NPKS Stress in Barley

Michael Nielsen
Aalborg University

Research in remote sensing for surveillance of crops are still confined to simple correlations between NVDI indexes or faulty estimates of Leaf Area Index and the disease or nutrient deficiency the given researcher is trying to predict. The trouble is that these methods only estimate whether something is wrong and if the controlled crops were provoked into having a certain stress it is obvious that the results are fine.
I will try to model the differences in the symptoms of N, P, K and S deficiency. Lene K. Christensen found that the symptoms are distinguishable using spectroscopy at tips and bases. I will use 3D computer vision to extract the data, instead. The symptoms are listed here:

Nutrient	***Typical Symptoms (all of them result in lower growth rate)***
N	Lighter green or pale leaves, Steeper leaves, Red bases, symptoms starting at the bottom
P	Red bases. Starting from bottom leaves, darker green at first and then red all over the leaves. If it is very serious the bottom leaves are yellow. Inhibits number of leaves.
K	Starting from the bottom, white-yellow tips, many brown-yellow spots all over the leaves and in the veins.
S	Like N, but when N is not stressed, the symptoms appear on top leaves first.

Table 1 shows the plant experimental setup. The rows were repeated in 2 locations, ie. distribution A has two rows A1 and A2.

Table 1. Nutrient distribution for experiments A-E,H. kg per hectar

Row	**N**	**P**	**K**	**S**	**Stress**
A1-2	0	0	0	0	NPKS
B1-2	60	0	60	25	P
C1-2	60	10	0	25	K
D1-2	60	10	60	25	None
E1-2	120	20	120	50	Gluttony
H1-2	60	0	0	0	PKS

Each row was a location, so there were 12 locations in total, and 36 plants.

Three cameras were used for the 3D reconstruction in order to extract leaf area, steepness, and height information of bases and tips. The cameras were sensitive to infrared light and they were mounted with filter wheels (Red, Green and narrowband 660 nm, 710 nm, and 770 nm). They were normalised against light-shading using the sum of the 5 bands. This means that if a green point turns red, the red component goes up a little (or not at all if the change also result in an intensity decrease) but the others go down, so the change has to be seen in relation to the others.

The average healthy leaf colour was:

Red	***Green***	***660 nm***	***710 nm***	***770 nm***
11.07	21.32	9.02	27.3	31.29

In order to predict the symptoms it was necessary to look at bottom tip and base and top tip and base as a whole. Thus, tip and base points were ranked by height and labelled as top or bottom leaves (by k-means clustering (k=2)). The 2 topmost and two bottommost tips and bases were paired as repeated measures for one plant.

For this preliminary statistical analysis, steepness, area, and normalised -Red, -Green, and -NIR (770 nm) spectral point measures were investigated.

Area estimated from the 3D model is correlated with the real area $R^2 = 0.92$.

Steepness is a new measure that has not been used before, so the experiments in this paper will show the tangibility of this measure. It is approximated with the height of the topmost tip minus the height of the topmost base divided by the ground plane distance.

Bases and tips were manually segmented. The bases, especially those in the bottom were difficult to segment from the images, as they were usually occluded. The closest point to the bases were chosen instead. This yields an insecurity to the label “topmost” and “bottommost” position of the bases. Furthermore, the topmost tip may not belong to the topmost base, etc.

The data looks like this:

1 obsId Refresh Categories Help

	Name	n	missing	min	max
1	obsId	576	0	1	576
2	plantId	576	0	1	36
3	location	576	0	0	11
4	steep	576	0	-0.87	146.8
5	normsteep	576	0	0	3.55
6	area	576	0	47.2	381.4
7	N	576	0	0	120
8	P	576	0	0	20
9	K	576	0	0	120
10	S	576	0	0	50
11	NS	576	0	0	1
12	PS	576	0	0	1
13	KS	576	0	0	1
14	SS	576	0	0	1
15	Rtiptop	576	0	9.83	15.07
16	Gtiptop	576	0	17.96	27.23
17	s660tiptop	576	0	6.22	13.08
18	s710tiptop	576	0	19.03	31.71
19	s770tiptop	576	0	24.39	35.57
20	Rtipbottom	576	0	9.2	24.14
21	Gtipbottom	576	0	15.27	25.06
22	s660tipbottom	576	0	5.67	19.97
23	s710tipbottom	576	0	16.82	39.09
24	s770tipbottom	576	0	16.9	36.45
25	Rbasetop	576	0	8.71	14.55
26	Gbasetop	576	0	18.17	27.3
27	s660basetop	576	0	4.62	16.4
28	s710basetop	576	0	21.82	37.36
29	s770basetop	576	0	22.39	41.25
30	Rbasebottom	576	0	8.45	20.06
31	Gbasebottom	576	0	17.35	25.41
32	s660basebottom	576	0	4.58	17.55
33	s710basebottom	576	0	20.73	39.1
34	s770basebottom	576	0	19.07	39.79

