.5	0.0006615	6.61
F-4	970.3	20.5449	0.3873	18.9623	0.3907	895.6	0.0008956	8.96
F-5	991.1	20.6072	0.2988	19.2725	0.4014	926.9	0.0009269	9.27
F-6	965.5	20.5251	0.3233	19.0468	0.2350	896.0	0.0008960	8.96
F-7	1007.8	20.6408	0.3678	19.2869	0.2164	941.7	0.0009417	9.42
F-8	284.9	20.6081	0.3317	19.2244	0.6082	265.8	0.0002658	2.66

Table S2: Agrovista UK chemical soil health indicator data for March 2021.

March-21			
Sample	**P (mg/l)**	**K (mg/l)**	**pH**
Zone 1	38.0	214.0	7.5
Zone 2	31.1	199.9	7.4
Zone 3	21.9	169.3	7.3
Zone 4	17.0	112.0	7.5
Mean	**27.0**	**173.8**	**7.4**
Minimum	**17.0**	**112.0**	**7.3**
Maximum	**38.0**	**214.0**	**7.5**

Table S3: Chemical soil health indicator data, determined using standard methods, for June, July, and August 2021 across the 3 phenological growth stages and oat grain quality (assessed by determination of β-glucan in oats) (a) organised per field layout and (b) organised per field zone.

(a)

Sample	Row	June 2021 (Stage 3 – Flowering)					July 2021 (Stage 4 – Grain filling)					August 2021 (Stage 5 -Ripening)					Oat Harvest
		NO_3- (mg/l)*	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	β-glucan (%)
A1	A	44	21	100	7.73	9.8	38	27	235	7.59	9.8	17	16	190	7.8	9.9	0.36
A2		51	23	105	7.74	11.2	9.7	13	205	7.57	9.3	0.9	10	150	7.79	8.2	0.21
A3		45	20	100	7.62	12.1	7.1	11	150	7.48	10.8	5.3	15	120	7.79	9.7	0.14
A4		9.7	20	110	7.84	12.7	7.5	18	150	7.6	10.5	23	11	200	7.92	10.1	0.74
A5		18	22	100	7.76	11.0	26	20	205	7.34	9.7	7.1	11	245	7.9	11.1	0.42
A6		23	13	80	7.78	11.1	32	9	105	7.51	11.2	16	11	130	7.79	10.1	1.08
A7		29	19	135	7.8	11.5	9.7	12	165	7.66	11.6	39	11	145	7.83	11.9	0.56
A8		31	20	145	7.89	11.0	28	22	170	7.69	10.0	26	20	190	7.84	9.8	0.70
B1	B	13	18	80	7.6	8.3	7.1	17	165	7.56	10.2	3.1	19	95	7.64	8.5	0.50
B2		27	20	100	7.63	10.6	14	17	200	7.56	10.0	8.9	14	125	7.75	8.2	1.18
B3		70	22	105	7.38	10.4	37	12	245	7.43	10.7	1.8	13	120	7.63	9.3	3.39
B4		23	20	100	7.65	11.9	33	11	175	7.69	10.8	4.4	16	185	7.85	11.8	1.34
B5		18	23	105	7.59	12.8	21	20	120	7.7	13.1	26	13	180	7.78	15.8	2.33

B6		8.9	13	90	7.58	9.6	3.5	25	100	7.84	11.9	2.7	16	100	7.84	9.6	0.51
B7		8.9	14	80	7.56	10.5	6.6	10	150	7.86	10.3	7.1	10	135	7.8	9.3	0.09
B8		4.4	11	115	7.18	6.6	4.4	10	175	7.8	10.4	4.4	9	155	7.64	9.6	0.15
C1	C	81	22	165	7.52	9.2	17	13	205	7.31	9.1	3.5	7	160	7.65	7.8	0.75
C2		59	25	150	7.64	9.9	12	16	215	7.43	9.0	2.7	12	210	7.73	5.7	1.41
C3		34	20	100	7.04	9.5	7.5	12	125	7.21	10.0	33	16	130	7.77	16.1	3.94
C4		19	20	90	7.2	10.8	34	10	215	7.28	10.8	7.5	18	135	7.84	12.0	0.96
C5		36	20	100	7.68	12.6	27	17	130	7.36	12.6	21	11	140	7.78	11.6	1.71
C6		13	23	90	7.52	10.8	9.3	9	150	7.57	9.8	18	16	200	7.63	9.5	1.68
C7		12	8	100	7.65	9.2	19	11	120	7.78	10.3	8.9	12	200	7.77	8.4	0.32
C8		34	9	110	7.59	9.9	16	18	150	7.82	10.4	7.1	12	95	7.82	11.0	2.19
D1	D	48	31	100	7.38	9.2	40	16	100	7.51	8.2	4.1	19	145	7.79	7.3	0.22
D2		23	33	132	7.61	10.1	8.4	20	105	7.63	11.7	7.9	19	185	7.77	9.0	0.95
D3		20	33	95	7.46	10.5	19	26	170	7.48	9.5	13	20	190	7.76	10.5	1.26
D4		20	27	90	7.7	11.5	15	11	110	7.49	9.8	39	15	200	7.82	10.3	0.79
D5		22	22	105	7.57	10.3	6.2	12	180	7.42	13.8	14	17	130	7.83	15.0	0.58
D6		46	18	95	7.66	12.8	26	16	155	7.57	13.3	23	16	195	7.85	10.0	2.11
D7		17	13	105	7.75	9.6	48	20	125	7.68	4.9	21	13	130	7.88	9.4	2.93
D8		36	17	105	7.78	10.0	24	20	130	7.78	11.0	14	13	165	7.87	10.3	1.86

