Udbytteprognose i vinterhvede

Af Mette Kramer Langgaard,
Specialkonsulent, SEGES Innovation

3. Oktober 2022

STØTTET AF
Promilleafgiftsfonden for landbrug

SEGES INNOVATION

Agenda

- Udbytteprognosen i CropManager
- Hvorfor arbejder vi med at forudsige udbytte?
- Grundlaget for nuværende model i CropManager
- Forskning på området
- Videreudvikling af udbytteprognosen

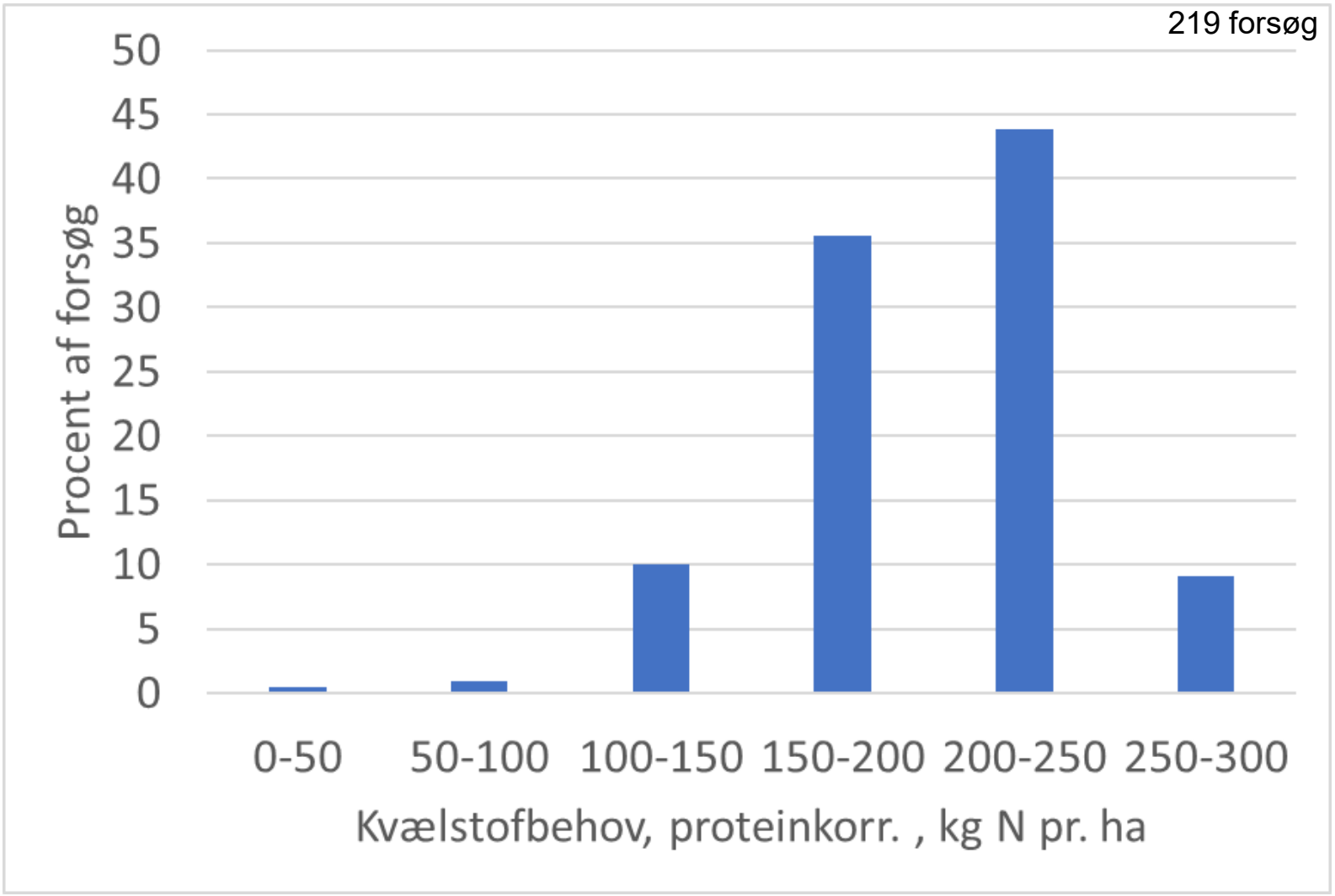
SEGES INNOVATION

Udbytteprognose i CropManager

SEGES INNOVATION

Hvorfor er udbytteforudsigelse overhovedet interessant?

