Den Europæiske Landbrugsfond for Udvikling af Landdistrikterne:
Danmark og Europa investerer i landdistrikterne

Miljø- og Fødevareministeriet
Landbrugsstyrelsen

Den Europæiske Landbrugsfond
for Udvikling af Landdistrikterne

LDP 2020

Se EU-Kommissionen, Den Europæiske Landbrugsfond for Udvikling af Landdistrikterne

TEKNOLOGISK
INSTITUT

Agro Food Park 15,
Skejby
8200 Aarhus N
Tlf. +45 7220 2000
info@teknologisk.dk

Notat: Satellitdata-baseret prædiktion af TS-udbytte og TS-procent i majs og slætgræs

31-01-2019
Philipp Trénel, ph.d., seniorkonsulent, specialist
phtr@teknologisk.dk, +45 72 20 33 92
Marlene Trinderup, konsulent
matr@teknologisk.dk, +45 72 20 32 97
Søren Kjærgaard Boldsen, Ph.D. Specialist
sobo@teknologisk.dk, Mobil +45 72 20 24 89

Formål

- At undersøge, om sentinel-2-satellitdata effektivt kan bruges til at bestemme tørstofudbytte (TS-udbytte i ton/ha) og tørstofprocent (TS%) i danske grovfodermarker på baggrund af et af SEGES indhentet datasæt over observerede TS-udbytter og TS% i majs- (N=1461) og slætgræsmarker (N=1734) fra 2017 og 2018. Der undersøges også, om tilføjelse af andre data, f.eks. klimadata, jordbundsdata, sortsnavn, forfrugt, m.m., kan forbedre prædiktionen.

Overordnede konklusioner

- **Majs**: de her anvendte modeller kunne på baggrund af satellitdata opnå en testsæt-prædiktionsnøjagtighed (RMSEP) for TS-udbytte og TS% i majs på hhv. ±2.3t/ha og 2.7%-point.
- **Græs:** For græs forelå der kun tilstrækkeligt datamateriale for slætnr. 1 til 3. Her var de fundne testsæt-prædiktionsnøjagtigheder (RMSEP) for modellerne på baggrund af satellitdata for TS-udbytte på ±1.4t/ha, ±1.1t/ha og ±0.9t/ha for slætnr. 1 til 3.
- For begge afgrøder udviste modellerne bygget på yderligere datakilder udover satellitdata generelt bedre prædiktioner på træningsdata (højere R^2), særligt som effekt af inkludering af klimavariabler, men denne forbedring i prædiktionen kunne typisk ikke findes i testsæt-prædiktionen. Dette indikerer formentligt, at der ikke foreligger tilstrækkeligt datamateriale til at vurdere de mere komplekse modeller.
- Funktionel PCA kan anbefales som en god strategi, såfremt det ønskes at implementere satellitdata i modeller, da denne metode kan håndtere og delvist kompensere for at data kan være meget sparsomme pga. skydække.
- Det understreges, at de foreliggende dataanalyser bygger på en række eksklusionskriterier, dataudvælgelseskriterier og andre modelrelaterede valg, og at det på nuværende tidspunkt ikke er undersøgt, om modellerne kan forbedres ved ændringer i nogle af de anvendte kriterier og trufne valg.

Metoder

Datasæt og datahåndtering

Markdata

- Markpolygondata blev hentet fra landbrugsstyrelsen, www.lbst.dk
- Udbytte og tørstof og andre markrelaterede data (markdata) blev leveret af SEGES i form af et excelark. Derudover blev der leveret et excelark med sortsinformation.
- **Dataeksklusion**:
    - Observationer med høstdatoer tidligere end 61 dage før 1. oktober og 1.juni for hhv. majs og græs er ekskluderet.
    - Majsdata for marker med flere høsttidspunkter blev ekskluderet.
- Minimum antal observationer indenfor **sort** (varietyName) er sat til 20, ellers er sorten betegnet som "Ukendt". Derudover blev der generet en variabel for registrerede sorter (reg_ varietyName) ved at kombinere varietyName med 0/1-variablen registered.
- For alle græsobservationer blev der estimeret et forventet **slæt-nummer** og en forventet **slætstrategi** på følgende måde: forventede fordelinger af slæthøstdatoer for slæt-numre 1 til 5 blev fundet udfra alle marker med et komplet sæt af 5 slæt (max slætnummer = 5). Alle slæt blev tilordnet det slætnummer under hvis fordeling den er mest sandsynligt og under de yderligere betingelser minimum og maksimum antal dage mellem to efterfølgende græsslæt på hhv. 18 og 35 dage. Slætstrategi blev estimeret som hhv. 5-slæt-strategi og 4-slæt-strategi ud fra betingelsen at max antal slæt skal være lig det højst estimerede slæt og ukendt ellers.