Experiments

The experiments can be divided into two parts. The first part modelled the effect of NPKS stress on area and steepness. N, P, K, and S were treated as binary variables (1 = stressed when nutrition amount in table 1 is zero). The second part modelled the tendencies of the spectral data.

First a model of the leaf area with the following hierarchy. Location is assumed relevant because area is not a direct indication of nutrition stress. Many other factors from the location affect the area.

Location |
------ Plant

$$\text{area}_{ij} \sim \text{N}(XB, \Omega)$$

$$\text{area}_{ij} = \beta_{0ij}\text{cons} + 2.782(57.211)\text{NS}_j + -115.216(49.546)\text{PS}_j + -110.328(49.546)\text{KS}_j + 159.441(75.683)\text{SS}_j$$

$$\beta_{0ij} = 211.874(28.606) + u_{0j} + e_{0ij}$$

$$\left[u_{0j}\right] \sim \text{N}(0, \Omega_u) : \Omega_u = \left[2978.636(1338.947)\right]$$

$$\left[e_{0ij}\right] \sim \text{N}(0, \Omega_e) : \Omega_e = \left[883.451(255.030)\right]$$

$$-2*loglikelihood(IGLS\ Deviance) = 375.281(36\ \text{of}\ 36\ \text{cases in use})$$

This shown a large impact on the area from random location effect. This could be soil compactness and a strange phenomenon that causes the plant to develop multiple stems when there is a lot of nutrition stress!

Intervals and tests show the significance of the fixed nutrition parameters:

	# 1	# 2	# 3	# 4
fixed : cons	0.000	0.000	0.000	0.000
fixed : ns	1.000	0.000	0.000	0.000
fixed : ps	0.000	1.000	0.000	0.000
fixed : ks	0.000	0.000	1.000	0.000
fixed : ss	0.000	0.000	0.000	1.000
constant(k)	0.000	0.000	0.000	0.000
function result(f)	2.782	-115.216	-110.328	159.441
f-k	2.782	-115.216	-110.328	159.441
chi sq. (f-k)=0. (1df)	0.002	5.408	4.958	4.438
+/- 95% sep.	112.111	97.091	97.091	148.308
+/- 95% joint	176.225	152.616	152.616	233.124

joint chi sq test(4df) = 7.837

random fixed # of functions 4 Calc Help

For 1DF chi sq. should be above 3.8 to be significant ($p<0.05$). Nitrogen is actually the worst predictor and P is the best one. P reduces the number of leaves growing from the stem, not only reduced growth rate like the others. N stress never occurs alone in the data set, so that may cause it to be insignificant (or given by the others).

Significance is the entire model by chi sq test: 7.837 at 4df: $p = 0.09$.

If I reduce the model:

$$\text{area}_{ij} \sim \text{N}(XB, \Omega)$$

$$\text{area}_{ij} = \beta_{0ij}\text{cons} + -115.216(49.551)\text{PS}_j + -110.328(49.551)\text{KS}_j + 160.832(70.076)\text{SS}_j$$

$$\beta_{0ij} = 211.874(28.608) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 2979.286(1339.210) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \text{N}(0,\ \Omega_e) \ : \ \Omega_e = \begin{bmatrix} 883.451(255.031) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = **375.283(36 of 36 cases in use)**

Log likelihood did not become worse by reducing the model. If the location effect is removed, the -2*loglikelihood = 399, so G^2 = 399-375 = 24 which on the chi sq is very significant.

The intercept of the B0 is above the average leaf area of all the plants. The fact that S stress has a positive effect is weird. Well, S stress does not occur alone in the experiment setup (table 1) so its effect may simply counter the additive negative effect from the other stresses (e.g. P + K) that did not really affect the real leaf area that much.