E1	E	14	28	95	7.76	10.1	23	19	215	7.81	9.4	21	20	225	7.69	8.2	0.62
E2		22	20	90	7.77	9.5	6.6	20	135	7.73	9.6	4.9	14	125	7.72	7.9	0.37
E3		54	24	115	7.64	9.2	2.7	13	110	7.83	9.0	18	14	155	7.72	9.3	0.71
E4		1.3	22	100	7.67	14.8	13	13	120	7.78	10.6	8.8	9	110	7.72	9.6	0.74
E5		41	38	120	7.48	10.0	1.8	13	105	7.45	10.5	35	13	120	7.72	10.1	0.44
E6		19	19	90	7.51	11.0	8.4	16	100	7.72	10.8	16	16	160	7.86	9.5	0.47
E7		1.8	26	95	7.87	11.6	12	15	165	7.83	10.6	7.1	12	140	7.89	9.1	0.98
E8		16	21	140	7.84	10.8	24	16	140	7.83	10.3	35	12	165	7.87	10.0	0.80
F1	F	48	25	110	7.11	11.3	40	36	240	7.48	10.5	19	23	230	7.73	7.6	1.00
F2		29	25	110	7.52	10.0	15	24	130	7.57	9.9	23	21	255	7.64	9.3	1.22
F3		5.8	29	180	7.42	9.0	39	21	215	7.58	8.8	11	20	125	7.82	9.0	0.43
F4		6.6	35	120	7.79	10.6	36	23	235	7.69	7.6	4.4	27	225	7.83	10.1	0.71
F5		8.9	29	105	7.75	11.3	13	20	135	7.68	17.6	15	25	215	7.8	9.8	0.92
F6		12	21	85	7.78	13.0	4.9	13	165	7.69	11.4	27	26	185	7.86	12.9	1.34
F7		48	17	80	7.89	13.2	9.3	13	108	7.76	11.0	15	14	100	7.88	11.1	0.84
F8		2.7	20	115	7.76	10.1	0.9	18	120	7.86	12.2	0	20	125	7.86	10.0	1.06
Mean		**27**	**22**	**107**	**7.62**	**10.7**	**18**	**17**	**157**	**7.61**	**10.5**	**14**	**15**	**161**	**7.79**	**10.0**	**1.04**
Minimum		**1.3**	**8**	**80**	**7.04**	**6.6**	**0.9**	**9**	**100**	**7.21**	**4.9**	**0**	**7**	**95**	**7.63**	**5.7**	**0.09**
Maximum		**81**	**38**	**180**	**7.89**	**14.8**	**48**	**36**	**245**	**7.86**	**17.6**	**39**	**27**	**255**	**7.92**	**16.1**	**3.94**

(b)

ZONE 1																
	June 2021					July 2021					August 2021					
Sample	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	β-glucan (%)
C2	59	25	150	7.64	9.9	12	16	215	7.43	9.0	2.7	12	210	7.73	5.7	1.41
D1	48	31	100	7.38	9.2	40	16	100	7.51	8.2	4.1	19	145	7.79	7.3	0.22
D2	23	33	132	7.61	10.1	8.4	20	105	7.63	11.7	7.9	19	185	7.77	9.0	0.95
D3	20	33	95	7.46	10.5	19	26	170	7.48	9.5	13	20	190	7.76	10.5	1.26
E1	14	28	95	7.76	10.1	23	19	215	7.81	9.4	21	20	225	7.69	8.2	0.62
E2	22	20	90	7.77	9.5	6.6	20	135	7.73	9.6	4.9	14	125	7.72	7.9	0.37
E3	54	24	115	7.64	9.2	2.7	13	110	7.83	9.0	18	14	155	7.72	9.3	0.71
F1	48	25	110	7.11	11.3	40	36	240	7.48	10.5	19	23	230	7.73	7.6	1.00
F2	29	25	110	7.52	10.0	15	24	130	7.57	9.9	23	21	255	7.64	9.3	1.22
F3	5.8	29	180	7.42	9.0	39	21	215	7.58	8.8	11	20	125	7.82	9.0	0.43
F4	6.6	35	120	7.79	10.6	36	23	235	7.69	7.6	4.4	27	225	7.83	10.1	0.71
Mean	**30**	**28**	**118**	**7.55**	**9.9**	**22**	**21**	**170**	**7.61**	**9.4**	**12**	**19**	**188**	**7.75**	**8.5**	**0.81**
Min	**5.8**	**20**	**90**	**7.11**	**9.2**	**2.7**	**13**	**100**	**7.43**	**7.6**	**2.7**	**12**	**125**	**7.64**	**5.7**	**0.22**
Max	**59**	**35**	**180**	**7.79**	**11.3**	**40**	**36**	**240**	**7.83**	**11.7**	**23**	**27**	**255**	**7.83**	**10.5**	**1.26**
ZONE 2																
	June 2021					July 2021					August 2021					
Sample	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO^{3-} (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	β-glucan (%)
A1	44	21	100	7.73	9.8	38	27	235	7.59	9.8	17	16	190	7.80	9.9	0.36
A2	51	23	105	7.74	11.2	9.7	13	205	7.57	9.3	0.9	10	150	7.79	8.2	0.21
B1	13	18	80	7.6	8.3	7.1	17	165	7.56	10.2	3.1	19	95	7.64	8.5	0.50
B2	27	20	100	7.63	10.6	14	17	200	7.56	10.0	8.9	14	125	7.75	8.2	1.18
B3	70	22	105	7.38	10.4	37	12	245	7.43	10.7	1.8	13	120	7.63	9.3	3.39
B4	23	20	100	7.65	11.9	33	11	175	7.69	10.8	4.4	16	185	7.85	11.8	1.34

C1	81	22	165	7.52	9.2	17	13	205	7.31	9.1	3.5	7	160	7.65	7.8	0.75
C3	34	20	100	7.04	9.5	7.5	12	125	7.21	10.0	33	16	130	7.77	16.1	3.94
C4	19	20	90	7.2	10.8	34	10	215	7.28	10.8	7.5	18	135	7.84	12.0	0.96
D4	20	27	90	7.7	11.5	15	11	110	7.49	9.8	39	15	200	7.82	10.3	0.79
D5	22	22	105	7.57	10.3	6.2	12	180	7.42	13.8	14	17	130	7.83	15.0	0.58
E4	1.3	22	100	7.67	14.8	13	13	120	7.78	10.6	8.8	9	110	7.72	9.6	0.74
E5	41	38	120	7.48	10.0	1.8	13	105	7.45	10.5	35	13	120	7.72	10.1	0.44
F5	8.9	29	105	7.75	11.3	13	20	135	7.68	17.6	15	25	215	7.80	9.8	0.92
Mean	**32**	**23**	**105**	**7.55**	**10.7**	**18**	**14**	**173**	**7.50**	**10.9**	**14**	**15**	**148**	**7.76**	**10.5**	**1.15**
Min	**1.3**	**18**	**80**	**7.04**	**8.3**	**1.8**	**10**	**105**	**7.21**	**9.1**	**0.9**	**7**	**95**	**7.63**	**7.8**	**0.21**
Max	**81**	**38**	**165**	**7.75**	**14.8**	**38**	**27**	**245**	**7.78**	**17.6**	**39**	**25**	**215**	**7.85**	**16.1**	**3.94**