Satellitdata

- Satellitdata blev hentet for 2017 og 2018 over hele Danmark fra www.esa.int via procedurer udviklet af Teknologisk Institut og dækker over følgende spektralbånd: B02 til B08, B08A, samt skydække (CLD). Derudover blev der ud fra de tilgængelige bånd beregnet NDVI og NDRE.
- Satellitbilleder er hentet ned fra ESA's Sentinel-2 satellit via Copernicus Open Access Hub. Data er fra Level 2A - Bottom-Of-Atmosphere (BOA).
- Der er 15 "Tiles" (billeder) som tilsammen dækker Danmark. I 2017 var der kun én satellit, som overfløj Europa med ca. 3 dages mellemrum. I 2018 var der to satellitter, som overfløj Europa med 2 dages mellemrum. For 2017 er der satellitbilleder fra 1. april til 1. november. Satellitbilleder før 1. april 2017 har et andet format (Level 1C - Top-Of-Atmosphere) og kræver derfor ekstra behandling for at danne Level 2A.
- Der er trukket 8 bånd (B02, B03, B04, B05, B06, B07, B08, B8A) samt et skydækkeindeks (CLD) ud af satellitbillederne.
- **NDVI**- og **NDRE**-indekserne er beregnet på pixel-niveau som:
    - NDVI = (B08 – B04) / (B08 + B04)
    - NDRE = (B8A – B05) / (B8A + B05)
- Billederne er koblet med oplysninger om markpolygonerne og for hver markpolygon/satellitoverflyvningsdato-kombination er der beregnet gennemsnit og spredning indenfor marken for de 8 bånd (B02.mean, B02.sd, B03.mean, etc.), skydækkeindekset (CLD.mean, CLD.sd), NDVI (NDVI.mean, NDVI.sd) og NDRE (NDRE.mean, NDRE.sd). Derudover blev der beregnet en spline-udglattet skydækkeindeks (CLD_smooth), et NDRE2 indeks, hvor NDRE2 = (B07.mean – B06.mean) / (B07.mean + B06.mean) og ratioen af NDRE over NDVI som NDRE_NDVI_ratio = NDRE.mean / NDVI.mean.
- **Dataeksklusion**:

    - Marker med et antal satellitbillede-pixels svarende til et areal mindre end 0.5 ha (< 32 pixels) blev ekskluderet, se appendiks.
    - Satellitbilleder før 15.05 og 15.04. for hhv. majs og græs.
    - Satellitbilleder med en NDVI.mean < 0.2, se Fig. 1.
    - Satellitbilleder med en CLD.mean > 5%, se Fig. 1.
    - For majs: marker med færre end ét satellitbillede per perioderne "maj-juni", "juli", "august" og såfremt høsten falder senere end 30.09. også "september".
    - For græs vurderet per slæt: marker med færre end 2 satellitbilleder i perioden før det aktuelle slæt.
- Indenfor hver afgrøde blev satellitvariablerne undersøgt for indbyrdes **korrelationer** ved at beregne alle parvise Pearson-korrelationer, se Fig. 2 for majs. Variablerne blev derefter grupperet i grupper af højt-korrelerede variabler vha. en hierarkisk clusteranalyse af Pearson-korrelationerne og fra hver variabelgruppe blev der udvalgt minimum en variabel i det endelige kandidatsæt af satellitvariablerne. Kandidatsættet af satellitvariablerne indeholder følgende variabler: NDVI.mean, NDRE.mean, NDRE2.mean, NDRE_NDVI_ratio, NDVI.sd, NDRE.sd, CLD_smooth, CLD.mean, B05.mean, B8A.sd.
- For hver mark (majs) og mark-slætnummer-kombination (græs) blev der defineret **vækstperioder** over hvilke satellitdata skal evalueres: en vækstperiode er perioden mellem starttidspunkt og høst, hvorved starttidspunktet er 15.05. for majs, 15.04. for græs/første slæt og for alle andre græs-slætnummer-kombinationer slætdato+1 af foregående slæts dato, dog maksimalt 35 dage siden.
- **Funktionel PCA-analyse**: For alle marker (majs) hhv. alle mark-slætnummer-kombinationer (græs) blev der over vækstperioderne kørt en funktionel PCA-analyse for hver kandidatsatellitvariabel. Antal PCA-komponenter *K* blev bestemt som det antal komponenter der tilsammen forklarer ≥95% af variationen i kandidatsatellitvariablens funktioner over vækstperioden. PC-scores for de første *K* PCA-komponenter blev ekstraheret for alle marker hhv. mark-slætnummer-kombinationer og tilføjet datasættet som NDVI.mean_PC1, NDVI.mean_PC2, etc. Funktionel PCA-analyse (fPCA) er velegnet til at opsamle det væsentligste af den tidsmæssige variation i satellitdata, estimerer en mean-funktion over den samlede vækstperiode og kvantificerer de enkelte markers hhv. mark-slætnummer-kombinationers afvigelse fra mean-funktionen. Dermed leverer metoden også en udglatning af den tidsmæssige udvikling i variablen (f.eks. NDVI). Se Fig. 3 for fPCA for majs.