	# 1	# 2	# 3
fixed : cons	0.000	0.000	0.000
fixed : ps	1.000	0.000	0.000
fixed : ks	0.000	1.000	0.000
fixed : ss	0.000	0.000	1.000
constant(k)	0.000	0.000	0.000
function result(f)	-115.216	-110.328	160.832
f-k	-115.216	-110.328	160.832
chi sq. (f-k)=0. (1df)	5.406	4.957	5.268
+/- 95% sep.	97.100	97.100	137.320
+/- 95% joint	138.522	138.522	195.900

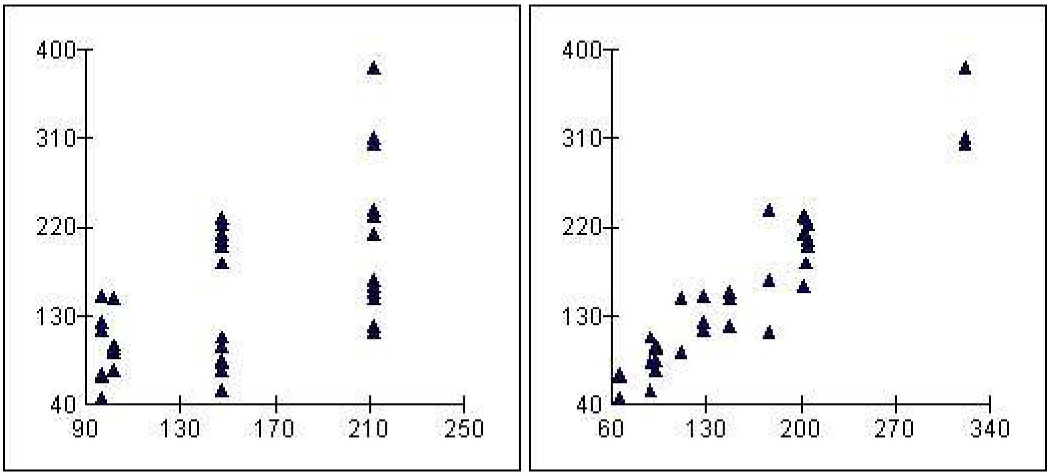
joint chi sq test(3df) = 7.833

random   fixed   # of functions 3   Calc   Help

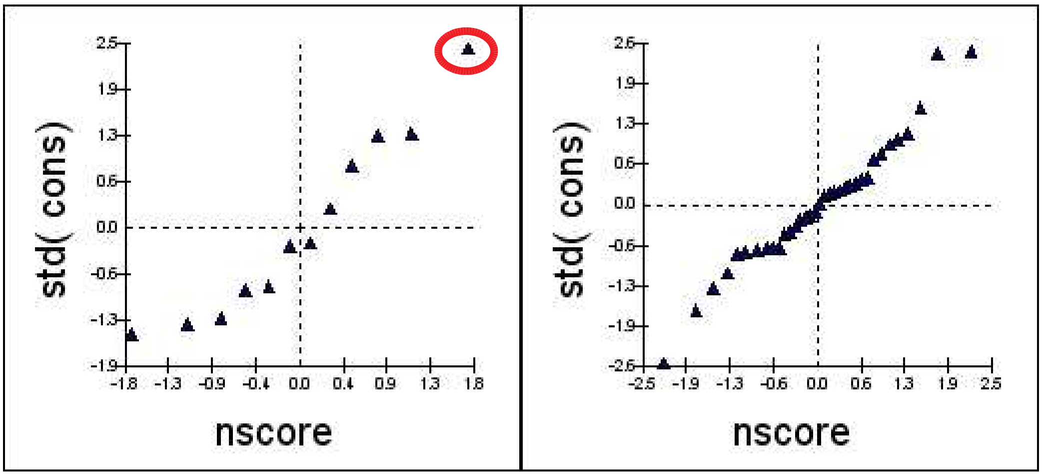
The chi sq. test for the entire model is now p = 0.049

ICC(location) is 2979 / (2979 + 883) = 0.77. So it is probable to get similar results when sampling leaf area from the same location.