Variation i kvælstofbehov i vinterhvede (2011-2020)



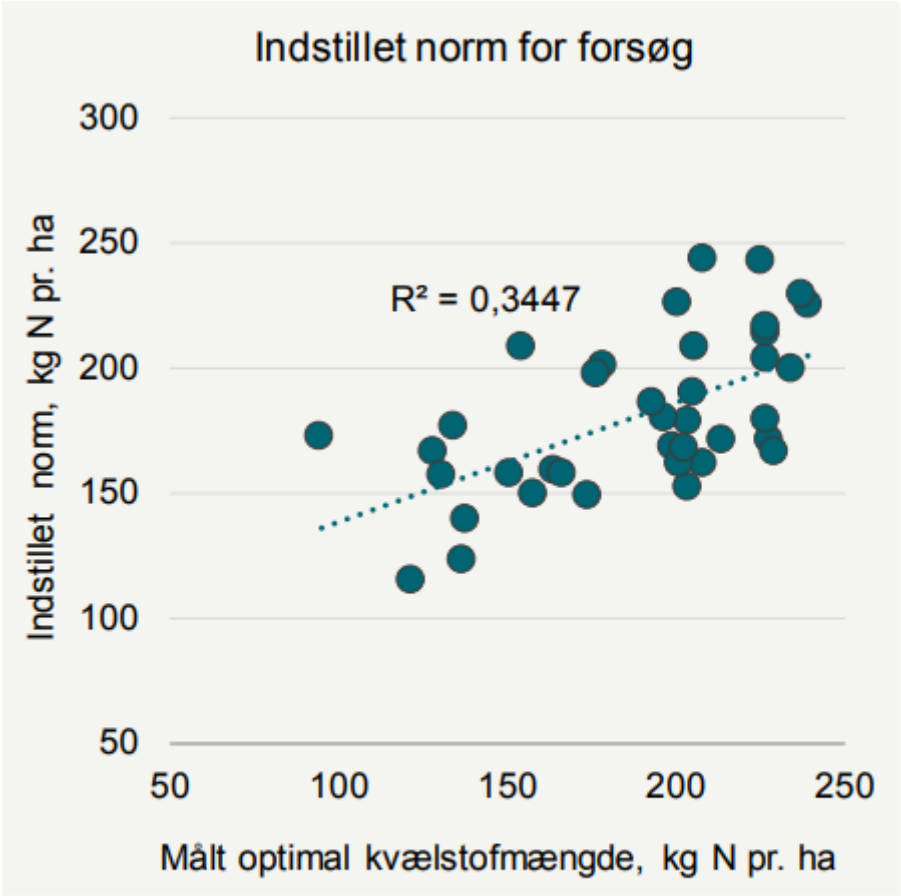
Nuværende model for bestemmelse af N-behov:

