

# Optimek 2015 - Harvest Yield Analysis

Morten S. Laursen

March 12, 2016

## Table of Contents

Dataset description .....	1
File paths used for processing.....	1
Tables contained in database .....	1
Reading required tables.....	2
Modeling.....	2
Definition of outliers .....	2
Full model from last year (2014) on 2015 data .....	2
Final model from last year (2014) on 2015 data (Invalid reduction) .....	4
Model spatial dependency based on the tram lines.....	4
Adding zones to the field based on aerial photos based on previous years.....	8
Doing the same for the other year 2014 using Krak.....	10
Making a spatial model, which uses the zones .....	11
Other aerial photos.....	21
Making a full model of the yield, utilizing the spatial model.....	23
Spatial location of parcel combinations .....	32
Adding a interaction between the spatial correction and the rowdistance .....	33

## Dataset description

### File paths used for processing

```
rootFolder <- "/home/morten/Optimek" #adjust for your setup
databaseFile <- "OptiMek Fullscale Field Trial - Foulum 2015.accdb"
bingDsn <- "bing_zoner"
krakDsn <- "krak_zoner"
```

### Tables contained in database

```
## [1] "RoboWeedSupport Status"
## [2] "tbl_AllSubParcels_xy_SubParcelID_ParcelID_GlobalID"
## [3] "tbl_ArcMap_Optimek2015_SuperParcels_polygons"
```

```

## [4] "tbl_ArcMapTramlines_Optimek2015_SuperParcels"
## [5] "tbl_DroneVisualInspection_1MaybeBad_2Bad"
## [6] "tbl_GlobalID_CmBetweenPlants"
## [7] "tbl_NTFS_Faktor1_CropDensity"
## [8] "tbl_NTFS_Faktor2_RowDist"
## [9] "tbl_NTFS_Faktor3_SeedPattern"
## [10] "tbl_NTFS_HarvestDataFrom Agrotech"
## [11] "tbl_NTFS_NetParcelArea"
## [12] "tbl_NTFS_ParcelNo_RKNo"
## [13] "tbl_PhytonParcelID_CmBetweenPlants"
## [14] "tbl_plantDensities_SubParcelID_ParcelID_GlobalID"
## [15] "tbl_SuperParcelCountID_2_SuperParcelCountID_Wrong_ArcGIS"
## [16] "tbl_SuperParcelCountID_DriverError_2_ArcMap"
## [17] "Que_RoboWeedSupport Summary count"
## [18] "QueUnion_ArcMap_and_ArcMapTramlines"
## [19] "Que_TreatmentComb_NoParcelAndPlantDist_afterArcMapFailiure"
## [20] "que_NTFS"
## [21] "que_NTFS_HarvestTable4Agrotech"
## [22] "que_NTFS_HarvestTableAgroTech_Crosstab"
## [23] "que_NTFS_Harvest_TSperHa_Crosstab"
## [24] "que_NTFS_Harvest_TSperHa"
## [25] "que_NTFS_Harvest_Grouping"

```

## Reading required tables

```

NTFS_Harvest <- dbReadTable(dbConn, "que_NTFS_Harvest_TSperHa")
marked_Outliers <- dbReadTable(dbConn,
"tbl_DroneVisualInspection_1MaybeBad_2Bad")

```

## Modeling

### Definition of outliers

Outliers are selected based on the ortophoto based on manual estimation. The outliers are identified on the lack of homogeneous canopy e.g. caused by missing plants.

```

outliers <- NTFS_Harvest$DB_GlobalID %in% marked_Outliers$GlobalID

```

### Full model from last year (2014) on 2015 data

The model were reduced during the analysis in 2014 e.g. removing interaction effects. No spatial corrections were performed. A very dense EM38 sampling (soil conductivity) used as covariate didn't have any significant effect. See the Analysis report [here](#).

```

filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines"

##
## Call:
## lm(formula = `TS_YieldParcel (ton/ha)` ~ RowDist *
NTFS_F1_ThousandSeedsPerHectar +

```

```

##      SeedPattern, data = NTFS_Harvest[filter, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.5305 -0.9783  0.1407  1.0574  5.3465
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)    7.879e+00  4.065e-01  19.382
## RowDist       -4.507e-03  8.828e-04  -5.105
## NTFS_F1_ThousandSeedsPerHectar  3.565e-03  2.691e-03   1.325
## SeedPatternSync  1.699e-01  1.290e-01   1.318
## RowDist:NTFS_F1_ThousandSeedsPerHectar  4.213e-05  7.049e-06   5.977
##              Pr(>|t|)
## (Intercept)    < 2e-16 ***
## RowDist       4.39e-07 ***
## NTFS_F1_ThousandSeedsPerHectar    0.186
## SeedPatternSync    0.188
## RowDist:NTFS_F1_ThousandSeedsPerHectar 3.82e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.598 on 629 degrees of freedom
## Multiple R-squared:  0.3566, Adjusted R-squared:  0.3525
## F-statistic: 87.17 on 4 and 629 DF,  p-value: < 2.2e-16

```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	7.88	0.41	19.38	0.00
RowDist	0.00	0.00	-5.10	0.00
NTFS_F1_ThousandSeedsPerHectar	0.00	0.00	1.32	0.19
SeedPatternSync	0.17	0.13	1.32	0.19
RowDist:NTFS_F1_ThousandSeedsPerHectar	0.00	0.00	5.98	0.00

From the full model from last year it becomes apparent that we cannot reuse the reduced model from 2014 as we have an interaction effect between row distance and the seed density.

The model indicate the yield increase as the row distance decreases. The interaction between the RowDistance and the seed density indicate the yield increases as the seed density increases but this effect is increasing as the row distance increases (opposite the main effect of the rowdistance). However further analysis is indeed needed since the latter result is highly doubtful due to a clear spatial difference in the crop height and density in form of a very low level in the SE corner of the field increasing moving NW (See model Spatial dependency). Early in the season this was realised and as a consequence the tramline tracks for every 27 m were included a harvest parcels trying to create the basis for modelling the spatial yield baseline level in the field.

## Final model from last year (2014) on 2015 data (Invalid reduction)

The reduced model from 2014 which clearly cannot be used for the 2015 data... so this is just for show.

```
filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines"

##
## Call:
## lm(formula = `TS_YieldParcel (ton/ha)` ~ RowDist +
##   NTFS_F1_ThousandSeedsPerHectar,
##   data = NTFS_Harvest[filter, ])
##
## Residuals:
##   Min       1Q   Median       3Q      Max
## -8.7106 -1.0369  0.1492  1.1246  5.4505
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.2079871  0.2845031  21.820  <2e-16 ***
## RowDist           0.0002421  0.0003919   0.618   0.537
## NTFS_F1_ThousandSeedsPerHectar 0.0182261  0.0011181  16.301  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.641 on 631 degrees of freedom
## Multiple R-squared:  0.3199, Adjusted R-squared:  0.3178
## F-statistic: 148.4 on 2 and 631 DF,  p-value: < 2.2e-16
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	6.21	0.28	21.82	0.00
RowDist	0.00	0.00	0.62	0.54
NTFS_F1_ThousandSeedsPerHectar	0.02	0.00	16.30	0.00

## Model spatial dependency based on the tram lines

As we have observed a high variation based on the location of a parcel within the field we have chosen to investigate if we can model the spatial effect by modeling the yield in the

tram lines based solely on the position in the field.

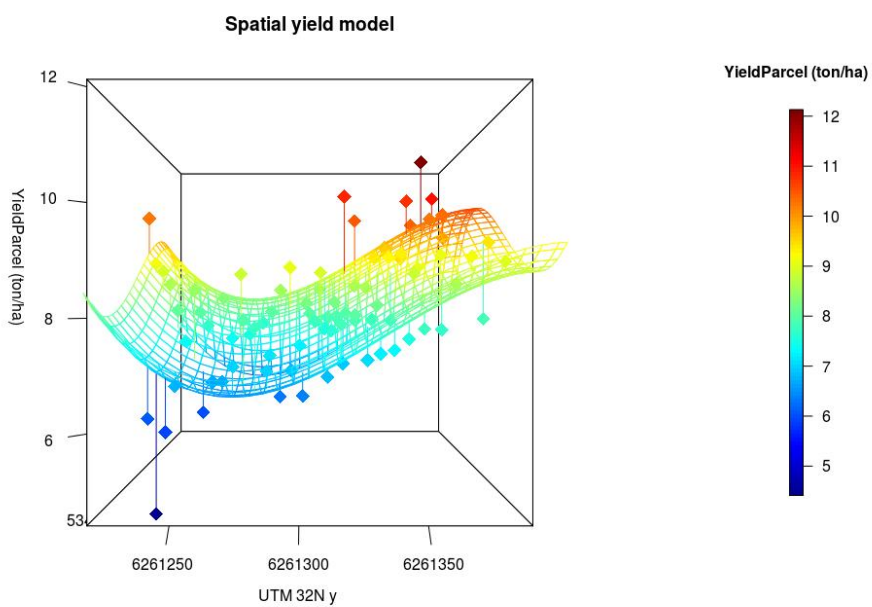
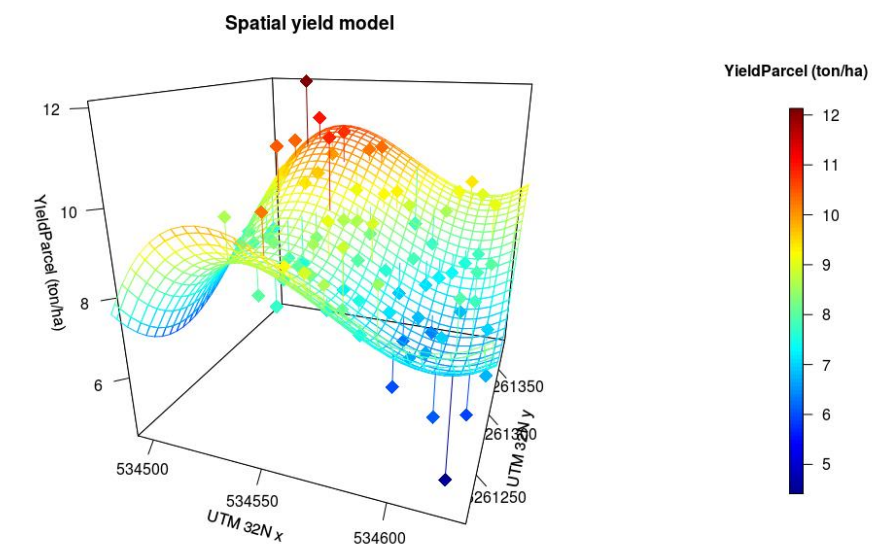


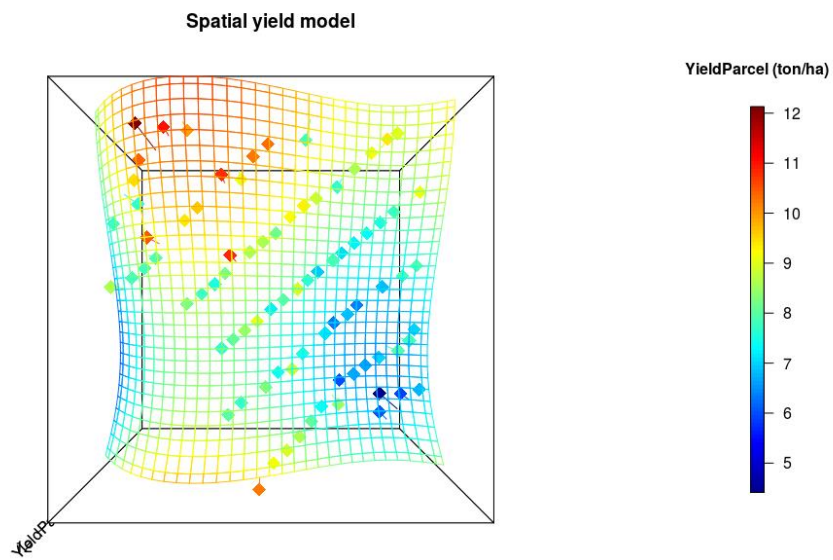
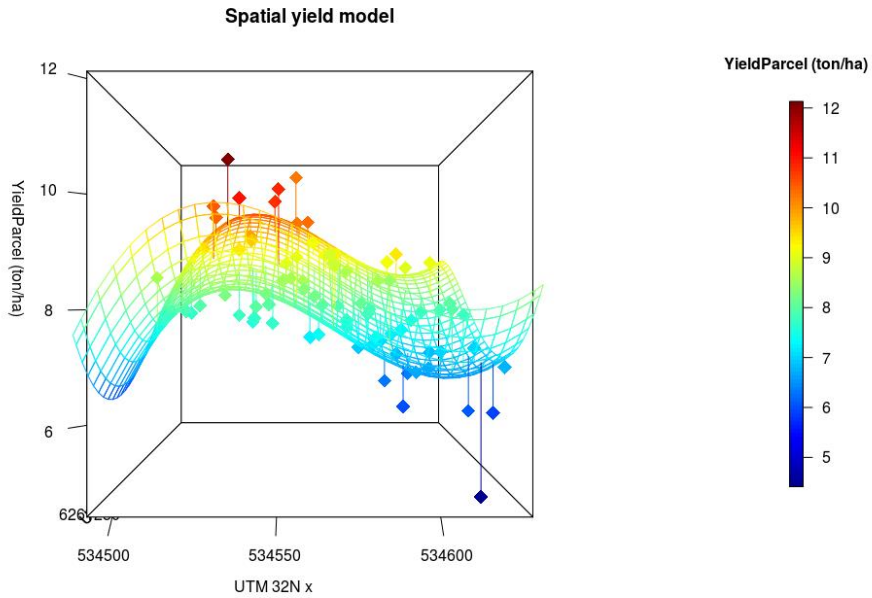
The DM yields in 81 tramline parcels with 500 mm row distance and 200 mm seed distance (10 plants per m<sup>2</sup>) were used trying to fit a 3. Order polynomial in the S → N (y) and E → W (x) direction. The interactions were not used fitting a plane.