Predicted values against correct values without (left) and with (right) known location. Then the location is unknown, the variance increases with increasing area:



Residuals for location level (left) and plant level (right) look ok:



Lets have a look at the steepness with all stress predictors and the hierarchy:

Location |
------ Plant

$$\text{normsteep}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\text{normsteep}_{ij} = \beta_{0ij}\text{cons} + 0.825(0.343)\text{NS}_j + 0.493(0.297)\text{PS}_j + 0.132(0.297)\text{KS}_j + -0.675(0.453)\text{SS}_j$$

$$\beta_{0ij} = 0.208(0.171) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.000(0.000) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 0.352(0.083) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 64.599(36 of 36 cases in use)

It is clear that the location random effect is irrelevant. Some of the nutrient stresses look irrelevant as well. Lets do a chi sq test:

	# 1	# 2	# 3	# 4
fixed : cons	0.000	0.000	0.000	0.000
fixed : ns	1.000	0.000	0.000	0.000
fixed : ps	0.000	1.000	0.000	0.000
fixed : ks	0.000	0.000	1.000	0.000
fixed : ss	0.000	0.000	0.000	1.000
constant(k)	0.000	0.000	0.000	0.000
function result(f)	0.825	0.493	0.132	-0.675
f-k	0.825	0.493	0.132	-0.675
chi sq. (f-k)=0. (1df)	5.797	2.764	0.197	2.218
+/- 95% sep.	0.671	0.581	0.581	0.888
+/- 95% joint	1.055	0.914	0.914	1.396

joint chi sq test(4df) = 9.541

random fixed # of functions 4 Calc Help

chi.sq. For 1DF should be above 3.8 to be significant $p<0.05$ and only Nitrogen is that here. Significance of the entire model 9.541 at 4 df $p = 0.05$.

So the new simple model is:

$$\text{normsteep}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\text{normsteep}_{ij} = \beta_{0i}\text{cons} + 0.660(0.278)\text{NS}_j$$

$$\beta_{0i} = 0.323(0.113) + e_{0ij}$$

$$\left[e_{0ij}\right] \sim \mathrm{N}(0, \Omega_e) : \Omega_e = \left[0.385(0.091)\right]$$

$$-2{*}loglikelihood(IGLS\ Deviance) = 67.808(36 \text{ of } 36 \text{ cases in use})$$

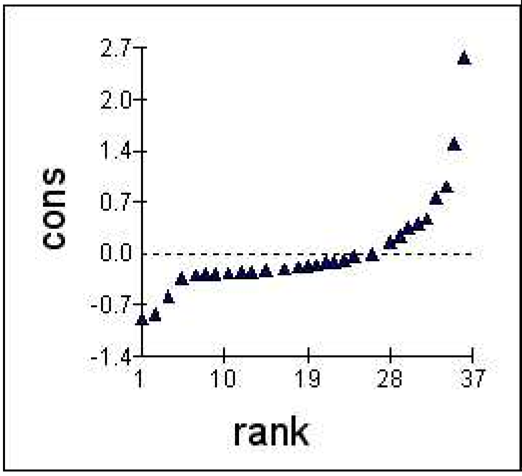
	# 1
fixed : cons	0.000
fixed : ns	1.000
constant(k)	0.000
function result(f)	0.660
f-k	0.660
chi sq. (f-k)=0. (1df)	5.656
+/- 95% sep.	0.544
+/- 95% joint	0.544

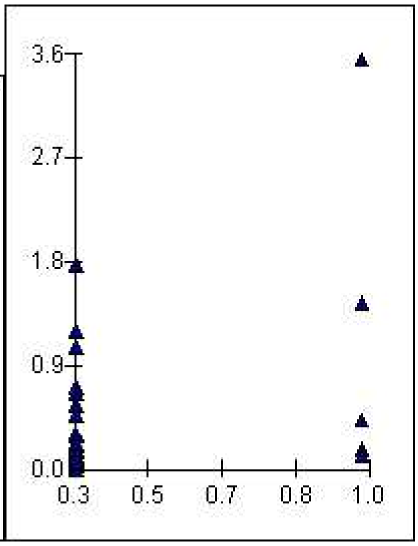
joint chi sq test(1df) = 5.656

random fixed # of functions 1

Significance of the entire model 5.656 at 1 df $p = 0.017$

Residuals (left) and predictions (right) look strange, but this may be because it is predicting a continuous variable with a binary predictor. It can only predict a positive trend in the steepness.