Andre datakilder

- **Klimadata** fra 2017 og 2018 for alle punkter af det danske Kvadratnet blev leveret af SEGES i form af et excelark. Klimavariablerne daglig gns. temperatur (airtemp), daglig akkumuleret nedbør (prec) og solindstråling (glorad) blev for hvert år 2017 og 2018 modelleret i en GAM-model, klimavariabel = s(geoX, geoY), med udglatningsgraden af den todimensionelle spline bestemt ved generaliseret krydsvalidering, for at interpolere til markdataernes lokationer.
- For hver mark (majs) og mark-slætnummer-kombination (græs) blev der beregnet kumulerede klimavariabler over vækstperioderne: cumTemp, cumPrec og cumGlorad.
- Funktionel PCA-analyse: Ligeledes blev der for hver mark (majs) og hver mark-slætnummer-kombination (græs) over vækstperioderne kørt en funktionel PCA-analyse for hver klimavariabel og PC-scores fra de første fire fPCA-komponenter blev ekstraheret for alle marker og tilføjet datasættet (airtemp_PC1, airtemp_PC2, etc.)

- Jordbundsdata (JBnr) blev hentet fra hjemmesiden Den Danske Jordklassificering, http://dca.au.dk/forskning/den-danske-jordklassificering/, som *.shp-fil og for hver mark blev der fundet JBnr, som det JBnr-polygon i shapefilen der indeholder markens koordinater.

Endelig datahåndtering

- Markdata, sortsdata, fPC-satellitdata, fPC-klimadata og JBnr. blev flettet på markniveau for majs og mark-slætNr-niveau for græs.
- Det endelige datasæt bestod af 1.230 majs-observationer (marker) og 2.786 græs-observationer fra i alt 1.457 marker. Tabel 1 viser summary-statistics for datasættet.