```
filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines"

##
## Call:
## lm(formula = `TS_YieldParcel (ton/ha)` ~ poly(x_center, 3) +
##     poly(y_center, 3), data = NTFS_Harvest[filter, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.55457 -0.57781 -0.03065  0.47368  1.90172
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      8.25452    0.09303   88.726 < 2e-16 ***
## poly(x_center, 3)1 -4.63463    0.89107  -5.201 1.66e-06 ***
## poly(x_center, 3)2 -2.57202    0.85439  -3.010 0.003554 **
## poly(x_center, 3)3  3.02566    0.84596   3.577 0.000614 ***
## poly(y_center, 3)1  5.17726    0.87516   5.916 9.19e-08 ***
## poly(y_center, 3)2  3.59309    0.85085   4.223 6.70e-05 ***
## poly(y_center, 3)3 -1.69229    0.86590  -1.954 0.054386 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8425 on 75 degrees of freedom
## Multiple R-squared:  0.6258, Adjusted R-squared:  0.5958
## F-statistic: 20.9 on 6 and 75 DF, p-value: 3.025e-14
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	8.25	0.09	88.73	0.00
poly(x_center, 3)1	-4.63	0.89	-5.20	0.00
poly(x_center, 3)2	-2.57	0.85	-3.01	0.00
poly(x_center, 3)3	3.03	0.85	3.58	0.00
poly(y_center, 3)1	5.18	0.88	5.92	0.00
poly(y_center, 3)2	3.59	0.85	4.22	0.00
poly(y_center, 3)3	-1.69	0.87	-1.95	0.05





Studying the model and the residuals it is clear the spatial model is needed minimizing bias in the statistical model analysing the effect of the row distance, seed pattern, and crop density. Still the model seems not sufficient visually studying the residuals especially in the x direction. This sparklet the idea that previous trials within the field cause bias in the spatial yield level the latter model only based on the x - y direction were not able to catch.

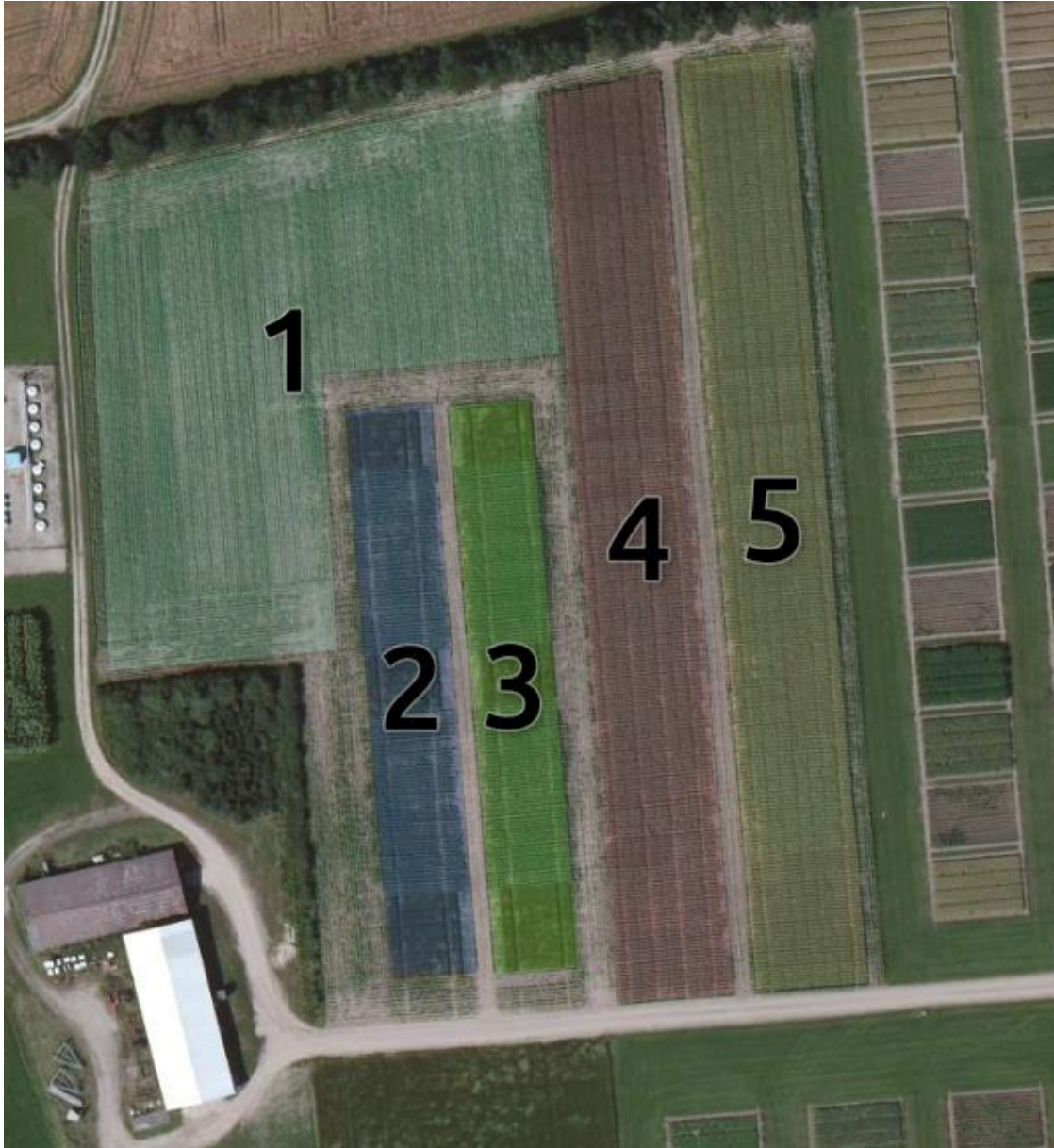
## Adding zones to the field based on aerial photos based on previous years

In order to improve the spatial map, we wish to include information about how the field was divided in 2013 and 2014. We do this by using aerial photos freely available online, these photos comes from Krak maps and Microsoft virtual earth.



Based on the ortho map from 2013 add a set of zones as a factor.





```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/morten/Optimek", layer: "bing_zoner"
## with 5 features
## It has 1 fields
```

Doing the same for the other year 2014 using Krak



We add a set of zones as a factor



```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/morten/Optimek", layer: "krak_zoner"
## with 2 features
## It has 1 fields
```

The assignment of the parcels were done in a relatively simple manner based on the center coordinate of the parcel. This is a relatively rough approach since a parcel may have different fractions of area within the different zones. This is however ignored (but not forgotten) in the following analysis. Furthermore the interaction between the x and y direction is included into the model.

### Making a spatial model, which uses the zones

```
filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines"
```

```
##
## Call:
```

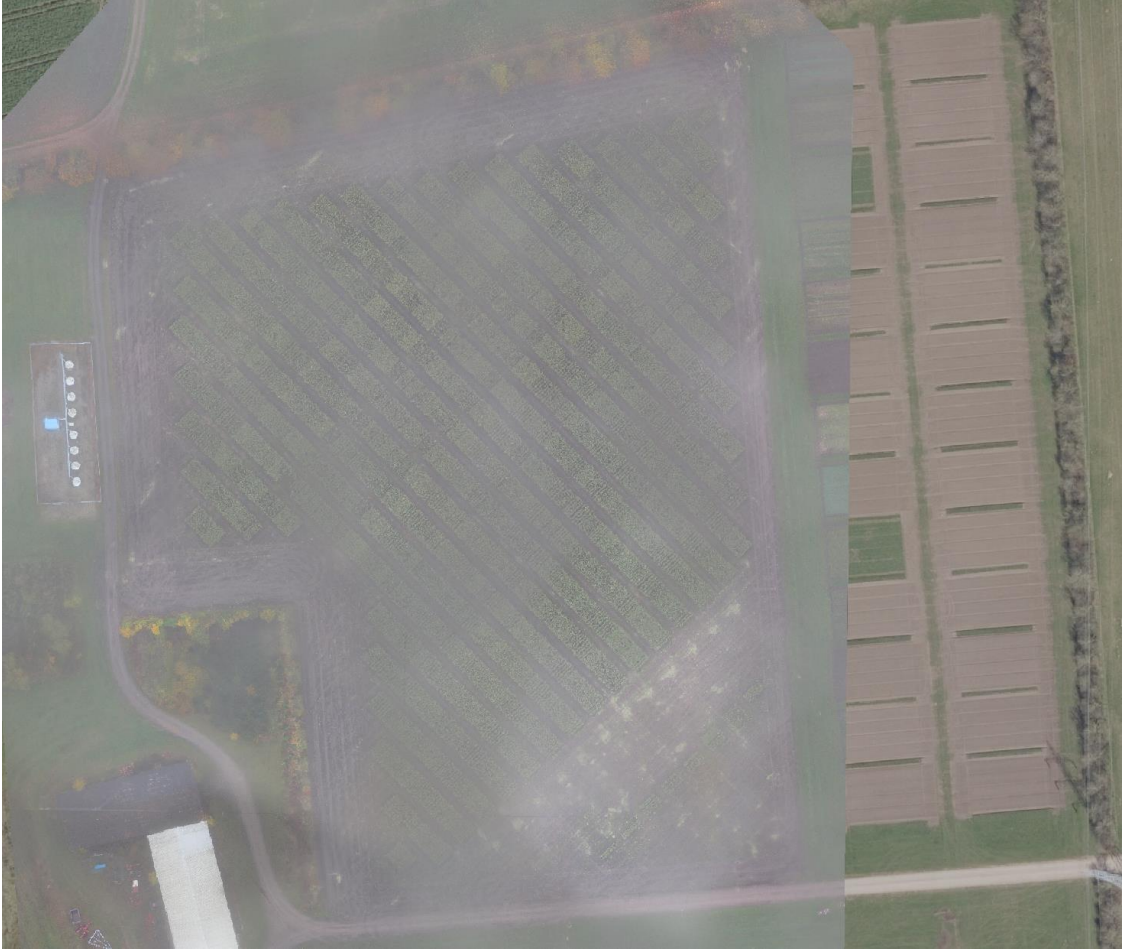
```

## lm(formula = TS_YieldParcel..ton.ha. ~ poly(x_center, 2) * poly(y_center,
##      2) + bingZone + krakZone, data = NTFS_Harvest[filter, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.71622 -0.45887 -0.00732  0.36948  2.13252
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)          6.1269    0.6445   9.506 4.79e-14
## poly(x_center, 2)1    -2.6454    2.5296  -1.046 0.299416
## poly(x_center, 2)2    -1.9275    1.9864  -0.970 0.335342
## poly(y_center, 2)1     6.3753    1.5630   4.079 0.000123
## poly(y_center, 2)2     1.4158    1.0613   1.334 0.186694
## bingZone1              0.8130    0.4522   1.798 0.076673
## bingZone2              0.4845    0.4676   1.036 0.303842
## bingZone3              0.3463    0.3694   0.938 0.351846
## bingZone4             -0.6901    0.3067  -2.250 0.027727
## bingZone5              0.1238    0.5203   0.238 0.812711
## krakZone2              2.0007    0.6263   3.195 0.002136
## poly(x_center, 2)1:poly(y_center, 2)1 -16.4907   19.6507  -0.839 0.404346
## poly(x_center, 2)2:poly(y_center, 2)1  28.0578   14.0165   2.002 0.049360
## poly(x_center, 2)1:poly(y_center, 2)2   2.6109   13.5713   0.192 0.848026
## poly(x_center, 2)2:poly(y_center, 2)2 -25.7888   12.6500  -2.039 0.045434
##
## (Intercept)          ***
## poly(x_center, 2)1
## poly(x_center, 2)2
## poly(y_center, 2)1          ***
## poly(y_center, 2)2
## bingZone1                  .
## bingZone2
## bingZone3
## bingZone4                  *
## bingZone5
## krakZone2                  **
## poly(x_center, 2)1:poly(y_center, 2)1
## poly(x_center, 2)2:poly(y_center, 2)1 *
## poly(x_center, 2)1:poly(y_center, 2)2
## poly(x_center, 2)2:poly(y_center, 2)2 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7641 on 67 degrees of freedom
## Multiple R-squared:  0.725, Adjusted R-squared:  0.6675
## F-statistic: 12.61 on 14 and 67 DF, p-value: 9.116e-14