ZONE 3

	June 2021					**July 2021**					**August 2021**					
Sample	**NO_3^- (mg/l)**	**P (mg/l)**	**K (mg/l)**	**pH**	**SOM (%)**	**NO_3^- (mg/l)**	**P (mg/l)**	**K (mg/l)**	**pH**	**SOM (%)**	**NO_3^- (mg/l)**	**P (mg/l)**	**K (mg/l)**	**pH**	**SOM (%)**	**β-glucan (%)**
A3	45	20	100	7.62	12.1	7.1	11	150	7.48	10.8	5.3	15	120	7.79	9.7	0.14
A4	9.7	20	110	7.84	12.7	7.5	18	150	7.60	10.5	23	11	200	7.92	10.1	0.74
A5	18	22	100	7.76	11.0	26	20	205	7.34	9.7	7.1	11	245	7.90	11.1	0.42
B5	18	23	105	7.59	12.8	21	20	120	7.70	13.1	26	13	180	7.78	15.8	2.33
B6	8.9	13	90	7.58	9.6	3.5	25	100	7.84	11.9	2.7	16	100	7.84	9.6	0.51
C5	36	20	100	7.68	12.6	27	17	130	7.36	12.6	21	11	140	7.78	11.6	1.71
C6	13	23	90	7.52	10.8	9.3	9	150	7.57	9.8	18	16	200	7.63	9.5	1.68
D6	46	18	95	7.66	12.8	26	16	155	7.57	13.3	23	16	195	7.85	10.0	2.11
E6	19	19	90	7.51	11.0	8.4	16	100	7.72	10.8	16	16	160	7.86	9.5	0.47
F6	12	21	85	7.78	13.0	4.9	13	165	7.69	11.4	27	26	185	7.86	12.9	1.34
mean	23	20	97	7.65	11.8	14	17	142	7.59	11.4	17	15	173	7.82	11.0	1.15
min	8.9	13	85	7.51	9.6	3.5	9	100	7.34	9.7	2.7	11	100	7.63	9.5	0.14
max	46	23	110	7.84	13.0	27	25	205	7.84	13.3	27	26	245	7.92	15.8	2.33

Zone 4																
	June 2021					July 2021					August 2021					
Sample	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	NO_3^- (mg/l)	P (mg/l)	K (mg/l)	pH	SOM (%)	β-glucan (%)
A6	23	13	80	7.78	11.1	32	9	105	7.51	11.2	16	11	130	7.79	10.1	1.08
A7	29	19	135	7.80	11.5	9.7	12	165	7.66	11.6	39	11	145	7.83	11.9	0.56
A8	31	20	145	7.89	11.0	28	22	170	7.69	10.0	26	20	190	7.84	9.8	0.70
B7	8.9	14	80	7.56	10.5	6.6	10	150	7.86	10.3	7.1	10	135	7.80	9.3	0.09
B8	4.4	11	115	7.18	6.6	4.4	10	175	7.80	10.4	4.4	9	155	7.64	9.6	0.15
C7	12	8	100	7.65	9.2	19	11	120	7.78	10.3	8.9	12	200	7.77	8.4	0.32
C8	34	9	110	7.59	9.9	16	18	150	7.82	10.4	7.1	12	95	7.82	11.0	2.19
D7	17	13	105	7.75	9.6	48	20	125	7.68	4.9	21	13	130	7.88	9.4	2.93
D8	36	17	105	7.78	10.0	24	20	130	7.78	11.0	14	13	165	7.87	10.3	1.86
E7	1.8	26	95	7.87	11.6	12	15	165	7.83	10.6	7.1	12	140	7.89	9.1	0.98
E8	16	21	140	7.84	10.8	24	16	140	7.83	10.3	35	12	165	7.87	10.0	0.80
F7	48	17	80	7.89	13.2	9.3	13	108	7.76	11.0	15	14	100	7.88	11.1	0.84
F8	2.7	20	115	7.76	10.1	0.9	18	120	7.86	12.2	0	20	125	7.86	10.0	1.06
mean	**20**	**16**	**108**	**7.72**	**10.4**	**18**	**15**	**140**	**7.76**	**10.3**	**15**	**13**	**144**	**7.83**	**10.0**	**1.04**
min	**1.8**	**8**	**80**	**7.18**	**6.6**	**0.9**	**9**	**105**	**7.51**	**4.9**	**0**	**9**	**95**	**7.64**	**8.4**	**0.09**
max	**48**	**26**	**145**	**7.89**	**13.2**	**48**	**22**	**175**	**7.86**	**12.2**	**39**	**20**	**200**	**7.89**	**11.9**	**2.93**

Table S4: Investigation of β-glucan analysis (a) Extraction efficiency of β-glucan in different extraction stages, (b) Conversion efficiency of β-glucan to glucose by acid hydrolysis, and (c) Analytical performance parameters for analysing glucose in oat beta-glucan by colorimetric assay.

(a)

β-glucan extraction in different stages	**Extraction efficiency (%) (N = 6), mean ± SD (individual values)**
Base level - no spike (mM)	-
Stage 1 - Before alkali extraction	84.4* (85.2, 83.5)
Stage 2 - Before acid neutralization	90 ± 5 (97.4, 85.1, 90.7, 90.2, 92.5, 83.9)
Stage 3 - Before freeze drying	95 ± 6 (97.3, 97.9, 89.4, 88, 105.2, 92.4)
Stage 4 - Before acid hydrolysis	105 ± 12 (114.3, 113, 96.6, 89.7, 98.5, 119.7)

*N = 2

Note: The extraction efficiency was determined by spiking 15 mM of glucose to the oat samples prior to the four stages of the extraction method. This provides information on the amount of β-glucan extracted from oats at each stage, indicating the efficiency of the extraction method. Stage 1 - Alkali extraction by 3% NaOH, 75 °C (1 h x 3) to obtain supernatant fraction consisting of β-glucan. Stage 2 - Acid neutralisation by 0.1 M and 0.01 M Tris-HCl to obtain gel deposit of β-glucan. Stage 3 – Freeze drying to remove excess solvents in the gel deposit of β-glucan. Stage 4 - acid hydrolysis converts β-glucan to glucose monomers.

(b)

β-glucan to glucose by acid hydrolysis	**Conversion efficiency (%) n = 6, mean ± SD (individual values)**
Stage 4 - Acid hydrolysis	102 ± 8 (114.4, 98.6, 103.6, 95.7, 108.2, 93.0)

Note: The conversion efficiency of beta-glucan was determined by spiking 15 mM of β-glucan to the samples prior to acid hydrolysis. This allows to determine if the β-glucan linkages in the oats sample are efficiently broken to glucose monomers by acid hydrolysis method.

(c)

Analyte	N	Linear regression (Y = mx + c)	Correlation coefficient (R^2)	Limit of detection (LOD) mg/L	Limit of quantification (LOQ) mg/L	Precision (RSDs %) low concentration	Precision (RSDs %) high concentration
Glucose	6	Y = 0.05973x - 0.00645	0.9992	33.6	101.9	9.5	1.9
β-glucan	6			30.3	91.7		

Note: The free glucose (Mr = 180) determined was adjusted by multiplying by a conversion factor of 0.9, to account for the difference in molecular weight of glucose monomers bound in a polysaccharide of β-glucan (Mr = 162).