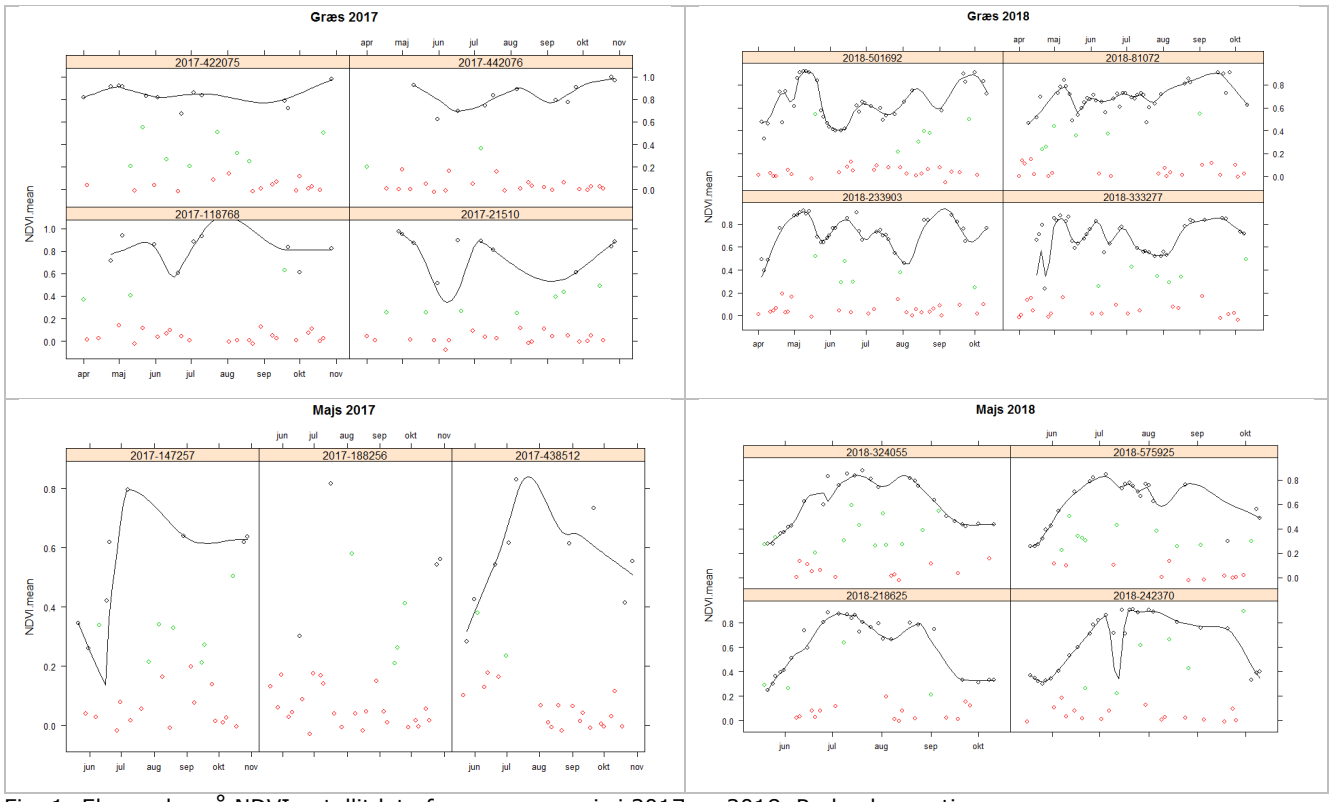


Fig. 1: Eksempler på NDVI-satellitdata for græs og majs i 2017 og 2018. Røde observationer er ekskluderet jf. eksklusionskriteriet NDVI.mean < 0.2. Grønne observationer er ekskluderet jf. eksklusionskriteriet CLD.mean (skydække) > 5%. Sorte linjer er en loess-regression igennem de sorte observationer, der opfylder inklusionskriterier. I græsdata fra 2018 kan de enkelte slæt genkendes; dette er ikke muligt i 2017. Majsdata fra 2017 er også mere sparsomme, men den overordnet forventede NDVI-udvikling kan genkendes, om end med betydelig færre detaljer sammenlignet med 2018.

Tabel 1: Summary-statistics for det endelige majs-datasæt.

```
Majs (N = 1.230)           2017 (N = 453)             2018 (N = 777)

 Areal(ha)-2017   Areal(ha)-2018     Høstdag-2017    Høstdag-2018
 Min.   : 0.960   Min.   : 0.520   Min.   :-16.00   Min.   :-45.0
 1st Qu.: 3.390   1st Qu.: 3.310   1st Qu.:  7.00   1st Qu.:-30.0
 Median : 5.380   Median : 5.800   Median : 13.00   Median :-25.0
 Mean   : 7.024   Mean   : 7.188   Mean   : 11.38   Mean   :-22.3
 3rd Qu.: 9.260   3rd Qu.: 9.430   3rd Qu.: 16.00   3rd Qu.:-14.0
 Max.   :39.350   Max.   :39.790   Max.   : 38.00   Max.   : 24.0
                                      (antal dage efter 01.10)

Ngrowthdays-2017 N_growthdays-2018 TSudbytte(t/ha)-2017 TSudbytte(t/ha)-2018
 Min.   :124.0      Min.   : 95.0      Min.   : 1.258      Min.   : 2.134
 1st Qu.:147.0      1st Qu.:110.0      1st Qu.:10.164      1st Qu.: 8.780
 Median :153.0      Median :115.0      Median :11.707      Median :11.034
 Mean   :151.4      Mean   :117.7      Mean   :11.802      Mean   :10.936
 3rd Qu.:156.0      3rd Qu.:126.0      3rd Qu.:13.505      3rd Qu.:12.923
 Max.   :178.0      Max.   :164.0      Max.   :22.746      Max.   :25.616

 TS%-2017             TS%-2018          Forfrugt-2017  Forfrugt-2018
 Min.   :25.70        Min.   :25.01        andet: 13      andet: 36
 1st Qu.:31.24        1st Qu.:31.00        græs : 74      græs :110
 Median :33.56        Median :33.11        korn : 71      korn :161
 Mean   :33.57        Mean   :33.29        majs :295      majs :470
 3rd Qu.:36.00        3rd Qu.:35.59
 Max.   :39.93        Max.   :39.96

 Økologi-2017  Økologi-2018 cumGlorad-2017   cumGlorad-2018    cumTemp-2017
 Nej:446       Nej:758       Min.   :1879     Min.   :1886     Min.   :1866
 Ja :  7       Ja : 19       1st Qu.:1989     1st Qu.:2074     1st Qu.:2099
                             Median :2054     Median :2149     Median :2175
                             Mean   :2051     Mean   :2143     Mean   :2166
                             3rd Qu.:2103     3rd Qu.:2218     3rd Qu.:2225
                             Max.   :2224     Max.   :2464     Max.   :2391

 cumTemp-2018     cumPrec-2017    cumPrec-2018      Sort-2017
 Min.   :1687    Min.   :348.3   Min.   :101.4   Ukendt      :371
 1st Qu.:1891    1st Qu.:494.0   1st Qu.:186.6   Emblem      : 22
 Median :1963    Median :549.2   Median :216.1   Augustus KWS: 17
 Mean   :2003    Mean   :547.8   Mean   :238.7   Ambition    :  8
 3rd Qu.:2120    3rd Qu.:603.5   3rd Qu.:291.7   Wizard      :  7
 Max.   :2572    Max.   :692.5   Max.   :393.7   Emmerson    :  6
                                                 (Other)     : 22
Sort-2018         JBnr-2017  JBnr-2018  Satellitdage-2017  Satellitdage-2018
Ukendt      :639    1-3:435    1-3:722  Min.   : 2.000     Min.   :11.0
Atrium      : 28    4-6: 18    4-6: 55  1st Qu.: 6.000     1st Qu.:21.0
Autens KWS  : 20                       Median : 7.000     Median :22.0
Emblem      : 20                       Mean   : 7.049     Mean   :22.2
Augustus KWS: 18                       3rd Qu.: 8.000     3rd Qu.:24.0
Ambition    : 17                       Max.   :15.000     Max.   :35.0
(Other)     : 35                       (antal dage med satellitbillede,
                                       der opfylder krav om CLD og NDVI)
```