```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	6.13	0.64	9.51	0.00

poly(x_center, 2)1	-2.65	2.53	-1.05	0.30
poly(x_center, 2)2	-1.93	1.99	-0.97	0.34
poly(y_center, 2)1	6.38	1.56	4.08	0.00
poly(y_center, 2)2	1.42	1.06	1.33	0.19
bingZone1	0.81	0.45	1.80	0.08
bingZone2	0.48	0.47	1.04	0.30
bingZone3	0.35	0.37	0.94	0.35
bingZone4	-0.69	0.31	-2.25	0.03
bingZone5	0.12	0.52	0.24	0.81
krakZone2	2.00	0.63	3.19	0.00
poly(x_center, 2)1:poly(y_center, 2)1	-16.49	19.65	-0.84	0.40
poly(x_center, 2)2:poly(y_center, 2)1	28.06	14.02	2.00	0.05
poly(x_center, 2)1:poly(y_center, 2)2	2.61	13.57	0.19	0.85
poly(x_center, 2)2:poly(y_center, 2)2	-25.79	12.65	-2.04	0.05

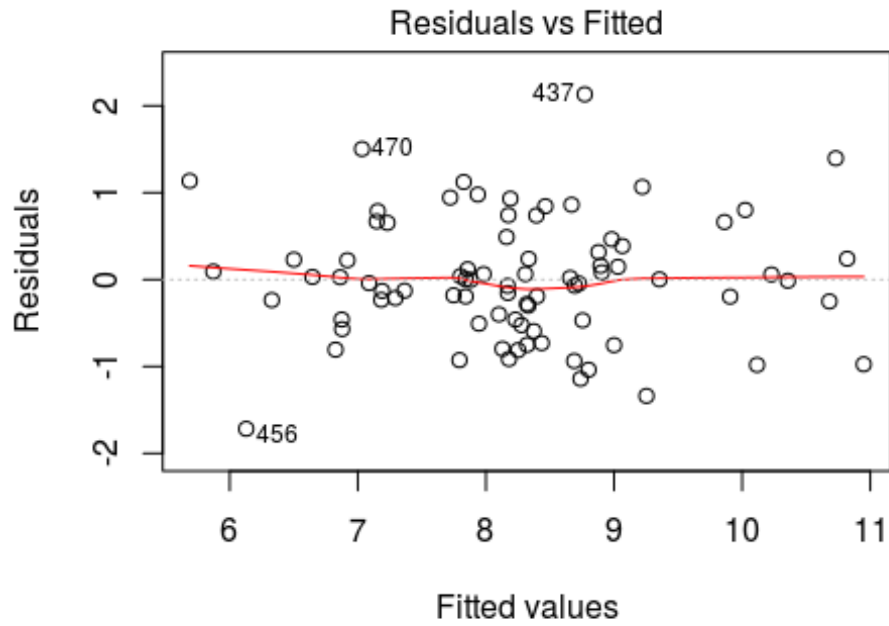




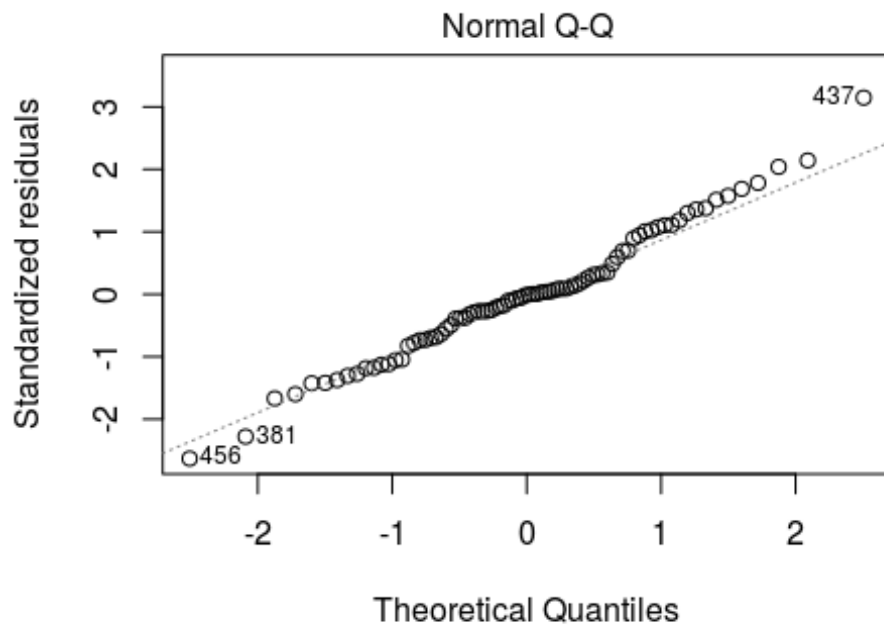


It is clear the Zones adds a significant contribution to the spatial model. The “krakZone2” have a significant higher yield level compared to “krakZone1” which is the effect of two different crops seeded in 2014 the year prior to the trial in 2015. In the 2013 two years prior to the trial the zones still have a significant effect on the yields measured in 2015 two years later. The “bingZone0” which were headland areas in 2013 seems to cause a lower yield in 2015 compared to the other zones except “bingZone4” which is the only zone having a significant lower yield compared to the headland zone. Still there seems to be a trend towards increasing yield effect from the headland zone “bingZone0” → “bingZone5” → “bingZone3” → “bingZone2” → “bingZone1”. However separately this effect is not significantly different from “bingZone0”. Studying the residual plot the spatial model seems relatively healthy.

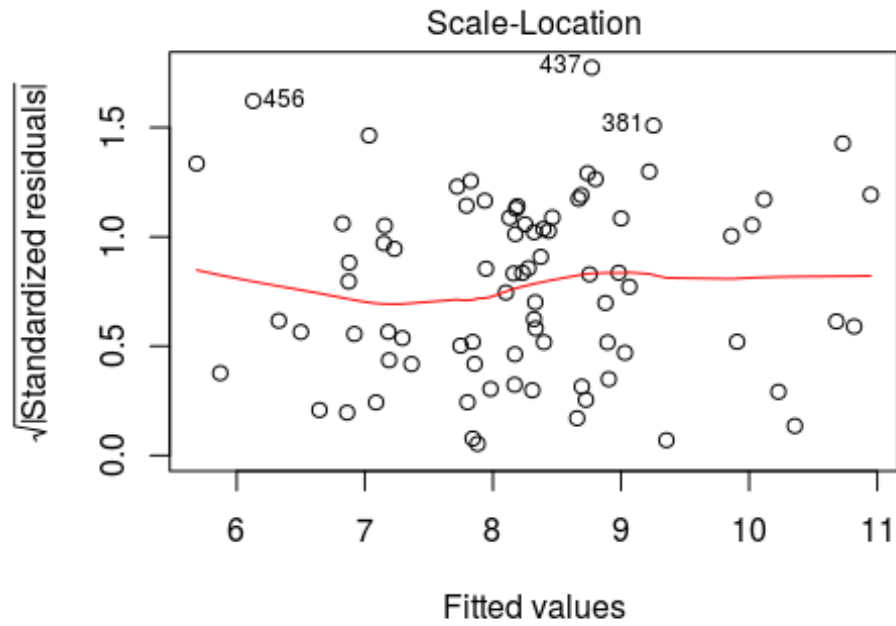




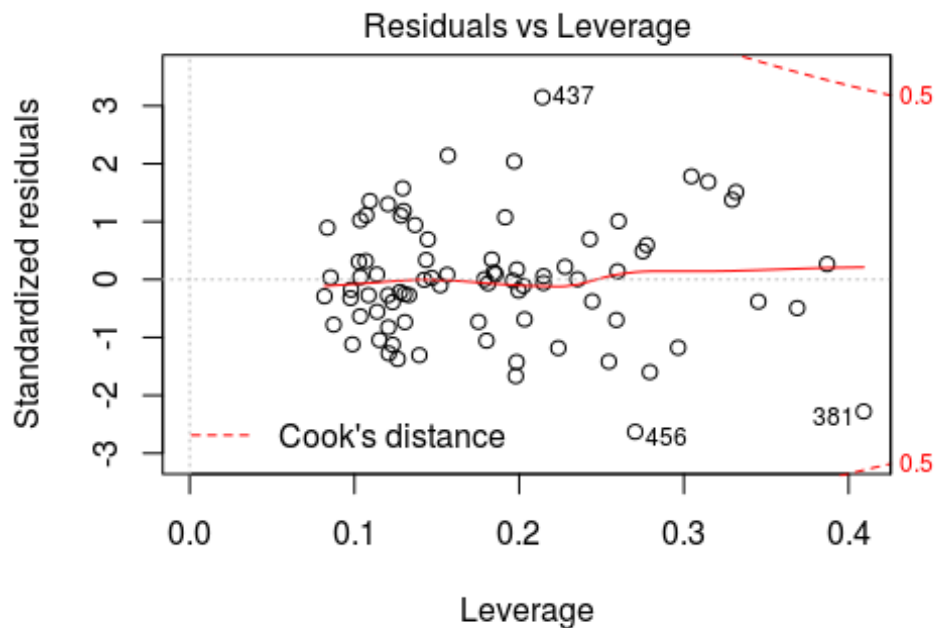
$n(\text{TS\_YieldParcel}..\text{ton.ha.} \sim \text{poly}(x\_center, 2) * \text{poly}(y\_center, 2) + \text{bir})$