- Forfrugt
- Jordtype
- Forventet udbytte
- Dyrkningshistorie
- N-prognose

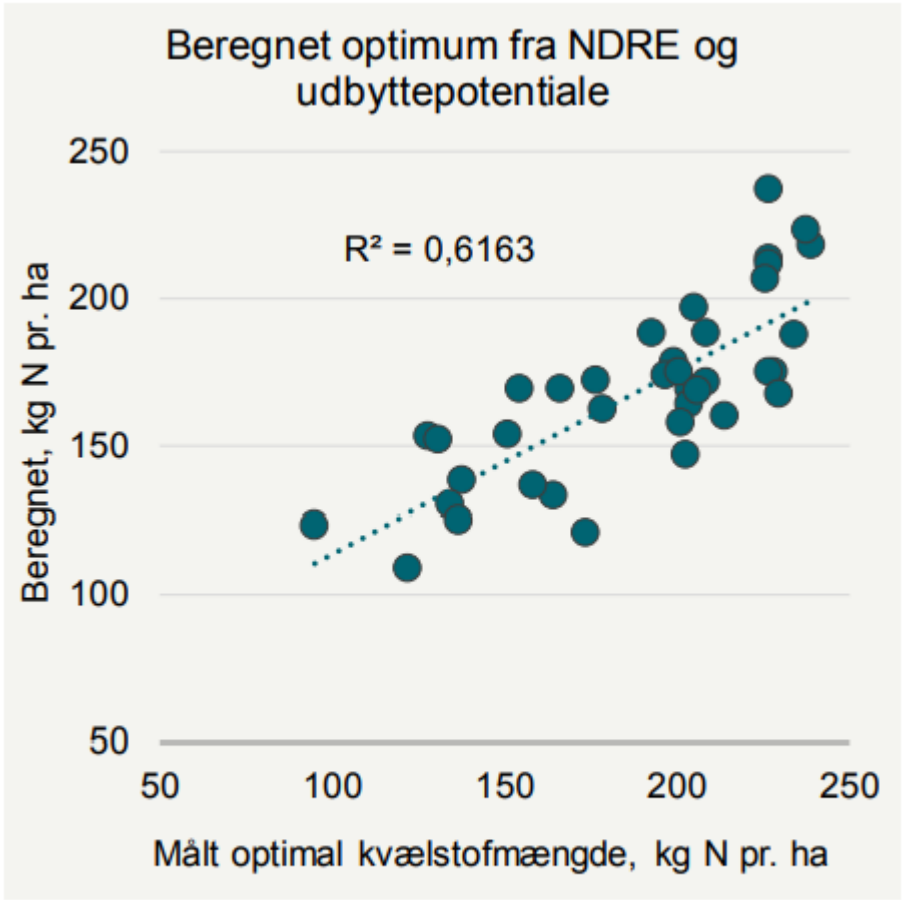
Anvendes i MarkOnline

SEGES INNOVATION

Optimal N ud fra udbytte og NDRE



FIGUR 19. Sammenhæng mellem kvælstofbehov ved beregningsmetoden til indstilling af kvælstofnormer og det målte kvælstofbehov.



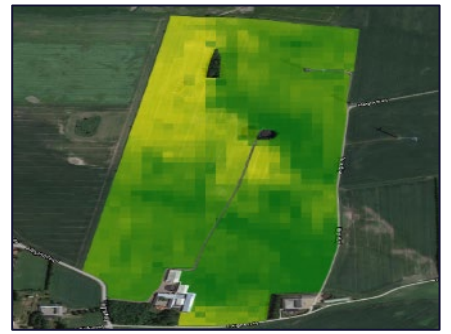
FIGUR 20. Sammenhæng mellem kvælstofbehov beregnet ud fra NDRE og udbytte og det målte kvælstofbehov.

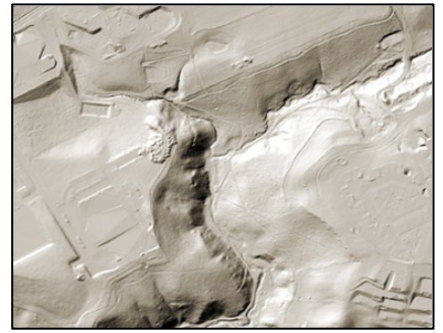
Kilde: Oversigt over landsforsøgene 2020, side 206

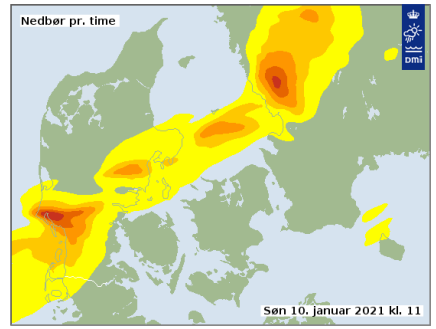
Datagrundlag for nuværende model i CropManager

Data

- Udbyttekort fra mejetærskere
- Satellitdata (L1C Sentinel data)
- Terrænhøjdedata (den Danske højdemodel)
- Vejrdata (DMI)