Tabel 2: Summary-statistics for det endelige græs-datasæt.

```
 Græs (N = 2.786)           2017 (N = 1.174)            2018 (N = 1.612)

SlætNr-2017   SlætNr-2018    Slætstrategi-2017  Slætstrategi-2018
 1:428            1:561            4     :168        4     :  38
 2:253            2:364            5     :152        5     : 142
 3:302            3:185            Ukendt:854        Ukendt:1432
 4: 71            4:251
 5:120            5:251
    (estimeret slætnr.)

Areal (ha)-2017  Areal (ha)-2018     Høstdag-2017     Høstdag-2018
 Min.   : 0.710   Min.   : 0.600   Min.   :-24.00   Min.   :-19.00
 1st Qu.: 3.583   1st Qu.: 3.567   1st Qu.: -7.00   1st Qu.: -9.00
 Median : 5.835   Median : 5.975   Median : 20.00   Median : 14.00
 Mean   : 7.514   Mean   : 7.687   Mean   : 30.70   Mean   : 40.63
 3rd Qu.: 9.820   3rd Qu.: 9.850   3rd Qu.: 54.75   3rd Qu.: 96.00
 Max.   :42.430   Max.   :39.340   Max.   :160.00   Max.   :153.00
                                      (antal dage efter 01.06)

Ngrowthdays-2017 Ngrowthdays-2018 TSudbytte(t/ha)-2017 TSudbytte(t/ha)-2018
 Min.   :20.00      Min.   :19.00      Min.   : 0.3861     Min.   :0.5004
 1st Qu.:31.00      1st Qu.:35.00      1st Qu.: 1.8438     1st Qu.:1.2599
 Median :35.00      Median :35.00      Median : 2.6206     Median :1.9811
 Mean   :33.17      Mean   :33.47      Mean   : 2.9844     Mean   :2.3228
 3rd Qu.:35.00      3rd Qu.:35.00      3rd Qu.: 3.8435     3rd Qu.:3.1107
 Max.   :35.00      Max.   :35.00      Max.   :17.7030     Max.   :9.5641

 TS%-2017             TS%-2018          Forfrugt-2017  Forfrugt-2018
 Min.   :20.90        Min.   :20.06         andet: 70      andet:187
 1st Qu.:30.00        1st Qu.:31.13         græs :636      græs :935
 Median :34.00        Median :35.66         korn :463      korn :484
 Mean   :34.68        Mean   :36.32         majs :  5      majs :  6
 3rd Qu.:39.00        3rd Qu.:40.30
 Max.   :50.00        Max.   :59.80

 Økologi-2017  Økologi-2018 cumGlorad-2017   cumGlorad-2018    cumTemp-2017
 Nej:956       Nej:1223      Min.   :113.3    Min.   :153.0    Min.   :242.0
 Ja :218       Ja : 389      1st Qu.:492.1    1st Qu.:436.6    1st Qu.:330.7
                             Median :555.3    Median :581.0    Median :412.4
                             Mean   :528.6    Mean   :544.2    Mean   :409.6
                             3rd Qu.:594.3    3rd Qu.:662.1    3rd Qu.:484.9
                             Max.   :670.1    Max.   :816.8    Max.   :573.4

 cumTemp-2018    cumPrec-2017     cumPrec-2018       Sort-2017
 Min.   :276.4   Min.   : 27.44   Min.   :  0.3256   Ukendt       :1156
 1st Qu.:422.6   1st Qu.: 52.43   1st Qu.: 31.5222   35 (blanding):   3
 Median :449.4   Median : 61.80   Median : 61.2118   45 (blanding):  15
 Mean   :484.4   Mean   : 73.36   Mean   : 73.2838   47 (blanding):   0
 3rd Qu.:556.5   3rd Qu.: 80.81   3rd Qu.: 89.6301
 Max.   :693.5   Max.   :209.91   Max.   :263.2334
                                                      (kun bekræftede,
                                                      registrerede sorter)

 Sort-2018           JBnr-2017  JBnr-2018 Satellitdage-2017 Satellitdage-2018
 Ukendt       :1582  1-3:1055   1-3:1464  Min.   :2.000     Min.   : 2.000
 35 (blanding):   7  4-6: 119   4-6: 148  1st Qu.:2.000     1st Qu.: 4.000
 45 (blanding):  13                       Median :2.000     Median : 7.000
 47 (blanding):  10                       Mean   :2.524     Mean   : 6.661
                                          3rd Qu.:3.000     3rd Qu.: 9.000
                                          Max.   :6.000     Max.   :17.000
                                          (antal dage med satellitbillede,
                                           der opfylder krav om CLD og NDVI)
```