Table S5: Statistical comparison of vegetation indices (GRVI, NDVI, GNDVI, SAVI, NDRE and Cl green) between multispectral UAV (MSI-UAV) and the hyperspectral ground reference data (Spectro-1 spectrometer) in (a) June (Stage 3- Flowering), (b) July (Stage 4 – Grain filling), and (c) August (Stage 5 - Ripening)

(a)

Sample	Spectro-1 spectrometer (GRVI)	MSI-UAV (GRVI)	P-value	MSI-UAV (NDVI)	P-value	MSI-UAV (GNDVI)	P-value	MSI-UAV (SAVI)	P-value	MSI-UAV (NDRE)	P-value	MSI-UAV (Cl green)	P-value
A6	0.5262	0.3300	0.0001	0.6295	0.0000	0.5983	0.0137	0.6544	0.0001	0.4926	0.0473	2.7459	0.0000
A4	0.4602	0.3512	0.0015	0.6673	0.0002	0.4855	0.1196	0.6545	0.0002	0.4457	0.2429	1.8112	0.0000
A2	0.4567	0.3561	0.0074	0.7385	0.0000	0.4674	0.3098	0.5077	0.0101	0.4706	0.2171	3.5470	0.0000
B6	0.5210	0.3957	0.0176	0.6296	0.0208	0.4821	0.1866	0.4984	0.2845	0.4303	0.0288	1.8456	0.0000
C6	0.4501	0.2873	0.0076	0.6668	0.0041	0.4803	0.2314	0.5644	0.0215	0.4791	0.2351	2.4125	0.0000
C3	0.5235	0.2486	0.0009	0.7139	0.0078	0.4901	0.2358	0.5556	0.2372	0.4081	0.0288	3.0511	0.0000
C1	0.4497	0.3548	0.0313	0.8138	0.0022	0.6409	0.0033	0.5743	0.0145	0.4149	0.1989	4.6234	0.0000
D5	0.5142	0.3300	0.0000	0.7257	0.0000	0.6511	0.0003	0.5872	0.0036	0.4196	0.0076	2.1216	0.0000
D3	0.3311	0.4509	0.0984	0.6722	0.0492	0.5448	0.0773	0.5358	0.0816	0.3941	0.2214	2.5926	0.0000

(b)

Sample	Spectro-1 spectrometer (GRVI)	MSI-UAV (GRVI)	P-value	MSI-UAV (NDVI)	P-value	MSI-UAV (GNDVI)	P-value	MSI-UAV (SAVI)	P-value	MSI-UAV (NDRE)	P-value	MSI-UAV (Cl green)	P-value
A6	0.1637	0.0021	0.1364	0.5698	0.0574	0.3822	0.1047	0.3872	0.1014	0.3518	0.1185	1.5022	0.0191
A4	0.2665	0.0075	0.0000	0.4209	0.0010	0.3694	0.0030	0.3367	0.0183	0.3646	0.0095	0.9730	0.0000
A2	0.2070	-0.0721	0.0110	0.3693	0.0306	0.2553	0.1858	0.3027	0.0770	0.4029	0.0212	0.6909	0.0001
B6	0.3189	-0.0164	0.0228	0.5469	0.0453	0.4431	0.1183	0.4378	0.1241	0.4482	0.1120	1.1627	0.0009
C6	0.2621	-0.0615	0.0011	0.3201	0.0891	0.2898	0.2306	0.2065	0.0932	0.3149	0.1016	0.6959	0.0000
C3	0.2893	0.1140	0.0567	0.3870	0.1682	0.2908	0.4916	0.1071	0.0965	0.3690	0.1977	0.4952	0.0889
C1	0.3472	0.0338	0.0417	0.2430	0.1901	0.1457	0.0807	0.1525	0.0850	0.1708	0.0983	0.3541	0.4738
D5	0.1838	-0.0207	0.0300	0.1836	0.1094	0.0868	0.0532	0.0841	0.0523	0.1701	0.0978	0.2741	0.2611
D3	0.2789	-0.0103	0.0012	0.3021	0.2844	0.2346	0.1587	0.2210	0.1173	0.2795	0.4945	0.3299	0.1405

(c)

Sample	Spectro-1 spectrometer (GRVI)	MSI-UAV (GRVI)	P-value	MSI-UAV (NDVI)	P-value	MSI-UAV (GNDVI)	P-value	MSI-UAV (SAVI)	P-value	MSI-UAV (NDRE)	P-value	MSI-UAV (CI green)	P-value
A6	-0.1356	-0.2793	0.0640	0.3902	0.0050	0.2032	0.0115	0.3207	0.0065	0.1613	0.0148	0.7405	0.0018
A4	-0.1707	-0.1890	0.3272	0.3299	0.0024	0.1336	0.0063	0.2640	0.0030	0.1502	0.0055	0.5618	0.0011
A2	-0.0714	-0.0796	0.4779	0.2607	0.1062	0.1311	0.1637	0.2696	0.1031	0.0886	0.1976	0.4235	0.0725
B6	-0.1064	-0.3540	0.0012	0.2994	0.0006	0.2635	0.0217	0.2441	0.0236	0.2205	0.0255	0.8685	0.0000
C6	-0.0011	-0.0915	0.1816	0.2620	0.0360	0.2635	0.0352	0.2441	0.0405	0.1695	0.0739	0.6388	0.0068
C3	-0.1507	-0.0806	0.1296	0.1563	0.0067	0.1775	0.0057	0.1693	0.0058	0.2002	0.0049	0.4812	0.0016
C1	-0.0941	-0.1833	0.1678	0.1516	0.0429	0.0887	0.0703	0.1573	0.0409	0.1092	0.0584	0.3131	0.0168
D5	-0.1944	-0.1575	0.0125	0.1544	0.0000	0.1537	0.0000	0.1599	0.0000	0.1304	0.0000	0.3456	0.0000
D3	-0.1197	-0.1002	0.0125	0.2360	0.0000	0.2175	0.0000	0.1443	0.0000	0.2213	0.0000	0.7521	0.0000

Table S6: The variables of hierarchical multinomial logistic regression model equation for the CI green VI yield which satisfies all the five yield classes.