$n(\text{TS\_YieldParcel}..\text{ton.ha.} \sim \text{poly}(x\_center, 2) * \text{poly}(y\_center, 2) + \text{bir})$

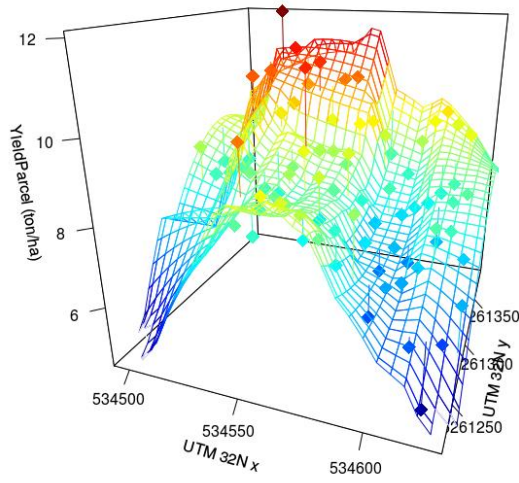


$n(\text{TS\_YieldParcel..ton.ha.} \sim \text{poly}(\text{x\_center}, 2) * \text{poly}(\text{y\_center}, 2) + \text{bir})$

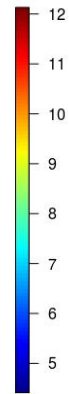


$n(\text{TS\_YieldParcel..ton.ha.} \sim \text{poly}(\text{x\_center}, 2) * \text{poly}(\text{y\_center}, 2) + \text{bir})$

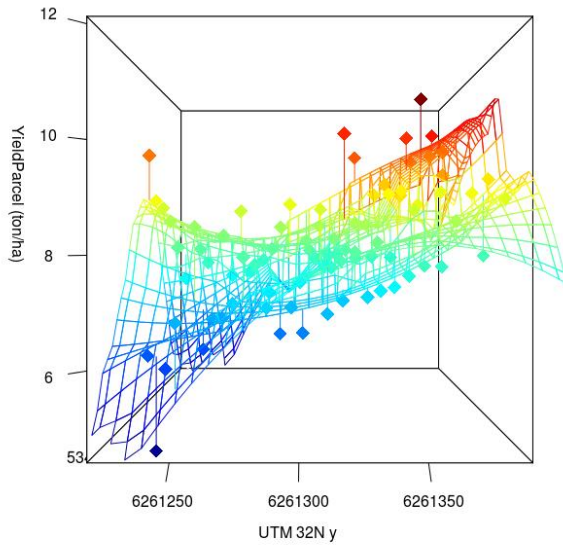
Spatial yield model



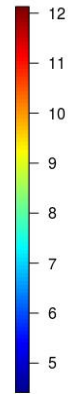
YieldParcel (ton/ha)



Spatial yield model



YieldParcel (ton/ha)



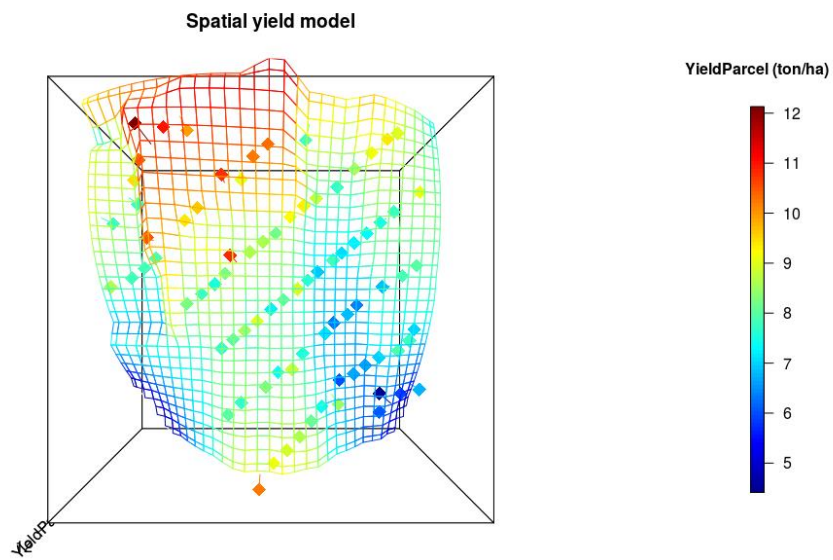
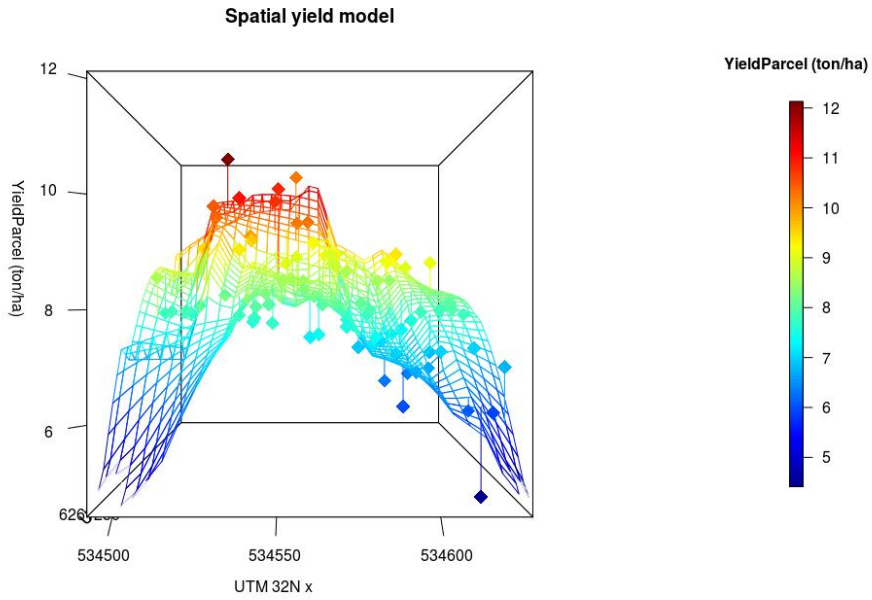
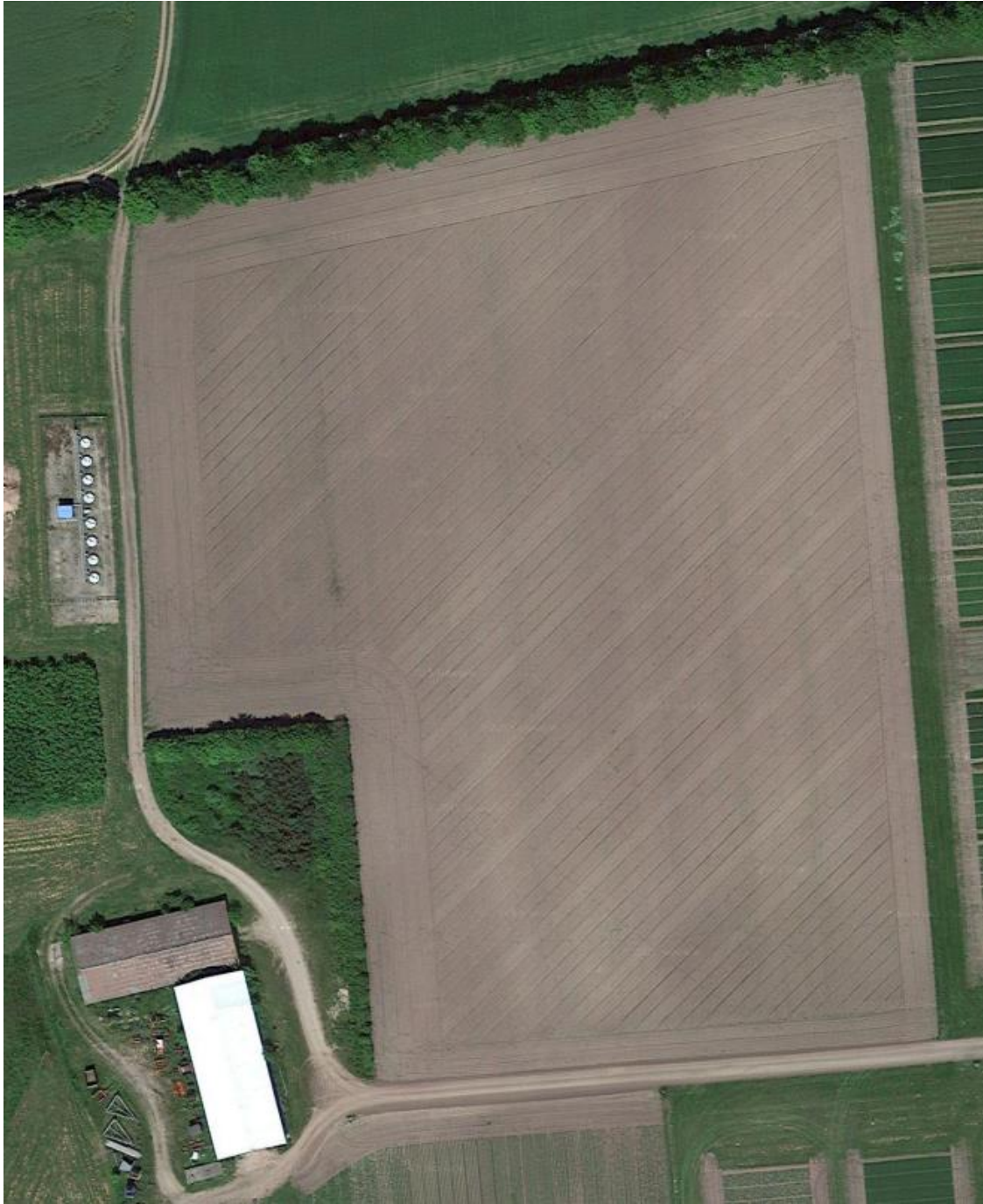


Image captured from a drone showing the spatial distribution within the field indicate the latter spatial model of the yield levels seems plausible.



### Other aerial photos

From google maps taken 12 June 2015. Notice weed patches in the NW part of the field in the transition between the Krakzones in 2015.



From Kort og Matrikelstyrelsen taken spring 2014, just before leafing.