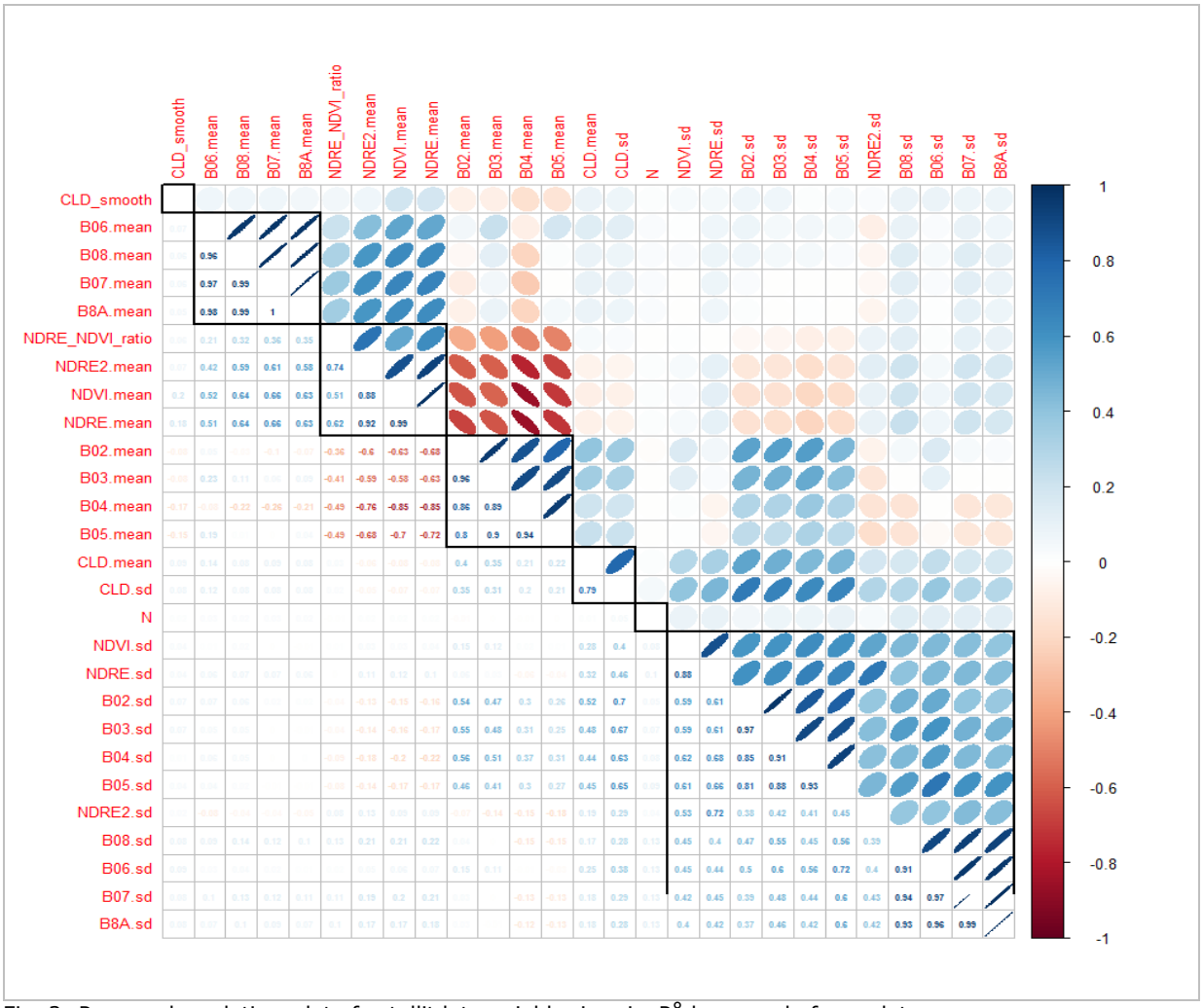
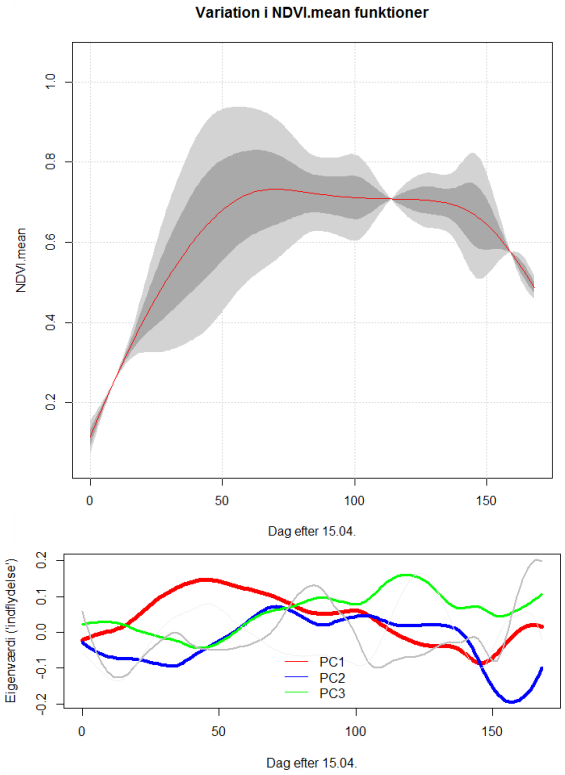