Equation variable	month	Equation	Equation 15		Equation 16		Equation 17		Equation 18	
		Feature/Variable	B-variable	Coefficient of β	β variable	Coefficient of β	β variable	Coefficient of β	β variable	Coefficient of β
Constant		Constant	β0	-4.54	β1	-5.79	β2	-4.38	β3	-146.66
X_1	June	Nitrate (mg/l)	β01	0.22	β11	-0.50	β21	-0.63	β31	-37.93
X_2		Potassium (mg/l)	β02	1.44	β12	2.77	β22	-3.15	β32	12.69
X_3		Phosphorus (mg/l)	β03	-4.56	β13	-4.40	β23	-0.01	β33	4.30
X_4		SOM (mg/l)	β04	-0.58	β14	1.33	β24	1.84	β34	-44.82
X_5		pH	β05	4.21	β15	-6.54	β25	-9.91	β35	179.75
X_6	July	Nitrate (mg/l)	β06	-0.57	β16	0.67	β26	-1.16	β36	53.20
X_7		Potassium (mg/l)	β07	2.03	β17	-0.54	β27	0.03	β37	-26.29
X_8		Phosphorus (mg/l)	β08	-0.75	β18	0.14	β28	-1.73	β38	-18.33
X_9		SOM (mg/l)	β09	-0.61	β19	3.19	β29	1.55	β39	-31.02
X_{10}		pH	β010	2.71	β110	0.30	β210	-2.18	β310	-272.26
X_{11}	August	Nitrate (mg/l)	β011	0.67	β111	-3.13	β211	1.70	β311	-10.56
X_{12}		Potassium (mg/l)	β012	-1.66	β112	0.62	β212	2.38	312	-27.49
X_{13}		Phosphorus (mg/l)	β013	1.30	β113	-2.04	β213	1.28	β313	16.87
X_{14}		SOM (mg/l)	β014	0.65	β114	-2.14	β214	-8.72	β314	108.45
X_{15}		pH	β015	-6.12	β115	3.24	β215	7.34	β315	130.97
X_{16}		β- glucan (%)	β016	-0.19	β116	3.60	β216	8.32	β316	-60.42

Table S7: Model Accuracy and McFadden pseudo R^2.

Model 1 - Multivariate normal distribution without bias correction	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.81	0.79
NDVI yield	0.6	0.58
CI green VI yield	0.88	0.76
Model 2 - Multivariate normal distribution with bias correction	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.78	0.84
NDVI yield	0.64	0.57
CI green VI yield	0.82	0.8
Model 3 - Multivariate lognormal distribution without bias correction	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.79	0.61
NDVI yield	0.52	0.51
CI green VI yield	0.78	0.7
Model 4 - Multivariate lognormal distribution with bias correction	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.86	0.79
NDVI yield	0.55	0.58
CI green VI yield	0.82	0.76
Model 5 - Multivariate skewed normal distribution with bias correction	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.76	0.86
NDVI yield	0.7	0.63
CI green VI yield	0.94	0.72
Model 6 - Linearly interpolated datapoint model	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.94	0.78
NDVI yield	0.56	0.58
CI green VI yield	1	0.83
Model 7 - Combination model	**Accuracy**	**McFadden pseudo R^2**
Measured yield	0.68	0.74
NDVI yield	0.72	0.61
CI green VI yield	0.84	0.76

Table S8: Specificity, sensitivity, and accuracy of model 1, model 2, model 3, model 4, model 5, and model 6 for measured, NDVI and CI green VI yields.

	Model 1 - Multivariate normal distribution without bias correction: Measure yield			Model 2 - Multivariate normal distribution with bias correction: Measured yield		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.91	0.98	0.97	0.91	0.98	0.97
Class 2 - low yield (4.97 – 6.18 t/ha)	0.84	0.98	0.95	0.84	0.98	0.95
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.81	0.93	0.91	0.81	0.93	0.91
Class 4 - high yield (7.11 – 7.62 t/ha)	0.99	0.95	0.96	0.99	0.95	0.96
Class 5 - very high yield (>7.62t/ha)	0.67	0.95	0.89	0.67	0.95	0.89
	Model 1 - Multivariate normal distribution without bias correction: NDVI yield			**Model 2 - Multivariate normal distribution with bias correction: NDVI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.48	0.97	0.87	0.15	0.98	0.81
Class 2 - low yield (4.97 – 6.18 t/ha)	0.42	0.87	0.78	0.27	0.86	0.74
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.37	0.89	0.78	0.84	0.86	0.86
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.83	0.86	1.00	0.83	0.86
Class 5 - very high yield (>7.62t/ha)	0.66	0.91	0.86	0.66	0.93	0.88
	Model 1 - Multivariate normal distribution without bias correction: CI green VI yield			**Model 2 - Multivariate normal distribution with bias correction: CI green VI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.57	0.98	0.90	0.77	0.97	0.93
Class 2 - low yield (4.97 – 6.18 t/ha)	0.44	0.98	0.87	0.70	0.97	0.91
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.95	0.84	0.86	0.74	0.98	0.93
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.92	0.94	0.99	0.87	0.89
Class 5 - very high yield (>7.62t/ha)	0.82	0.97	0.94	0.78	0.96	0.92

	Model 3 - Multivariate lognormal without bias correction: Measured yield	**Model 4 - Multivariate lognormal with bias correction: Measured yield**

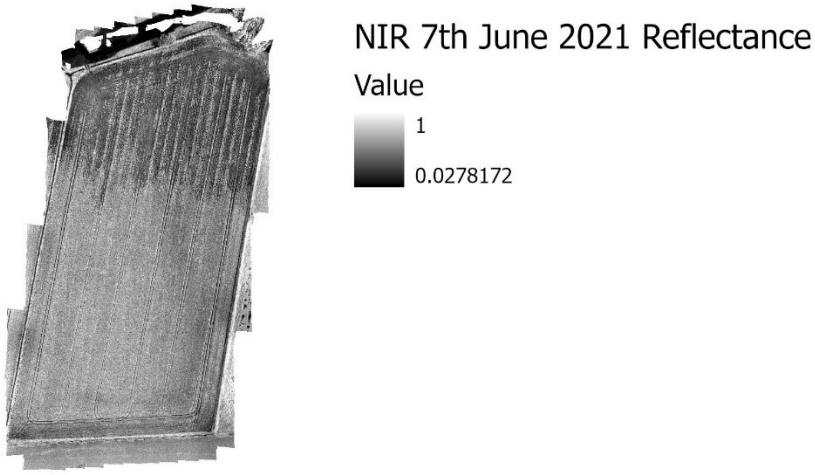
	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy
Class 1 - very low yield (0 – 4.97 t/ha)	0.45	0.99	0.88	0.59	0.98	0.90
Class 2 - low yield (4.97 – 6.18 t/ha)	0.40	0.96	0.86	0.95	0.94	0.94
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.50	0.96	0.87	0.74	0.98	0.93
Class 4 - high yield (7.11 – 7.62 t/ha)	0.96	0.67	0.73	0.96	0.88	0.89
Class 5 - very high yield (>7.62t/ha)	0.72	0.93	0.88	0.72	0.96	0.91
	Model 3 - Multivariate lognormal without bias correction: NDVI yield			**Model 4 - Multivariate lognormal with bias correction: NDVI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.46	0.97	0.87	0.42	0.96	0.85
Class 2 - low yield (4.97 – 6.18 t/ha)	0.28	0.87	0.75	0.34	0.89	0.78
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.24	0.86	0.74	0.59	0.82	0.77
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.84	0.87	1.00	0.93	0.94
Class 5 - very high yield (>7.62t/ha)	0.61	0.84	0.80	0.57	0.87	0.81
	Model 3 - Multivariate lognormal without bias correction: CI green VI yield			**Model 4 - Multivariate lognormal with bias correction: CI green VI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.47	0.97	0.87	0.56	0.96	0.88
Class 2 - low yield (4.97 – 6.18 t/ha)	0.44	0.93	0.83	0.59	0.95	0.88
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.85	0.87	0.86	0.93	0.88	0.89
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.93	0.94	0.99	0.96	0.96
Class 5 - very high yield (>7.62t/ha)	0.72	0.93	0.89	0.75	0.95	0.91