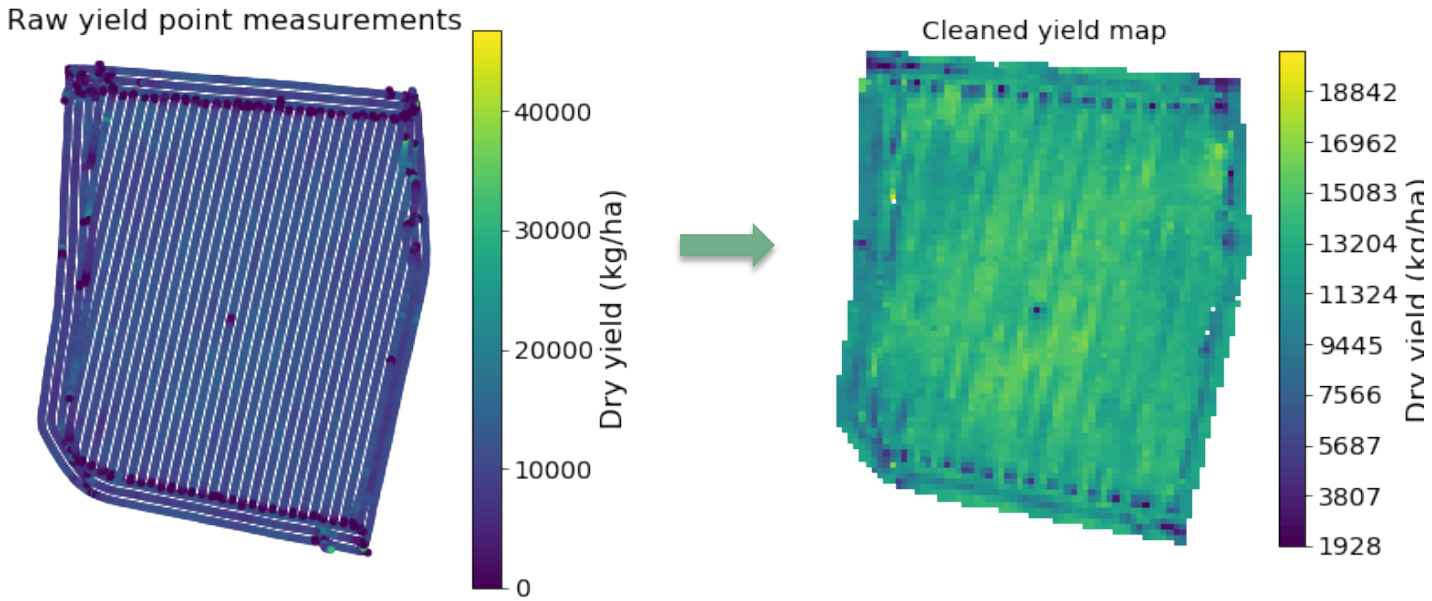




SEGES INNOVATION

Udbyttedata fra mejetærskeren

- Udbyttekort fra 14 bedrifter indsamlet i 2018
- Udbyttedata fra 2003-2017
- Satellitdata tilgængelig fra 2016 → vinterhvedemarker fra 2016 og 2017 udvalgt og oprenset
- Udbyttedata fra 106 vinterhvedemarker svarende til 1.125 ha.
- Oprensning af data