Fig. 2: Pearson korrelationsplot af satellitdatavariabler i majs. På baggrund af complete hierarchical clustering er der identificeret syv grupper af variabler, der er relativ uafhængige af hinanden, men udviser høj korrelation indbyrdes indenfor gruppen. Kandidatvariabler er udvalgt på baggrund af de syv fundne grupper.



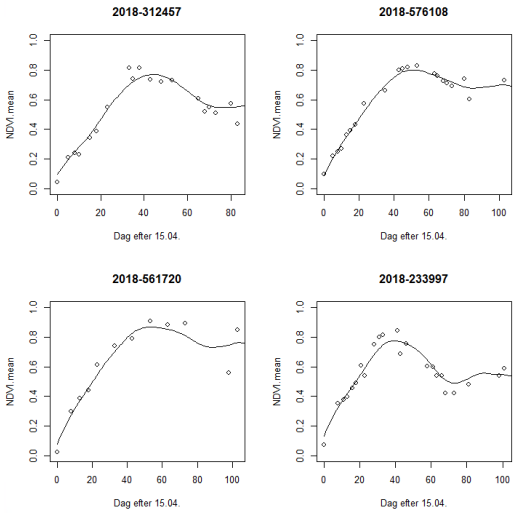


Fig. 3: Funktionel PCA-analyse for majs.

Model

- Der er analyseret en række modeller af stigende kompleksitet fra modeller med kun NDVI hhv. NDRE som input til modeller der indeholder satellitdata og de resterende foreligger data såsom klimadata og markrelaterede data (JBnr, forfrugt, m.m.).
- Der er kørt en randomForest model til identificering af kandidater til vekslevirkninger jf. R-pakken randomForestSRC.
- Modellerne fremgår af tabellerne 3 til 13.
- Alle analyserede modeller er GAM modeller med alle inputvariabler fitted som thin-plate spline funktioner med udglatningsgraden bestemt vha. generaliseret krydsvalidering.
- Modellernes performance blev evalueret vha. en 10-fold 2-times repeated cross validation på testdata.

Resultater

- Resultater fremgår af tabellerne 3 til 9. Ikke alle modeller kunne køres pga. for få data; disse er udeladt fra tabellerne.

Tabel 3: Satellitdata-baseret **TS-udbytte**-prædiktion i **majs** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo., NDVI	0.38	2.0	2.5±0.2
A2	geo. + NDRE	geo., NDRE	0.45	1.9	2.4±0.2
A3	geo. + NDVI + NDRE	geo., NDRE, NDRE/NDVI	0.49	1.8	2.4±0.2
A4	geo. + sat.	geo., NDRE, NDRE/NDVI, CLD	0.53	1.7	2.3±0.3
B1	geo. + sat. + dyrkn.	geo., NDRE, NDRE/NDVI, CLD, forfrugt, JBnr	0.52	1.7	2.3±0.3
B3	geo. + sat. + dyrkn. + klima + vekselvirkninger	geo., NDRE, NDRE/NDVI, CLD, forfrugt, glorad, airtemp, prec	0.56	1.7	2.3±0.3