	Model 5 - Skewed multivariate normal distribution with bias correction: Measured yield			Model 6 - Linearly interpolated datapoint: Measured yield		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.78	0.98	0.94	0.62	1.00	0.92
Class 2 - low yield (4.97 – 6.18 t/ha)	0.96	0.99	0.99	0.72	0.97	0.92
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.89	0.96	0.95	0.72	0.97	0.92
Class 4 - high yield (7.11 – 7.62 t/ha)	0.96	0.98	0.98	1.00	0.82	0.85
Class 5 - very high yield (>7.62t/ha)	0.75	0.91	0.88	0.86	0.98	0.95
	Model 5 - Skewed multivariate normal distribution with bias correction: NDVI yield			**Model 6 - Linearly interpolated datapoint: NDVI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.49	0.98	0.88	0.25	0.96	0.81
Class 2 - low yield (4.97 – 6.18 t/ha)	0.27	0.91	0.78	0.50	0.89	0.82
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.79	0.87	0.86	0.50	0.89	0.81
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.82	0.85	1.00	0.79	0.83
Class 5 - very high yield (>7.62t/ha)	0.65	0.96	0.90	0.69	0.94	0.89
	Model 5 - Skewed multivariate normal distribution with bias correction: CI green VI yield			**Model 6 – Linearly interpolated datapoint: CI green VI yield**		
	Specificity	**Sensitivity**	**Accuracy**	**Specificity**	**Sensitivity**	**Accuracy**
Class 1 - very low yield (0 – 4.97 t/ha)	0.57	0.97	0.89	0.75	1.00	0.95
Class 2 - low yield (4.97 – 6.18 t/ha)	0.21	0.97	0.82	0.51	0.95	0.86
Class 3 - medium yield (6.18 – 7.11 t/ha)	0.98	0.81	0.84	0.92	0.89	0.90
Class 4 - high yield (7.11 – 7.62 t/ha)	1.00	0.92	0.94	1.00	0.96	0.97
Class 5 - very high yield (>7.62t/ha)	0.85	0.98	0.95	0.98	1.00	1.00

Figure S1: Pseudo corrected raster NIR band. The NIR spectral band was divided by the maximum DN value of 65535 to convert reflectivity (a) to reflectance by performing to normalise the data and obtain values between 0 to 1

(a)

(b)



Hierarchical multinomial logistic regression model

The steps below are performed using MATLAB and uses the function in the folder named as 'hierarchical multinomial logistic regression model' in the supplementary materials.

1) To run the function below select 'Oats_yield_production.m' script hit run it. Ensure all the other function scripts specified in the folder of 'hierarchical multinomial logistic regression model' and the 'data' file from excel are present in the folder.

2) The data file consist of the measured yield, NDVI and CI green along with the soil nutrients and beta-glucan levels collected across three months of June, July and August 21.

3) The output after running the scripts will show the labels as below;

- r_corr_mat : correlation matrix between all pairs of features
- pV_mat : associated p-value for correlation values in r_corr_mat
- B_cell : B coefficient values across all partitions in Leave-one-out
- cross validation
- dev_cell : fitted logistic regression model deviance across all partitions in Leave-one-out cross validation
- pihat_dp : probability estimates for each of the unique labels in the
- data
- MI : labels assigned by the trained logistic regression model to each test datapoint
- pp : p-value representing model significance against an intercept only model
- pseudo_R2 - McFadden's pseudo R2 explains variability in dependant variable explained by independent predictors over an intercept only model
- accuracy - model prediction accuracy across all labels
- Val_all - Val_all(:,5)- Sensitivity; Val_all(:,6)- Specificity; Val_all(:,7)- Accuracy for each label.