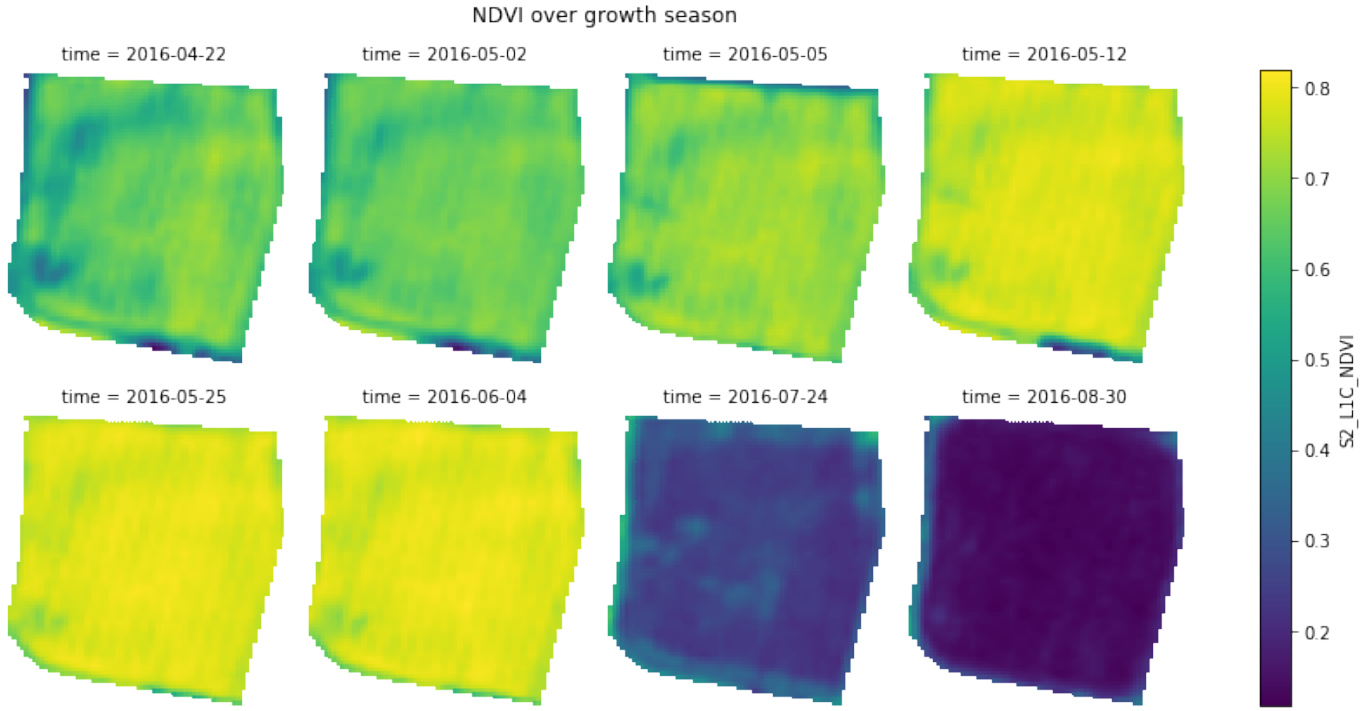
SEGES INNOVATION

Satellitdata Sentinel-2 L1C

fra 9 marts – 6 september i 2016 og 2017

- B01
- B02
- B03
- B04
- B05
- B06
- B07
- B08
- B08a
- B09
- B10
- B11
- B12
- NDVI
- NDRE
- MSAVI2

- Akkumuleret gennemsnit
- Integralet over tid
- Biomasse maks. og biomasse 14 og 30 dage efter