## Making a full model of the yield, utilizing the spatial model

The spatial model setting the yield level of the current location within the field is included in the full yield model with the treatment factors (Row Distance, Seeding Pattern, and Seeding density) as a covariate “spatialCorrection”. Furthermore seeding densities below 6 plants/ha and above 16 plants per hectare is excluded.

```
filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines" &
  (NTFS_Harvest$NTFS_F1_ThousandSeedsPerHectar >= 60) &
  (NTFS_Harvest$NTFS_F1_ThousandSeedsPerHectar <= 160)

##
## Call:
```

```

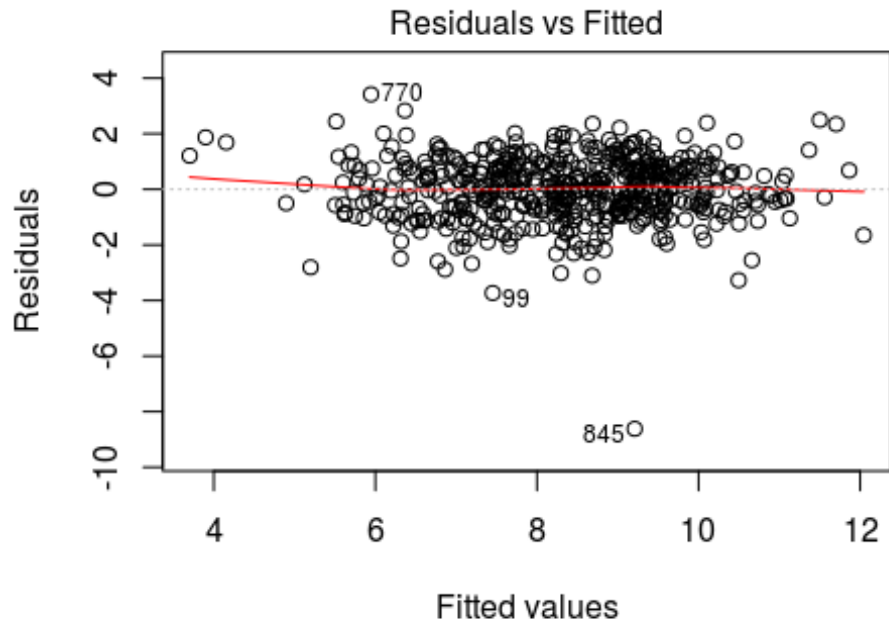
## lm(formula = TS_YieldParcel..ton.ha. ~ RowDist *
NTFS_F1_ThousandSeedsPerHectar +
##      poly(NTFS_F1_ThousandSeedsPerHectar, 2) + poly(RowDist, 2) +
##      RowDist * SeedPattern + spatialCorrection, data =
as(NTFS_Harvest[filter,
##      ], "data.frame"))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.6149 -0.6736  0.0218  0.7453  3.4192
##
## Coefficients: (2 not defined because of singularities)
##
##              Estimate Std. Error t value
## (Intercept)    -3.708e+00  6.768e-01  -5.478
## RowDist         3.623e-03  1.108e-03   3.268
## NTFS_F1_ThousandSeedsPerHectar
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)1          NA          NA          NA
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)2  -6.460e+00  1.131e+00  -5.713
## poly(RowDist, 2)1          NA          NA          NA
## poly(RowDist, 2)2    -3.021e+00  1.149e+00  -2.629
## SeedPatternSync      1.417e+00  3.110e-01   4.557
## spatialCorrection     8.726e-01  4.454e-02  19.590
## RowDist:NTFS_F1_ThousandSeedsPerHectar  -2.282e-05  9.355e-06  -2.439
## RowDist:SeedPatternSync  -2.480e-03  5.693e-04  -4.356
##
##              Pr(>|t|)
## (Intercept)    6.68e-08 ***
## RowDist         0.00115 **
## NTFS_F1_ThousandSeedsPerHectar
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)1          NA
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)2  1.86e-08 ***
## poly(RowDist, 2)1          NA
## poly(RowDist, 2)2    0.00882 **
## SeedPatternSync      6.47e-06 ***
## spatialCorrection    < 2e-16 ***
## RowDist:NTFS_F1_ThousandSeedsPerHectar  0.01505 *
## RowDist:SeedPatternSync  1.60e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.126 on 525 degrees of freedom
## Multiple R-squared:  0.6, Adjusted R-squared:  0.5939
## F-statistic: 98.45 on 8 and 525 DF,  p-value: < 2.2e-16

```

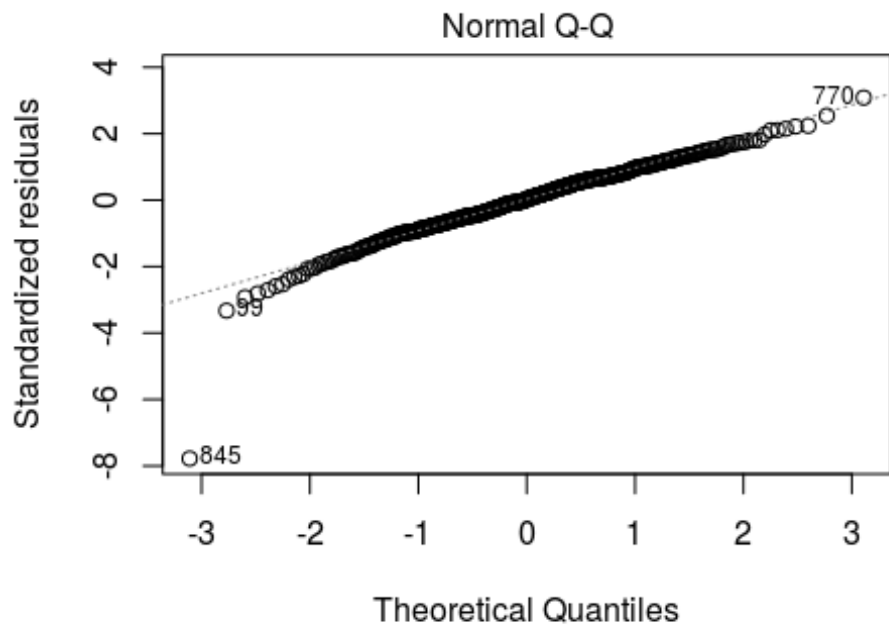
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-3.71	0.68	-5.48	0.00
RowDist	0.00	0.00	3.27	0.00
NTFS_F1_ThousandSeedsPerHectar	0.04	0.01	7.84	0.00
poly(NTFS_F1_ThousandSeedsPerHectar, 2)2	-6.46	1.13	-5.71	0.00



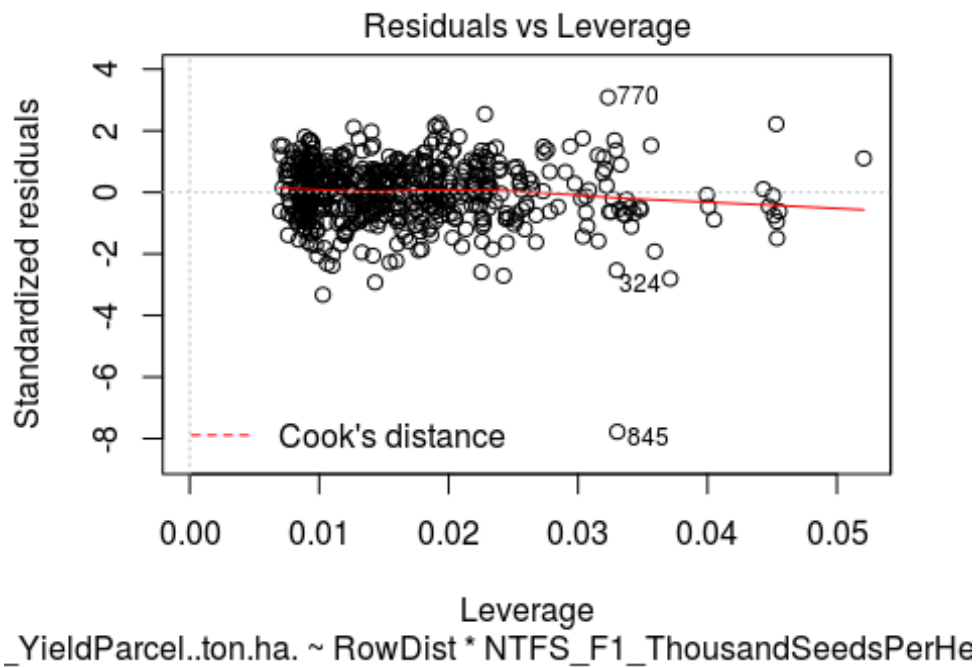
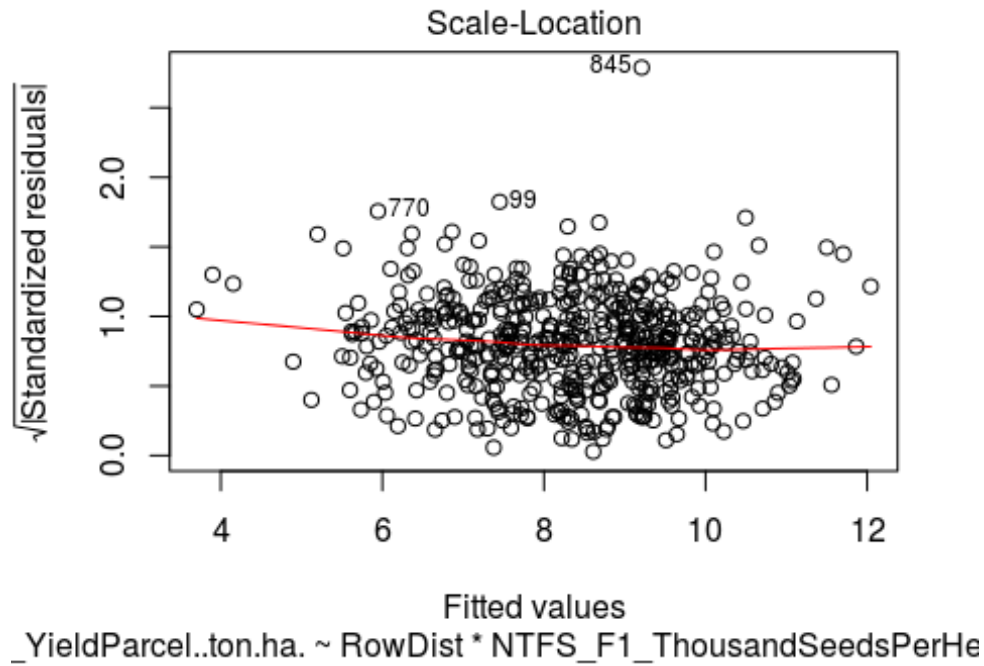
poly(RowDist, 2) <sup>2</sup>	-3.02	1.15	-2.63	0.01
SeedPatternSync	1.42	0.31	4.56	0.00
spatialCorrection	0.87	0.04	19.59	0.00
RowDist:NTFS_F1_ThousandSeedsPerHectar	0.00	0.00	-2.44	0.02
RowDist:SeedPatternSync	0.00	0.00	-4.36	0.00



$\text{YieldParcel.ton.ha.} \sim \text{RowDist} * \text{NTFS\_F1\_ThousandSeedsPerHe}$



$\text{\_YieldParcel..ton.ha.} \sim \text{RowDist} * \text{NTFS\_F1\_ThousandSeedsPerHe}$



The model presented above is the result of several tests with different interactions and shifts between having the treatment variables as factor variables and covariates. Studying the residual plots the model seems to do a decent job modelling the data except for a single

extreme with ID845. The Normal Q-Q plot is hanging a little for the lower values but we will accept this for now.

The actual yield can be considered as a mean yield levels, due to including the mean yield level from the spatial model as a covariate. The spatialCorrection is highly significant. The Seeding density or thousand plants per m2 (10 x plants per m2) "NTFS\_F1\_ThousandSeedsPerHectar" and the row distance "RowDistance" is highly significant with a 2. Order polynomial with a negative 2. Order → having a maximum yield. Furthermore the first order of NTFS\_F1\_ThousandSeedsPerHectar and "RowDist" interact with a negative value. This means as the seeding density and/or the row distance increase the contribution the yield decreases (remember to include the separate contributions from the two variables when try to comprehend the full model). As consequence this interpretation becomes relatively complex and plots of the model might help.