Script Oats_yield_production.m'

```
% Read  in Data from excel file
addpath('./Data/');
addpath('./functions/');


Sh1=xlsread('Data.xlsx',1);
Sh2=xlsread('Data.xlsx',2);
```

```
data=Sh1(:,[3:7,12:16,21:25]); % data consisting of only soil composition values in June, July and August


% Dependant variable
dep_var=Sh2(2:end,1:3);

B_glucan_aug=Sh2(2:end,4);

reco_labels= zeros(size(dep_var,1),size(dep_var,2));

% recoding variables in dep_var - 3 dependent variables
for i=1:size(dep_var,1)
   for j=1:size(dep_var,2)
      if dep_var(i,j)>7.62
         reco_labels(i,j)=5;
      elseif dep_var(i,j)>7.11
         reco_labels(i,j)=4;
      elseif dep_var(i,j)>6.18
         reco_labels(i,j)=3;
      elseif dep_var(i,j)>4.97
          reco_labels(i,j)=2;
      else
         reco_labels(i,j)=1;
      end
   end
end

% data matrix consisting of soil composition over the 3 months
data=[data B_glucan_aug];

%  variable cols and names
col=1:16;
names={'Nitrate (mg/l) - June 21','Potassium (mg/l) - June 21','Phosphorus (mg/l) - June 21','Organic matter (mg/l) - June 21','pH - June 21','Nitrate (mg/l) - July 21','Potassium (mg/l) - July 21','Phosphorus (mg/l) - July 21','Organic matter (mg/l) - July 21','pH - July 21','Nitrate (mg/l) - August 21','Potassium (mg/l) - August 21','Phosphorus (mg/l) - August 21','Organic matter (mg/l) - August 21','pH - August 21','\beta - glucan (%) - August 21'};

%% for 3 datasets and 3 different models

%% Model 1 - using mean and covariance of multivariate normal distribution
%% Model 2 - using robust mean and covariance of multivariate normal distribution
%% Model 3 - using mean and covariance of multivariate lognormal distribution
%% Model 4 - using robust mean and covariance to remove skew in the  multivariate lognormal distribution data
%% Model 5 - using the robust cov +mean for skewed multivariate normal distribution
%% Model 6 - using only linearly interpolated datapoints
```

```
for i=1:6

  tt=strcat('Case ',num2str(i), ' being processed.');

 disp(tt);

% disp('Model 1 completed');
%model development #1
[r_corr_mat1,pV_mat1,B_cell1, dev_cell1, stats_cell1, pihat_dp1, MI1, pp1, pseudo_R21,
accuracy1, Val_all1, selected_fs1, mu_grp1, std_grp1, Data_gen1,labels1] =
model_development(data,reco_labels,1,100,1,1,i);
 disp('Model 1 completed');

%model development #2
[r_corr_mat2,pV_mat2,B_cell2, dev_cell2, stats_cell2, pihat_dp2, MI2, pp2, pseudo_R22,
accuracy2, Val_all2,selected_fs2, mu_grp2, std_grp2, Data_gen2,labels2] =
model_development(data,reco_labels,2,100,1,1,i);
 disp('Model 2 completed');

%model development #3
[r_corr_mat3,pV_mat3,B_cell3, dev_cell3, stats_cell3, pihat_dp3, MI3, pp3, pseudo_R23,
accuracy3, Val_all3,selected_fs3, mu_grp3, std_grp3,Data_gen3,labels3] =
model_development(data,reco_labels,3,100,1,1,i);
 disp('Model 3 completed');

Model{i}.corrMat1=r_corr_mat1;
Model{i}.corrMat2=r_corr_mat2;
Model{i}.corrMat3=r_corr_mat3;

Model{i}.pV_mat1=pV_mat1;
Model{i}.pV_mat2=pV_mat2;
Model{i}.pV_mat3=pV_mat3;

Model{i}.B_cell1=B_cell1;
Model{i}.B_cell2=B_cell2;
Model{i}.B_cell3=B_cell3;

Model{i}.dev_cell1=dev_cell1;
Model{i}.dev_cell2=dev_cell2;
Model{i}.dev_cell3=dev_cell3;

Model{i}.stats_cell1=stats_cell1;
Model{i}.stats_cell2=stats_cell2;
Model{i}.stats_cell3=stats_cell3;

Model{i}.pihat_dp1=pihat_dp1;
Model{i}.pihat_dp2=pihat_dp2;
Model{i}.pihat_dp3=pihat_dp3;

Model{i}.MI1=MI1;
```

```
Model{i}.MI2=MI2;
Model{i}.MI3=MI3;

Model{i}.pp1=pp1;
Model{i}.pp2=pp2;
Model{i}.pp3=pp3;

Model{i}.pseudo_R21=pseudo_R21;
Model{i}.pseudo_R22=pseudo_R22;
Model{i}.pseudo_R23=pseudo_R23;

Model{i}.accuracy1=accuracy1;
Model{i}.accuracy2=accuracy2;
Model{i}.accuracy3=accuracy3;

Model{i}.Val_all1=Val_all1;
Model{i}.Val_all2=Val_all2;
Model{i}.Val_all3=Val_all3;

Model{i}.selected_fs1=selected_fs1;
Model{i}.selected_fs2=selected_fs2;
Model{i}.selected_fs3=selected_fs3;

Model{i}.mu_grp1=mu_grp1;
Model{i}.mu_grp2=mu_grp2;
Model{i}.mu_grp3=mu_grp3;

Model{i}.std_grp1=std_grp1;
Model{i}.std_grp2=std_grp2;
Model{i}.std_grp3=std_grp3;

Model{i}.Data_gen1=Data_gen1;
Model{i}.Data_gen2=Data_gen2;
Model{i}.Data_gen3=Data_gen3;

Model{i}.labels1=labels1;
Model{i}.labels2=labels2;
Model{i}.labels3=labels3;

end


% plot the correlation matrix change for CI green dataset for change in the
% covariance
% Feature set you want to look at in the plot
featureS=[1:5 16];

figure(1)
Data_m1d1=Model{1,1}.Data_gen1(labels1==1,:);
plotmatrix(Data_m1d1(featureS,featureS));
```

```
figure(2)
Data_m2d1=Model{1,2}.Data_gen1(labels1==1,:);
plotmatrix(Data_m2d1(featureS,featureS));

figure(3)
Data_m3d1=Model{1,3}.Data_gen1(labels1==1,:);
plotmatrix(Data_m3d1(featureS,featureS));

figure(4)
Data_m4d1=Model{1,4}.Data_gen1(labels1==1,:);
plotmatrix(Data_m4d1(featureS,featureS));

figure(5)
Data_m5d1=Model{1,5}.Data_gen1(labels1==1,:);
plotmatrix(Data_m5d1(featureS,featureS));

figure(6)
Data_m6d1=Model{1,6}.Data_gen1(labels1==1,:);
plotmatrix(Data_m6d1(featureS,featureS));
```

%% Mismatch in the labels assigned to the 3 dependant variables

% between 1, 2 and 3

```
mismatch_12=(sum(reco_labels(:,1)~=reco_labels(:,2)))/size(reco_labels,1);
mismatch_13=(sum(reco_labels(:,1)~=reco_labels(:,3)))/size(reco_labels,1);
mismatch_23=(sum(reco_labels(:,2)~=reco_labels(:,3)))/size(reco_labels,1);
```