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 4: Satellitdata-baseret **TS%**-prædiktion i **majs** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo., NDVI	0.28	2.6	2.8±0.2
A2	geo. + NDRE	geo., NDRE	0.24	2.7	2.8±0.1
A3	geo. + NDVI + NDRE	geo., NDRE, NDRE/NDVI	0.30	2.6	2.7±0.3
A4	geo. + sat.	geo., NDRE, NDRE/NDVI, CLD, B05	0.30	2.6	2.7±0.1
B1	geo. + sat. + dyrkn.	geo., NDRE, NDRE/NDVI, CLD, B05, høstdag, Ngrowthdays, økologi, sort	0.32	2.6	2.8±0.1
B3	geo. + sat. + dyrkn. + klima + vekselvirkninger	geo., NDRE, NDRE/NDVI, B05, glorad, airtemp, prec	0.62	1.9	2.3±0.1

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 5: Satellitdata-baseret **TS-udbytte**-prædiktion i **græs-1.slæt** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo.	0.15	1.3	1.4±0.2
A2	geo. + NDRE	geo.	0.15	1.3	1.4±0.3
A3	geo. + NDVI + NDRE	geo.	0.15	1.3	1.4±0.2
A4	geo. + sat.	geo., CLD	0.20	1.2	1.4±0.3
B2	geo. + sat. + dyrkn. + klima	geo., CLD, airtemp, glorad, prec	0.53	0.9	1.7±0.3

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 6: Satellitdata-baseret **TS-udbytte**-prædiktion i **græs-2.slæt** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo., NDVI	0.29	1.0	1.1±0.2
A2	geo. + NDRE	geo.,	0.24	1.1	1.1±0.2
A3	geo. + NDVI + NDRE	geo., NDRE, NDRE/NDVI	0.31	1.0	1.2±0.2
A4	geo. + sat.	geo., NDRE, NDRE/NDVI, CLD, B05	0.56	0.8	1.1±0.1
B1	geo. + sat. + dyrkn.	geo., NDRE, NDRE/NDVI, CLD, økologi, slætstrategi, høstdag + Ngrowthdays	0.67	0.7	0.9±0.3
B2	geo. + sat. + dyrkn. + klima	geo., NDRE, NDRE/NDVI, glorad, airtemp, prec	0.78	0.6	0.8±0.2
B3	geo. + sat. + dyrkn. + klima + vekselvirkninger	geo., NDRE, NDRE/NDVI, CLD, høstdag, glorad, airtemp, prec	0.78	0.6	0.8±0.2

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 7: Satellitdata-baseret **TS-udbytte**-prædiktion i **græs-3.slæt** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo.	0.30	0.8	0.9±0.1
A2	geo. + NDRE	geo.	0.31	0.8	0.9±0.1
A3	geo. + NDVI + NDRE	geo.	0.31	0.8	0.9±0.1
A4	geo. + sat.	geo., NDRE, NDVI, CLD, B05	0.39	0.7	0.9±0.2
B1	geo. + sat. + dyrkn.	geo., NDRE, NDVI, CLD, B05, økologi, høstdag, Ngrowthdays	0.48	0.7	0.8±0.1
B2	geo. + sat. + dyrkn. + klima	geo., NDRE, NDVI, CLD, B05, glorad, airtemp, prec	0.58	0.7	0.9±0.2
B3	geo. + sat. + dyrkn. + klima + vekselvirkninger	geo., NDRE, NDVI, CLD, B05, glorad, airtemp, prec	0.57	0.7	0.9±0.1

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 8: Satellitdata-baseret **TS-udbytte**-prædiktion i **græs-4.slæt** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo., NDVI	0.30	0.9	1±0.3
A2	geo. + NDRE	geo., NDRE	0.27	0.9	1±0.3
A3	geo. + NDVI + NDRE	geo., NDRE	0.27	0.9	1±0.3

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.

Tabel 9: Satellitdata-baseret **TS-udbytte**-prædiktion i **græs-5.slæt** for modeller af stigende kompleksitet. Prædiktionsperformance er angivet for både trainingsdata og testdata (10-fold 2-times repeated cross validation).

Model		Signifikante termer[1]	Training-sæt		Test-sæt
			R^2	RMSE (t/ha)	RMSEP (t/ha)
A1	geo. + NDVI	geo.	0.17	1.0	1.3±0.7
A2	geo. + NDRE	geo.	0.18	1.0	1.3±0.7
A3	geo. + NDVI + NDRE	geo., NDRE, NDRE/NDVI	0.19	1.0	1.4±0.7

[1]: angiver overordnet de termer, hvis features, f.eks. mean, sd, fPC1, m.m., og/eller interaktioner indgår signifikant i modellen.