Spectral Data

The 3 chosen spectral bands at each sample spot on the plant will be treated as multivariate responses because they are correlated in the way that they are normalised. Bands Red, Green and 770 are chosen. Band 660 and 710 are covered by the broad Red band. Red should pick up red and white symptoms, and green is needed for reference to the red. Band 770 is Near infrared and tells something about chlorophyll content. If all bands are used, the perfect correlation between them caues the regression to fails for all bands except the middle one (660 nm).
Location is assumed irrelevant for the spectral symptoms as these are direct indicators of stress. The repeated measures for the plants are treated as a new hierarchy. MLWin also treats the multivariate responses as a low level in the hierarchy:

Plant|
------Observation|
--- Response

There will be 4 independent regressions; one for each of tiptop, basetop, tipbottom, and basebottom per observation.
Instead of using S stress directly, I will investigate its interaction with N stress, because its unique symptoms would only be present is there are enough Nitrogen. So the term for S stress is SS*Not NS.

On the following four figures of the regressions, the most significant fixed effects are highlighted. Note that I use a relaxed definition of significant because I do not expect that all repeated measures within a plant has the symptom. I just want to see a trend that resembles the known symptoms (Refer to the symptom list). Most symptoms are most evident in the bottom of the plant. On the figures there are most highlighted effects on the tip bottom and base bottom models and it shows that the significant effects are indeed concentrated around the bottom of the plant.

S stress did not appear in the top, but this may be the recurrent problem that S is not isolated in the experimental design. P stress seems to turn the top base red. It may be stealing the effect from S stress, because S stress never occurs without P stress.

In the bottom of the plants, the chlorophyll (770 nm band) is heavily affected. Like in the area prediction, the S stress seems to counter the additive negative effects of the others.
P stress turns the bottom tip very red, and K stress strongly corrupts the leaf in both ends.
N stress (which only occurs in full stress) strongly effects the chlorophyll 770 nm base at the

bottom.

The covariance matrices on all four figures show negative correlations between colour bands. A better independent normalisation is necessary. Maybe use a blue band as the normalisation factor.

The ICCs between plants and ICCs between observations are almost the same. Does that mean that plant cluster is weakly defined? (Especially when ICC(obs) is larger than ICC(plant), because there is a weaker correlation between observations within the same plant).
The way the tips and bases are paired, we cannot expect to the symptoms on all the repeated measures.

$$\text{resp}_{1ijk} \sim \text{N}(XB,\ \Omega)$$
$$\text{resp}_{2ijk} \sim \text{N}(XB,\ \Omega)$$
$$\text{resp}_{3ijk} \sim \text{N}(XB,\ \Omega)$$
$$\text{resp}_{1ijk} = \beta_{0ijk}\text{cons.Rtiptop}_{ijk} + 1.479(0.463)\text{NS.Rtiptop}_{ijk} + 0.166(0.303)\text{PS.Rtiptop}_{ijk} + \text{-}0.396(0.303)\text{KS.Rtiptop}_{ijk} + 0.582(0.463)\text{SS.NNS.Rtiptop}_{ijk}$$
$$\beta_{0ijk} = 11.489(0.175) + v_{0k} + u_{0ijk}$$
$$\text{resp}_{2ijk} = \beta_{1ijk}\text{cons.Gtiptop}_{ijk} + 1.000(0.957)\text{NS.Gtiptop}_{ijk} + 0.923(0.627)\text{PS.Gtiptop}_{ijk} + \text{-}0.676(0.627)\text{KS.Gtiptop}_{ijk} + 0.615(0.957)\text{SS.NNS.Gtiptop}_{ijk}$$
$$\beta_{1ijk} = 21.778(0.362) + v_{1k} + u_{1ijk}$$
$$\text{resp}_{3ijk} = \beta_{2ijk}\text{cons.s770tiptop}_{ijk} + \text{-}1.395(1.210)\text{NS.s770tiptop}_{ijk} + \text{-}1.315(0.792)\text{PS.s770tiptop}_{ijk} + \text{-}0.028(0.792)\text{KS.s770tiptop}_{ijk} + 0.556(1.210)\text{SS.NNS.s770tiptop}_{ijk}$$
$$\beta_{2ijk} = 29.780(0.457) + v_{2k} + u_{2ijk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \end{bmatrix} \sim \text{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 0.343(0.087) & & \\ \text{-}0.086(0.127) & 1.447(0.370) & \\ \text{-}0.593(0.191) & \text{-}0.378(0.338) & 2.392(0.592) \end{bmatrix}$$