%% outlining the concentration range for different selected_fetures(i.e. nutrients) for each unique yield label

```
unique_labels=unique(reco_labels);


Range1=zeros(length(unique_labels),2,size(data,2));
Range2=zeros(length(unique_labels),2,size(data,2));
Range3=zeros(length(unique_labels),2,size(data,2));

Xaxis={};


for j=1:size(Model,2)
   for i=1:length(unique_labels)
      % model #1 all feature range
      Model{j}.Range1(i,1,:)=min(Model{j}.Data_gen1(labels1==i,:));
      Model{j}.Range1(i,2,:)=max(Model{j}.Data_gen1(labels1==i,:));
      Model{j}.Xaxis{1,i}=Model{j}.Data_gen1(labels1==i,:);
   %end
```

```
    % model #2 all feature range
   % for i=1:length(unique_labels)
       Model{j}.Range2(i,1,:)=min(Model{j}.Data_gen2(labels2==i,:));
       Model{j}.Range2(i,2,:)=max(Model{j}.Data_gen2(labels2==i,:));
       Model{j}.Xaxis{2,i}=Model{j}.Data_gen2(labels2==i,:);

    %end

    % model #3 all feature range
    %for i=1:length(unique_labels)
       Model{j}.Range3(i,1,:)=min(Model{j}.Data_gen3(labels3==i,:));
       Model{j}.Range3(i,2,:)=max(Model{j}.Data_gen3(labels3==i,:));
       Model{j}.Xaxis{3,i}=Model{j}.Data_gen3(labels3==i,:);
    end
end

all_Data1=[];
all_Data2=[];
all_Data3=[];

all_labs1=[];
all_labs2=[];
all_labs3=[];

for i=1:size(Model,2)
   all_Data1=[all_Data1; Model{i}.Data_gen1];
   all_labs1=[all_labs1; Model{i}.labels1];

   all_Data2=[all_Data2; Model{i}.Data_gen2];
   all_labs2=[all_labs2; Model{i}.labels2];

   all_Data3=[all_Data3; Model{i}.Data_gen3];
   all_labs3=[all_labs3; Model{i}.labels3];
end

[r_corr_mat_A1, pV_mat_A1,B_A1, dev_A1,devD_A1,dfeD_A1, stats_A1, pihat_A1,
MI_A1, pp_A1, pseudo_R2_A1, accuracy_A1, Val_A1, selected_features_A1,
train_data_A1, test_data_A1] = full_model_development(all_Data1,all_labs1,1);
[r_corr_mat_A2, pV_mat_A2,B_A2, dev_A2, devD_A2,dfeD_A2, stats_A2, pihat_A2,
MI_A2, pp_A2, pseudo_R2_A2, accuracy_A2, Val_A2, selected_features_A2,
train_data_A2, test_data_A2] = full_model_development(all_Data2,all_labs2,2);
[r_corr_mat_A3, pV_mat_A3,B_A3, dev_A3,devD_A3,dfeD_A3, stats_A3, pihat_A3,
MI_A3, pp_A3, pseudo_R2_A3, accuracy_A3, Val_A3, selected_features_A3,
train_data_A3, test_data_A3] = full_model_development(all_Data3,all_labs3,3);

% intersection of the features marked significant by 3 different data
% models
fs_intersect=intersect(intersect(selected_features_A1,selected_features_A2,'stable'),selected_
features_A3,'stable');
```

```
% Features significant across all 3 models
F_sig=[];
for i=1:length(fs_intersect)
   F_sig=[F_sig; string(names{fs_intersect(i)})];
end

for j=1:(size(Model,2)-1)
   %size(Model,2)
   %figure;
   if j==1 | j==2
   plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 200,names,'mvn',j);
   elseif j==3 | j==4
    plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 200,names,'logn',j);
   else
    plot_cl_pdfs(Model{j}.Xaxis,[5 5 5],1,6, 100,names,'mvsn',j);
   end
end

colorsS=[255,114,118;255,255,102;255, 165, 0;0, 255, 0;34,139,34];
colorsS=colorsS./255;
nn=get(gcf,'Number');
Model_6_title=[{'Linearly interpolated yield plot - Measured Data'},{'Linearly interpolated
yield plot - NDVI Data'},{'Linearly interpolated yield plot - CI green Data'}];
%scatter plot of last model
for i=1:size(Model{6}.Xaxis,1)
   figure(nn+i);
   for j=1:size(Model{6}.Xaxis,2)
      A=Model{6}.Xaxis{i,j};

scatter3(A(:,2),A(:,7),A(:,16),60,colorsS(j,:),'MarkerEdgeColor','k','MarkerFaceColor',colors
S(j,:));
      hold on;
      xticks([min(A(:,2)):30:max(A(:,2))]);
      yticks([min(A(:,7)):40:max(A(:,7))]);
      xlabel(names{2});
      ylabel(names{7});
      zlabel(names{16});
      box on;
      axis square;
      axis tight;
   end
   legend([{'Very low yield'},{'Low yield'},{'Medium yield'},{'High yield'},{'Very high
yield'}])
   %title(Model_6_title{1,i});
   set(gca,"FontSize",16)
   hYLabel = get(gca,'YLabel');
   set(hYLabel,'rotation',-
35,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignment','right','Position',[0
.32673749997845,-0.118479586512629,0])
```

```
   hXLabel = get(gca,'XLabel');

set(hXLabel,'rotation',20,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignm
ent','left','Position',[0.556156794063963,-0.0978,0])

end

%% Original data scatter - CI green
nn=get(gcf,'Number');
F={};
E=unique(reco_labels(:,3));
for i=1:length(E)
    F{i}=data(reco_labels(:,3)==E(i),:);
end


figure(nn+1);
for i=1:size(F,2)
   AA=F{i};

scatter3(AA(:,2),AA(:,7),AA(:,16),60,colorsS(i,:),'MarkerEdgeColor','k','MarkerFaceColor',c
olorsS(i,:));
   hold on;
   xticks([min(AA(:,2)):30:max(AA(:,2))]);
   yticks([min(AA(:,7)):40:max(AA(:,7))]);
   xlabel(names{2});
   ylabel(names{7});
   zlabel(names{16});
   box on;
   axis square;
   axis tight;
end
legend([{'Very low yield'},{'Low yield'},{'Medium yield'},{'High yield'},{'Very high
yield'}])
set(gca,"FontSize",16)
hYLabel = get(gca,'YLabel');
set(hYLabel,'rotation',-
35,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignment','right','Position',[0
.32673749997845,-0.118479586512629,0])
hXLabel = get(gca,'XLabel');
set(hXLabel,'rotation',20,'Units','normalized','VerticalAlignment','baseline','HorizontalAlignm
ent','left','Position',[0.556156794063963,-0.0978,0])


%% Univariate plot showing normal and 'ogk' mu and cov for same distribution
%% for Beta glucan variable class 2 in CI green

AAE=Model{1}.mu_grp3.mu_F2(1,16);
AAEc=sqrt(Model{1}.std_grp3.cov_F2(16,16));
AAE2=Model{2}.mu_grp3.mu_F2(1,16);
```

```
AAE2c=sqrt(Model{2}.std_grp3.cov_F2(16,16));

tt=normpdf([-1:0.01:4],AAE,AAEc);
tt1=normpdf([-1:0.01:4],AAE2,AAE2c);

nn=get(gcf,'Number');

figure(nn+1);
plot([-1:0.01:4],tt,'b-','LineWidth',2);
hold on;
plot([-1:0.01:4],tt1,'r-','LineWidth',2);
box on;
axis tight;
axis square;
xlabel(names{16});
ylabel('Probability density function');
ylim([0 1]);
legend([{'no bias correction'},{'bias correction'}],'Location','northeast');
set(gca,"FontSize",20)

%%% range Values for very high and very low yield from the original 48 data
data_hy=data(reco_labels(:,1)==5 & reco_labels(:,2)==5 & reco_labels(:,3)==5,:);

min_data_hy=min(data_hy);
max_data_hy=max(data_hy);

Range_valvH=num2cell([min_data_hy' max_data_hy']);
Range_finalvHigh=[names' Range_valvH];

data_ly=data(reco_labels(:,1)==1 & reco_labels(:,2)==1 & reco_labels(:,3)==1,:);

min_data_ly=min(data_ly);
max_data_ly=max(data_ly);

Range_valvL=num2cell([min_data_ly' max_data_ly']);
Range_finalvLow=[names' Range_valvL];
```