The Seeding pattern is also highly significant and interacts negatively with the row distance. The synchron pattern yields higher yields than the offset pattern ... BUT it as the row distance increases this effect might flip within the range of the model (6 to 16 plant per hectare). The yield contribution to the model for actually seeding row distances and seeding patterns will be:

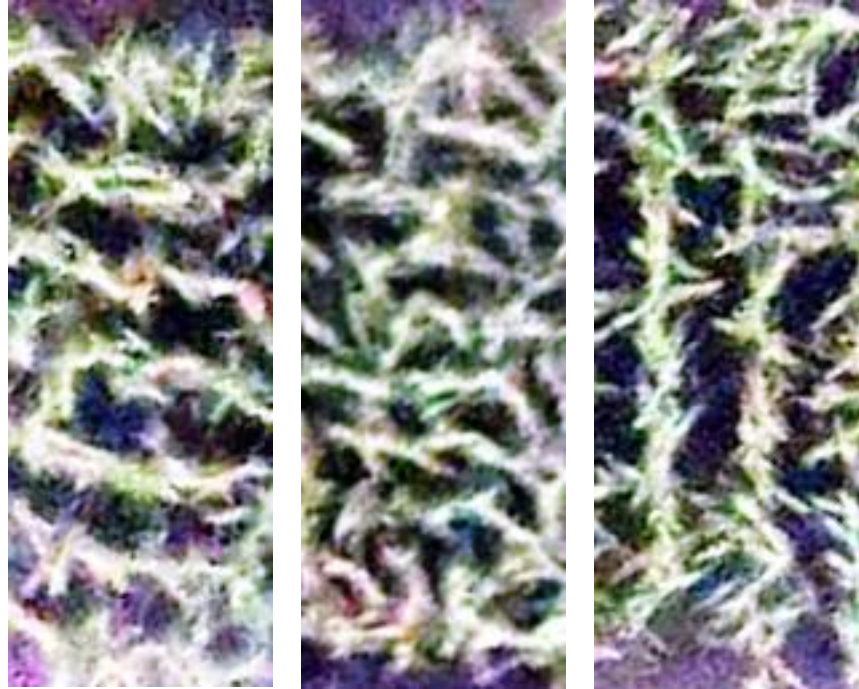
|\*\*\*\*|**Seeding Pattern**| |Row distance|Offset|Sync| |250| |375| |500| |625| |750\*\*| Morten indsæt tabel ala nedenfor ... gerne med indikator for signifikante forskelle

Seeding pattern Row distance offset sync 250 0.91 1.70 375 1.36 1.85 500 1.81 1.99 625 2.26 2.13 750 2.72 2.27

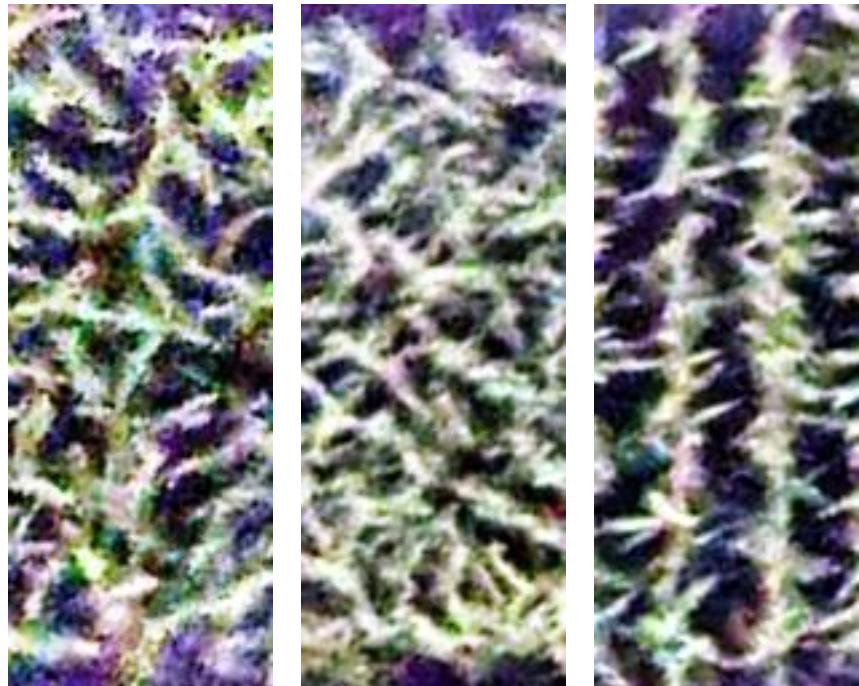
The table shows that the synchron pattern have a positive effect on the yield compared the the offset pattern at the lowest row distance. This is a kind of surprising meaning that the synch pattern will at narrow row distance and seeding densities create orthogonal rows indirectly in the SE ↔ NW direction.

<b>Pattern\RowWidth</b>	<b>250</b>	<b>375</b>	<b>750</b>
-------------------------	------------	------------	------------

**SYNC**



**Offset**

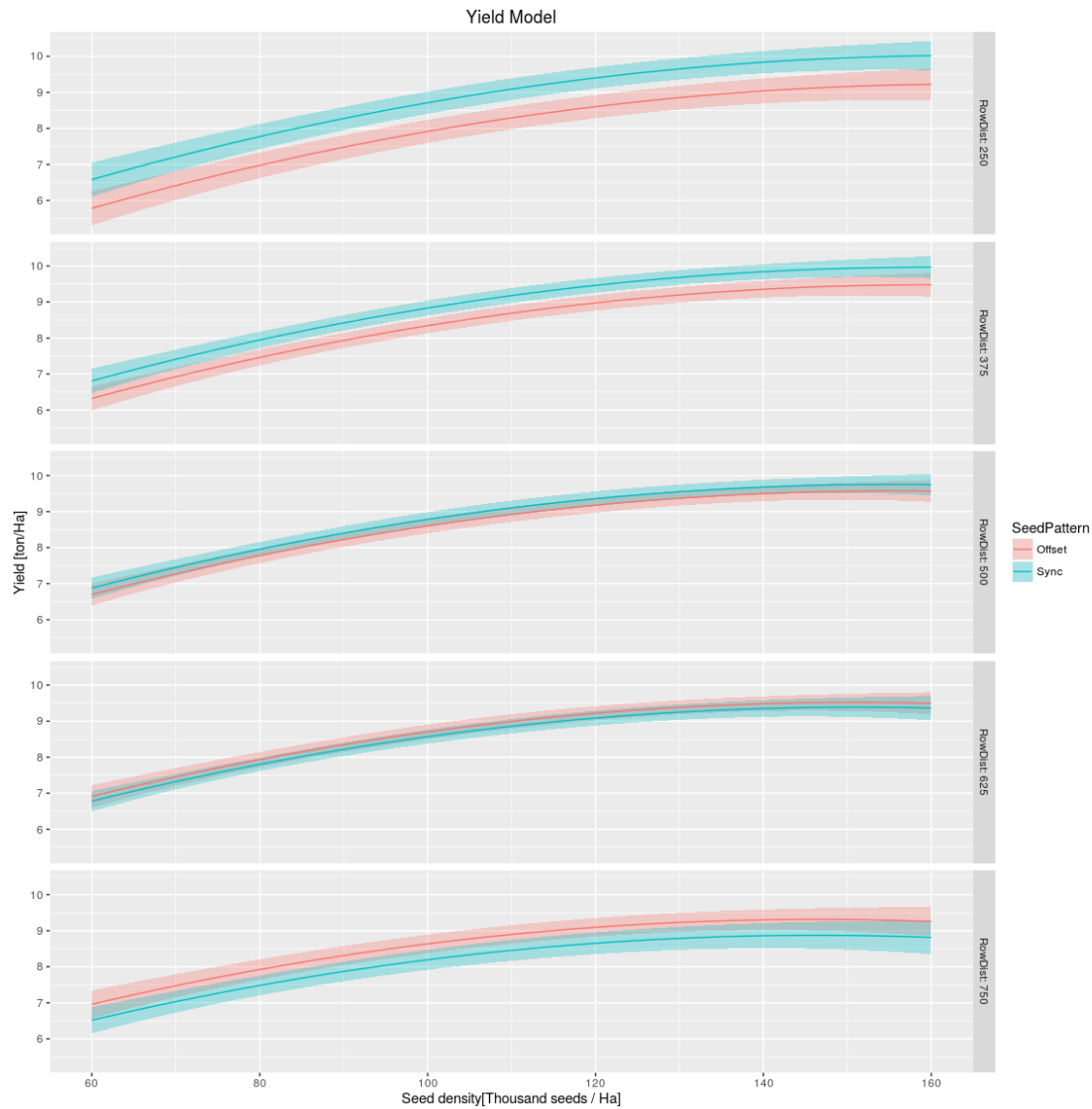


\*Note the orientation of the cropped illustrative parcels is rotated from the original SW  $\leftarrow$   $\rightarrow$  NE to S  $\leftrightarrow$  N.

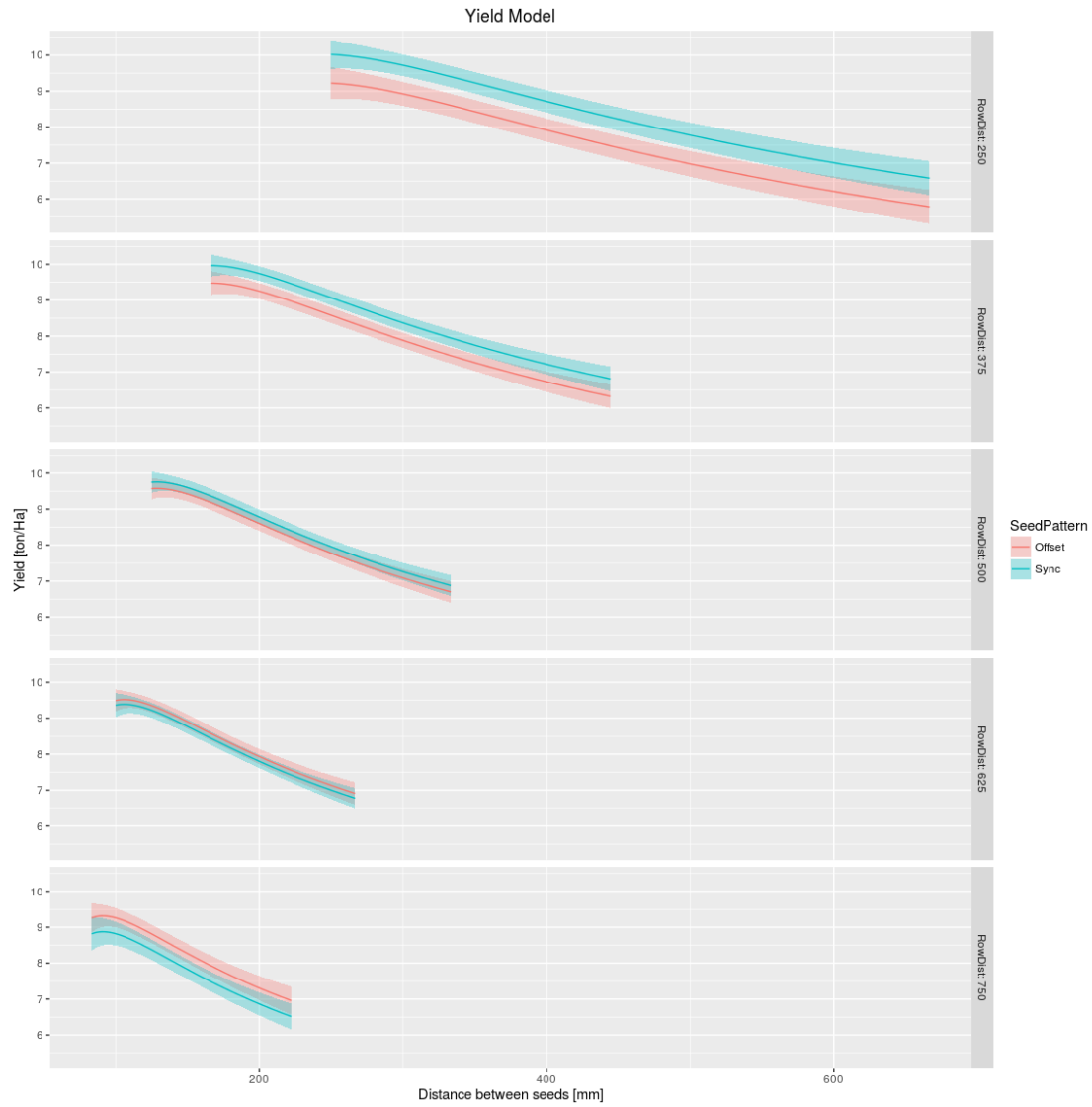
Below the model is illustrated by plotting the seed density versus the yield for the offset and sync pattern, respectively. The colored bands indicate the uncertainty of the predicted yields with 95%. If the the red and blue bands overlap they are not significantly different. Thus the yield is only significantly different for the row distance 250 and 375 mmm

especially in the middle range of the spanned range of seed densities (6 to 16 plants per m<sup>2</sup>)

```
## Warning in predict.lm(Full_Model_With_Zones_And_Spatial_Correction,  
newdata  
## = grid.pred, : prediction from a rank-deficient fit may be misleading
```