$$\begin{bmatrix} u_{0ijk} \\ u_{1ijk} \\ u_{2ijk} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.394(0.024) & & \\ 0.306(0.040) & 1.992(0.121) & \\ \text{-}0.570(0.044) & \text{-}0.455(0.085) & 1.875(0.114) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 5018.908(1728 of 1728 cases in use)

$$\mathrm{resp}_{1jk} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{resp}_{2jk} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{resp}_{3jk} \sim \mathrm{N}(XB,\ \Omega)$$

$$\mathrm{resp}_{1jk} = \beta_{0jk}\mathrm{cons.Rbasetop}_{ijk} + -0.474(0.470)\mathrm{NS.Rbasetop}_{ijk} + 0.826(0.307)\mathrm{PS.Rbasetop}_{ijk} + 0.221(0.307)\mathrm{KS.Rbasetop}_{ijk} + -0.453(0.470)\mathrm{SS.NNS.Rbasetop}_{ijk}$$

$$\beta_{0jk} = 10.504(0.178) + v_{0k} + u_{0jk}$$

$$\mathrm{resp}_{2jk} = \beta_{1jk}\mathrm{cons.Gbasetop}_{ijk} + -0.190(1.007)\mathrm{NS.Gbasetop}_{ijk} + 0.419(0.659)\mathrm{PS.Gbasetop}_{ijk} + -0.082(0.659)\mathrm{KS.Gbasetop}_{ijk} + 0.151(1.007)\mathrm{SS.NNS.Gbasetop}_{ijk}$$

$$\beta_{1jk} = 21.798(0.381) + v_{1k} + u_{1jk}$$

$$\mathrm{resp}_{3jk} = \beta_{2jk}\mathrm{cons.s770basetop}_{ijk} + -0.413(1.889)\mathrm{NS.s770basetop}_{ijk} + -2.610(1.237)\mathrm{PS.s770basetop}_{ijk} + -0.408(1.237)\mathrm{KS.s770basetop}_{ijk} + 0.614(1.889)\mathrm{SS.NNS.s770basetop}_{ijk}$$

$$\beta_{2jk} = 32.810(0.714) + v_{2k} + u_{2jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 0.336(0.089) & & \\ 0.485(0.162) & 1.611(0.410) & \\ -0.881(0.300) & -1.653(0.622) & 5.790(1.442) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.674(0.041) & & \\ 0.805(0.061) & 2.031(0.124) & \\ -1.263(0.098) & -2.492(0.177) & 5.255(0.320) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 5406.478(1728 of 1728 cases in use)

$$\mathrm{resp}_{1jk} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{resp}_{2jk} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{resp}_{3jk} \sim \mathrm{N}(XB,\ \Omega)$$
$$\mathrm{resp}_{1jk} = \beta_{0jk}\mathrm{cons.Rtipbottom}_{ijk} + -0.576(1.571)\mathrm{NS.Rtipbottom}_{ijk} + 2.109(1.028)\mathrm{PS.Rtipbottom}_{ijk} + 2.551(1.028)\mathrm{KS.Rtipbottom}_{ijk} + -3.499(1.571)\mathrm{SS.NNS.Rtipbottom}_{ijk}$$
$$\beta_{0jk} = 11.496(0.594) + v_{0k} + u_{0jk}$$
$$\mathrm{resp}_{2jk} = \beta_{1jk}\mathrm{cons.Gtipbottom}_{ijk} + -0.107(0.943)\mathrm{NS.Gtipbottom}_{ijk} + -0.643(0.617)\mathrm{PS.Gtipbottom}_{ijk} + -0.596(0.617)\mathrm{KS.Gtipbottom}_{ijk} + 0.941(0.943)\mathrm{SS.NNS.Gtipbottom}_{ijk}$$
$$\beta_{1jk} = 21.616(0.356) + v_{1k} + u_{1jk}$$
$$\mathrm{resp}_{3jk} = \beta_{2jk}\mathrm{cons.s770tipbottom}_{ijk} + -0.442(2.094)\mathrm{NS.s770tipbottom}_{ijk} + -1.908(1.371)\mathrm{PS.s770tipbottom}_{ijk} + -4.274(1.371)\mathrm{KS.s770tipbottom}_{ijk} + 5.278(2.094)\mathrm{SS.NNS.s770tipbottom}_{ijk}$$
$$\beta_{2jk} = 29.801(0.791) + v_{2k} + u_{2jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} 4.051(0.997) & & \\ -0.647(0.438) & 1.369(0.359) & \\ -4.781(1.256) & 0.718(0.575) & 6.998(1.772) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 2.849(0.173) & & \\ -0.598(0.117) & 2.467(0.150) & \\ -3.430(0.256) & -0.649(0.197) & 8.296(0.505) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 6965.098(1728 of 1728 cases in use)

$$\text{resp}_{1jk} \sim \text{N}(XB, \Omega)$$
$$\text{resp}_{2jk} \sim \text{N}(XB, \Omega)$$
$$\text{resp}_{3jk} \sim \text{N}(XB, \Omega)$$
$$\text{resp}_{1jk} = \beta_{0jk}\text{cons.Rbasebottom}_{ijk} + 0.598(1.030)\text{NS.Rbasebottom}_{ijk} + 0.927(0.674)\text{PS.Rbasebottom}_{ijk} + 1.321(0.674)\text{KS.Rbasebottom}_{ijk} + -1.257(1.030)\text{SS.NNS.Rbasebottom}_{ijk}$$
$$\beta_{0jk} = 10.295(0.389) + v_{0k} + u_{0jk}$$
$$\text{resp}_{2jk} = \beta_{1jk}\text{cons.Gbasebottom}_{ijk} + -0.852(1.059)\text{NS.Gbasebottom}_{ijk} + 0.935(0.693)\text{PS.Gbasebottom}_{ijk} + 1.377(0.693)\text{KS.Gbasebottom}_{ijk} + -1.313(1.059)\text{SS.NNS.Gbasebottom}_{ijk}$$
$$\beta_{1jk} = 20.508(0.400) + v_{1k} + u_{1jk}$$
$$\text{resp}_{3jk} = \beta_{2jk}\text{cons.s770basebottom}_{ijk} + -4.214(1.901)\text{NS.s770basebottom}_{ijk} + -1.116(1.245)\text{PS.s770basebottom}_{ijk} + -1.076(1.245)\text{KS.s770basebottom}_{ijk} + 2.108(1.901)\text{SS.NNS.s770basebottom}_{ijk}$$
$$\beta_{2jk} = 31.631(0.719) + v_{2k} + u_{2jk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \end{bmatrix} \sim \text{N}(0, \Omega_v) : \Omega_v = \begin{bmatrix} 1.716(0.428) & & \\ 0.482(0.322) & 1.841(0.453) & \\ -2.242(0.687) & -0.855(0.592) & 5.715(1.461) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \end{bmatrix} \sim \text{N}(0, \Omega_u) : \Omega_u = \begin{bmatrix} 1.630(0.099) & & \\ 0.010(0.062) & 1.284(0.078) & \\ -2.429(0.185) & 0.179(0.135) & 7.697(0.468) \end{bmatrix}$$

*-2*loglikelihood(IGLS Deviance)* = 6439.354(1728 of 1728 cases in use)

I tested the difference between using SS.NNS and just the SS, and there was no difference. The conclusion must be that the data cannot separate N and S. There was not enough combinations of nutrient distributions.

Preliminary Conclusions

Steepness needs to be better defined and compared to ground truth. Leaf area is more dependent on location than nutrition K and P and some arbitrary relation of N and S (that the data was unable to separate) was best predictors. Although additional information about the soil location is necessary for a prediction.

Colours were difficult to analyse. It seemed to be a big mess, because they were based on leaves with symptoms and leaves without. The data should be picked intelligently using pattern recognition, so that leaves with interesting spots (defined as spots that are discoloured) are chosen.

Overall Conclusion

The data was not the best suited data to use for these analyses. There were too mush fuzziness in the experimental design and variable definition and no definite ground truth. The clusters (higher level hierarchies) were not well represented. Binary predictors caused prediction values to be almost binary (finite number of possible outcomes), which made residuals very large and non-normal.