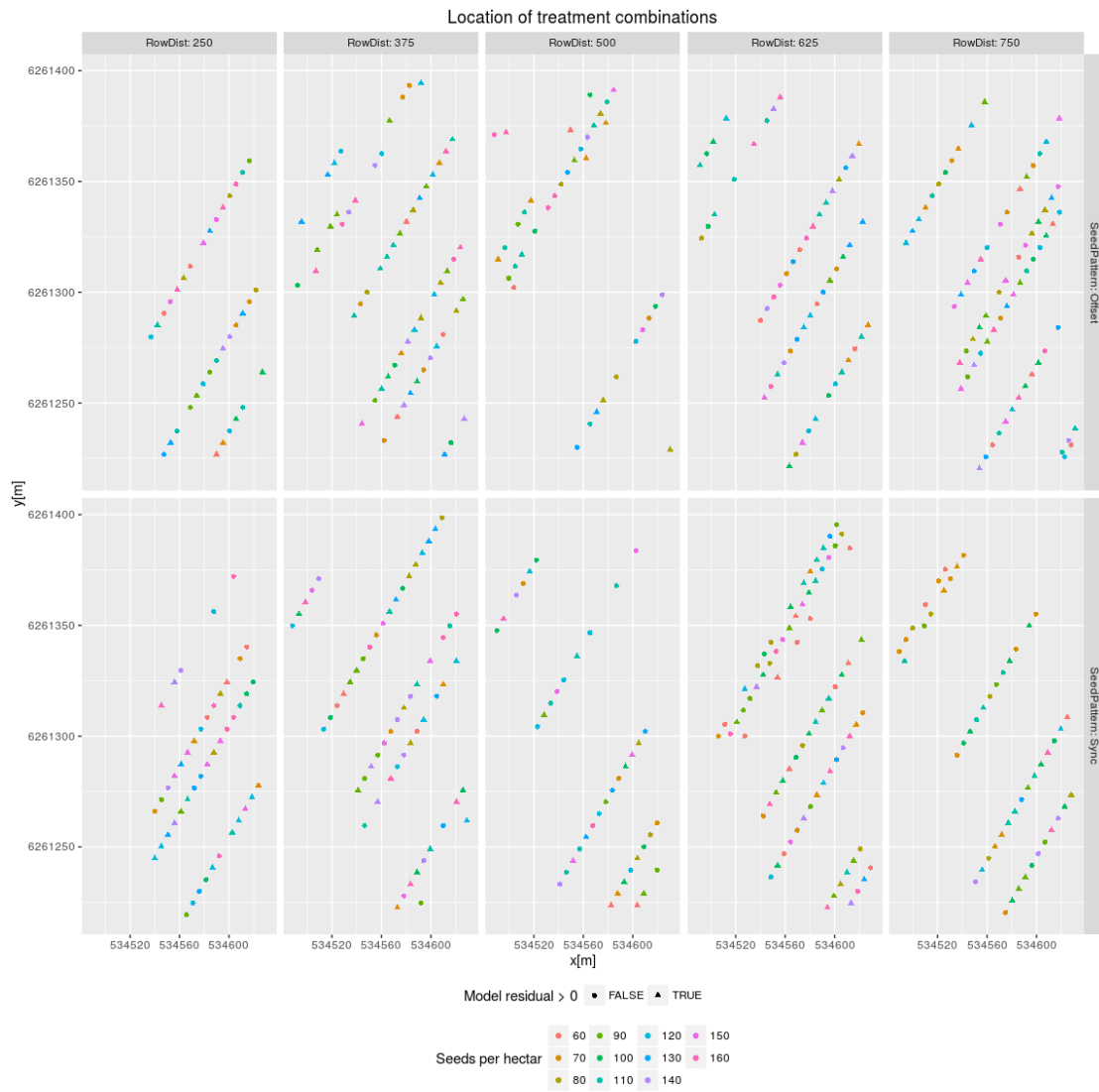


The latter figure can also be illustrated as the distance between the single crop plants in the row versus the yield. Note the highest crop densities is now to the left decreasing moving to the right.

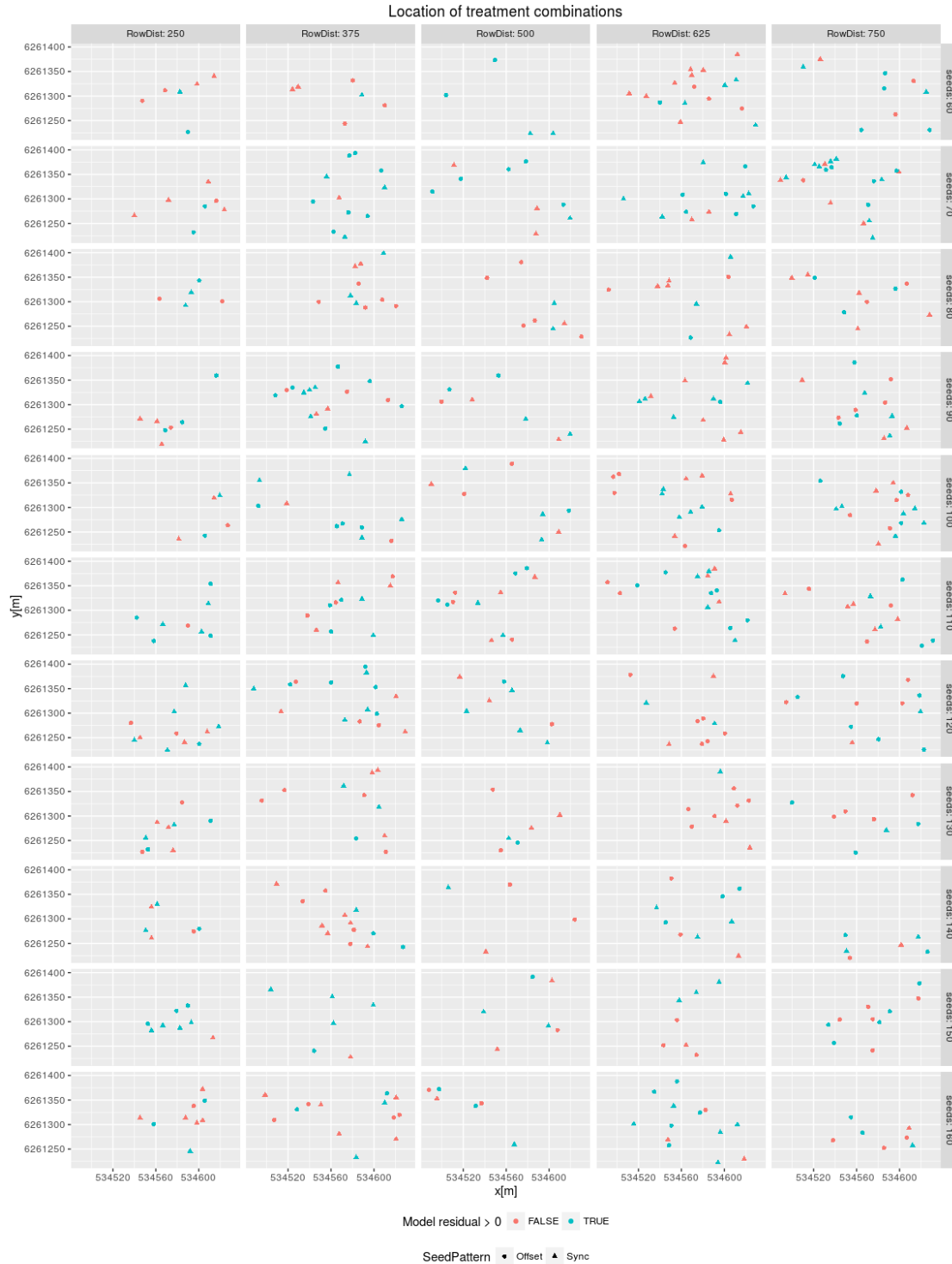


It is somewhat surprising the sync pattern performs better than the offset pattern at the narrower row distances... assuming the model is valid!

# Spatial location of parcel combinations







## Adding a interaction between the spatial correction and the rowdistance

```
filter <- !outliers & NTFS_Harvest$SeedPattern!="SyncTramLines" &
  (NTFS_Harvest$NTFS_F1_ThousandSeedsPerHectar >= 60) &
  (NTFS_Harvest$NTFS_F1_ThousandSeedsPerHectar <= 160)
```

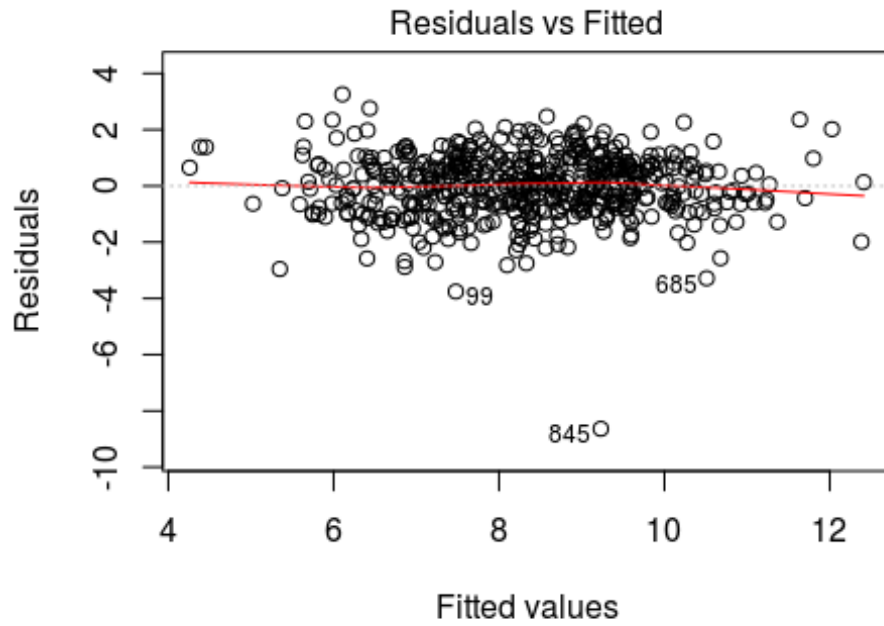
```

##
## Call:
## lm(formula = TS_YieldParcel..ton.ha. ~ RowDist *
NTFS_F1_ThousandSeedsPerHectar +
##   poly(NTFS_F1_ThousandSeedsPerHectar, 2) + poly(RowDist, 2) +
##   RowDist * SeedPattern + spatialCorrection +
NTFS_F1_ThousandSeedsPerHectar *
##   spatialCorrection, data = as(NTFS_Harvest[filter, ], "data.frame"))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.6365 -0.6512  0.0417  0.7454  3.2588
##
## Coefficients: (2 not defined because of singularities)
##
##              Estimate Std. Error
## (Intercept)   -8.341e-01  1.420e+00
## RowDist         3.727e-03  1.105e-03
## NTFS_F1_ThousandSeedsPerHectar
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)1
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)2
## poly(RowDist, 2)1
## poly(RowDist, 2)2
## SeedPatternSync
## spatialCorrection
## RowDist:NTFS_F1_ThousandSeedsPerHectar
## RowDist:SeedPatternSync
## NTFS_F1_ThousandSeedsPerHectar:spatialCorrection
##
##              t value Pr(>|t|)
## (Intercept)   -0.587 0.557263
## RowDist        3.374 0.000796 ***
## NTFS_F1_ThousandSeedsPerHectar
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)1
## poly(NTFS_F1_ThousandSeedsPerHectar, 2)2
## poly(RowDist, 2)1
## poly(RowDist, 2)2
## SeedPatternSync
## spatialCorrection
## RowDist:NTFS_F1_ThousandSeedsPerHectar
## RowDist:SeedPatternSync
## NTFS_F1_ThousandSeedsPerHectar:spatialCorrection
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.122 on 524 degrees of freedom
## Multiple R-squared:  0.604, Adjusted R-squared:  0.5972
## F-statistic: 88.81 on 9 and 524 DF,  p-value: < 2.2e-16

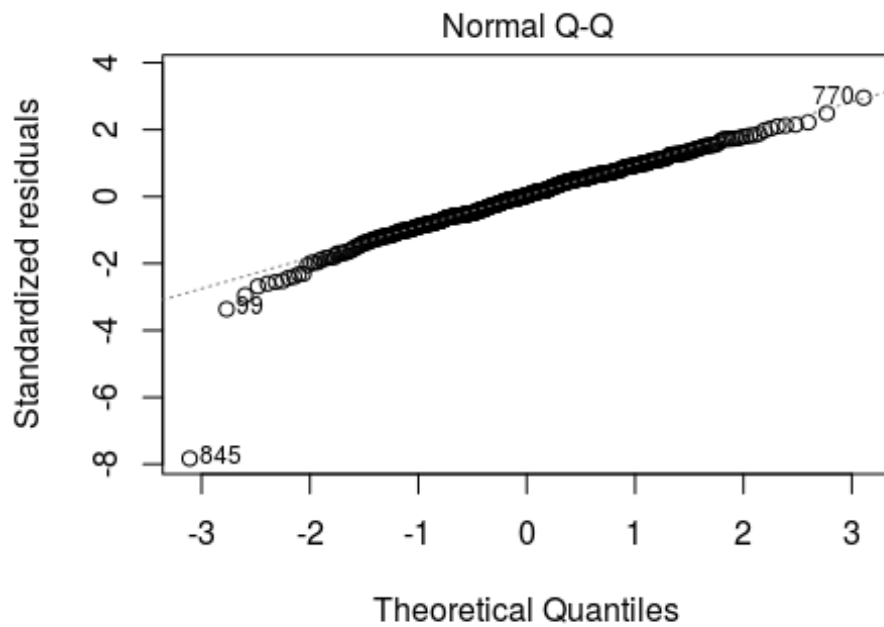
```

	Estimate	Std. Error	t value	Pr(> t )
--	----------	------------	---------	----------

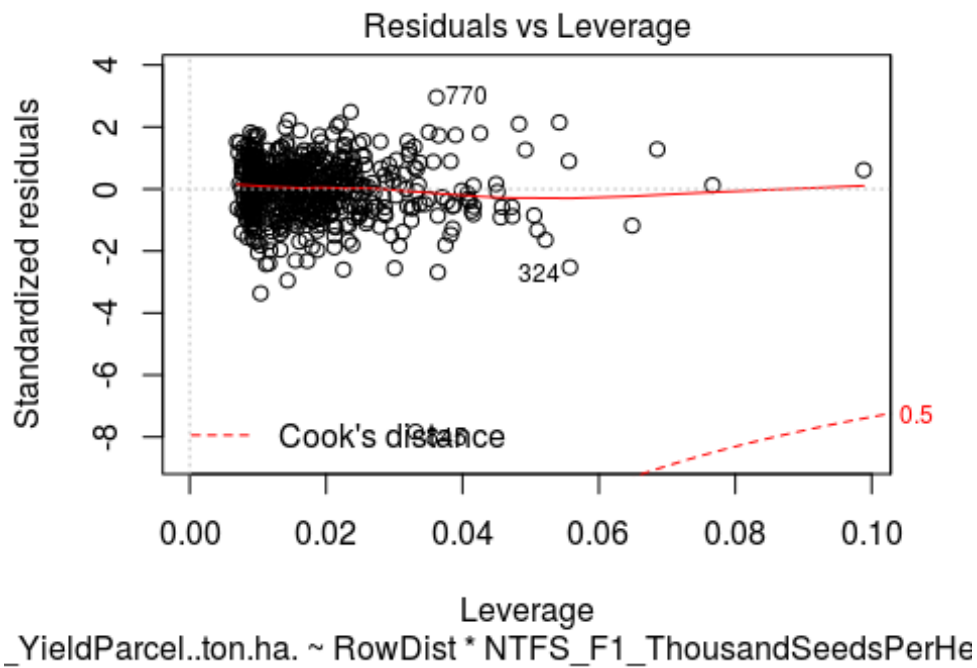
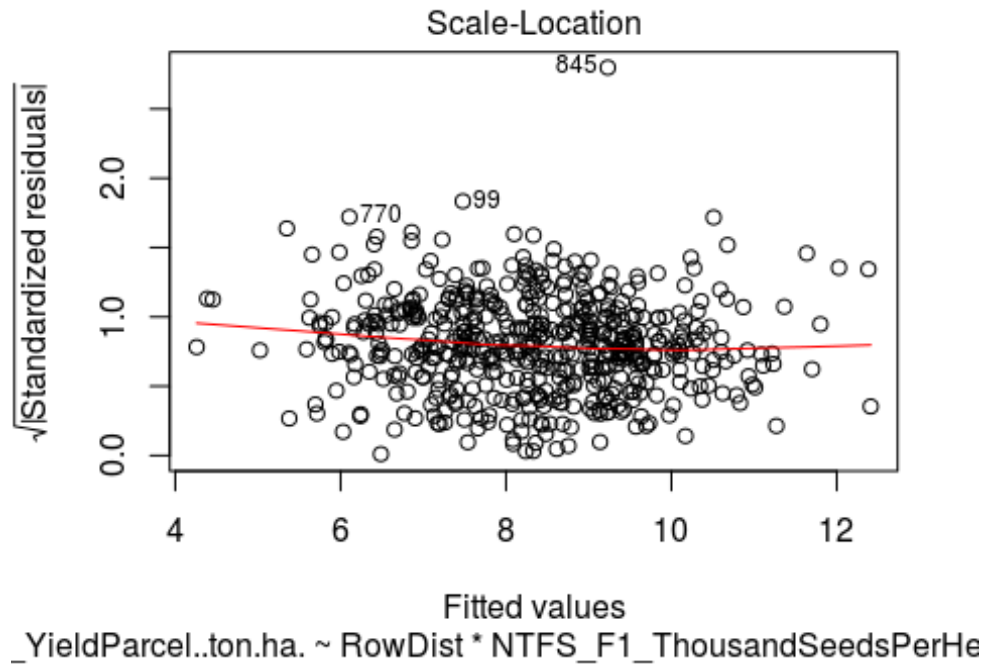
(Intercept)	-0.83	1.42	-0.59	0.56
RowDist	0.00	0.00	3.37	0.00
NTFS_F1_ThousandSeedsPerHectar	0.01	0.01	0.95	0.34
poly(NTFS_F1_ThousandSeedsPerHectar, 2)2	-6.50	1.13	-5.77	0.00
poly(RowDist, 2)2	-2.61	1.16	-2.25	0.02
SeedPatternSync	1.36	0.31	4.39	0.00
spatialCorrection	0.51	0.16	3.08	0.00
RowDist:NTFS_F1_ThousandSeedsPerHectar	0.00	0.00	-2.57	0.01
RowDist:SeedPatternSync	0.00	0.00	-4.10	0.00
NTFS_F1_ThousandSeedsPerHectar:spatialCorrection	0.00	0.00	2.30	0.02



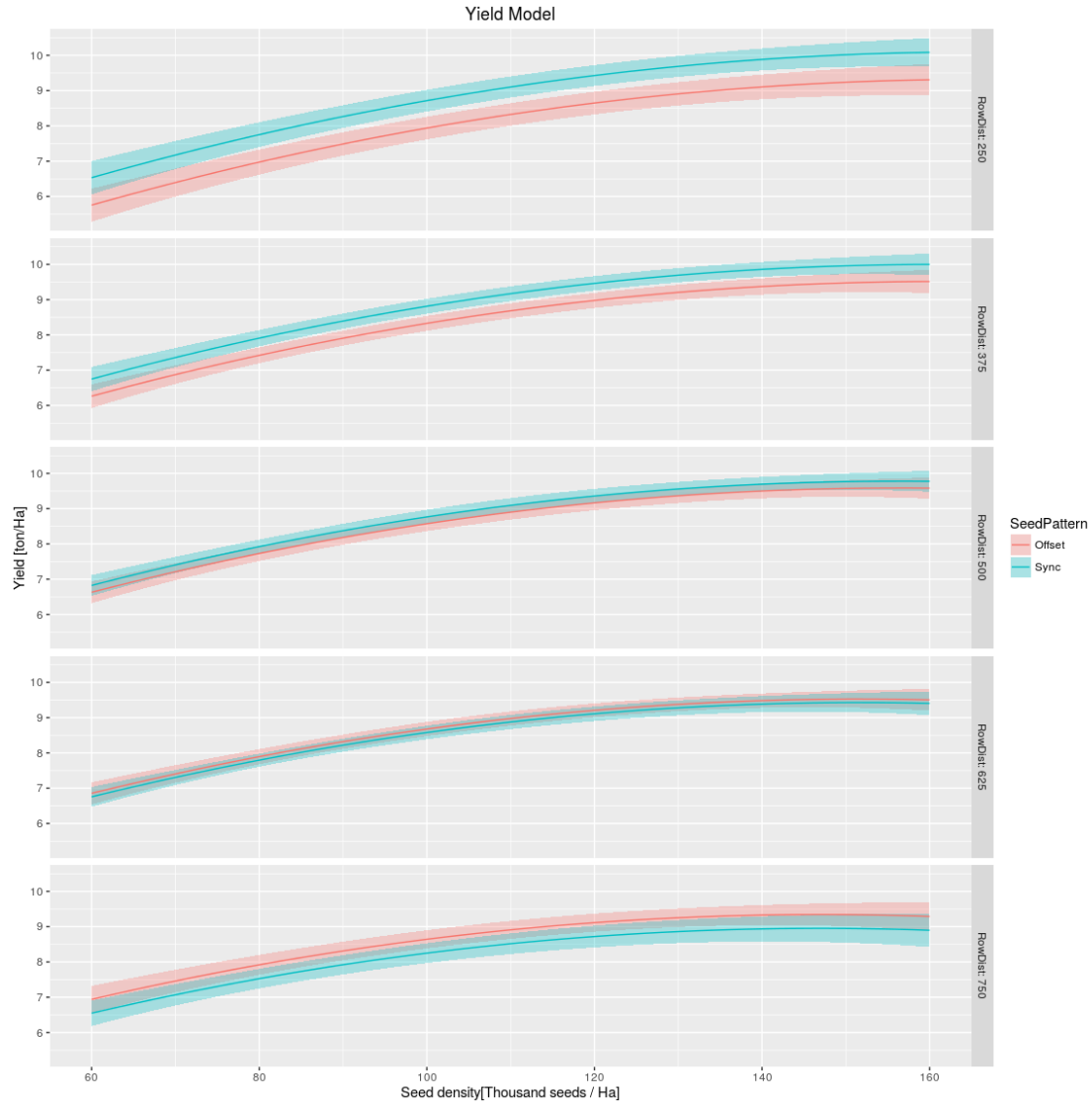
$\text{YieldParcel.ton.ha.} \sim \text{RowDist} * \text{NTFS\_F1\_ThousandSeedsPerHe}$



$\text{\_YieldParcel..ton.ha.} \sim \text{RowDist} * \text{NTFS\_F1\_ThousandSeedsPerHe}$



```
## Warning in predict.lm(Full_Model_With_Zones_And_Spatial_Correction,
newdata
## = grid.pred, : prediction from a rank-deficient fit may be misleading
```



The latter figure can also be illustrated as the distance between the single crop plants in the row versus the yield. Note the highest crop densities is now to the left decreasing moving to the right.

Yield Model