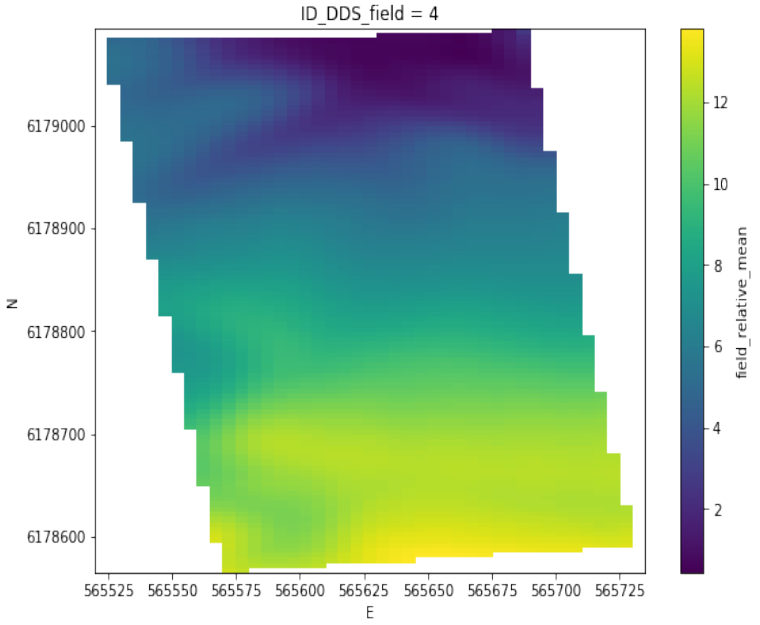


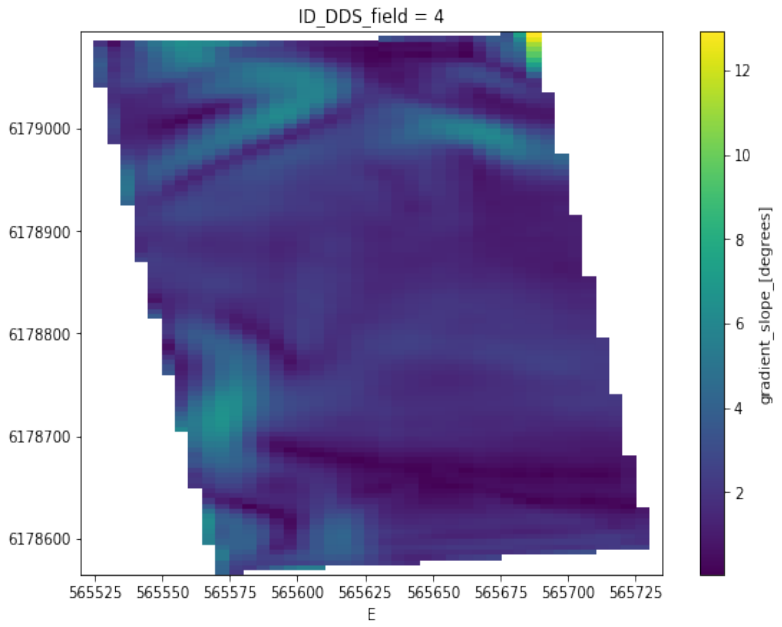
SEGES INNOVATION

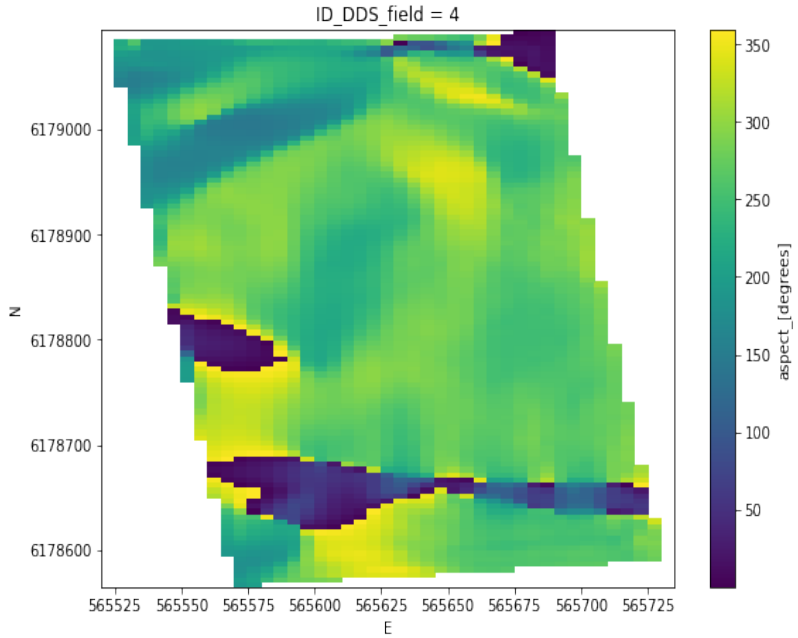
Terrænhøjde data

Den danske højde model (opløsning på 0,4 meter)

- Relativ højde i forhold til det laveste punkt i marken
- Hældning (0-90°)
- Orientering af hældning (0° = nord, 90 °= øst, 180° = syd og 270° = vest)







Vejrdata

- Hentet fra DMI (opløsning på 10 x 10 km) og summeret hver 14 dag.
  - Luft- og jordtemperatur
  - Nedbør
  - Fordampning
  - Indstråling

Følgende er afledte er beregnet: gennemsnit, standardafvigelse (SD), minimum, maksimum + akkumuleret værdier for nedbør og indstråling.

SEGES INNOVATION

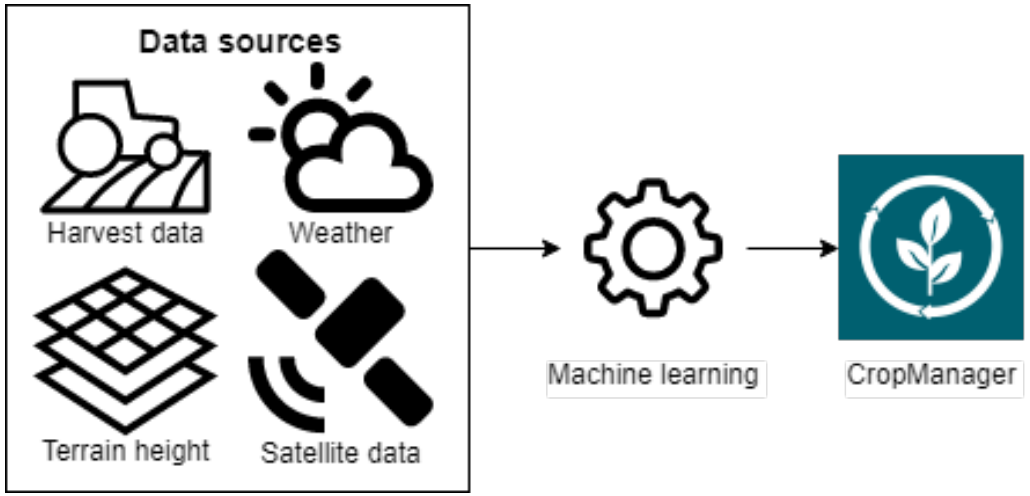
Datagrundlag og ML model

Ca. 450.000 observationer

- Træningsdata: ~**350.000** 5x5 meter pixels.
- Valideringsdata: ~**100.000** 5x5 meter pixels.

1343 variabler i modellen

ML algoritmen Gradient Boosting

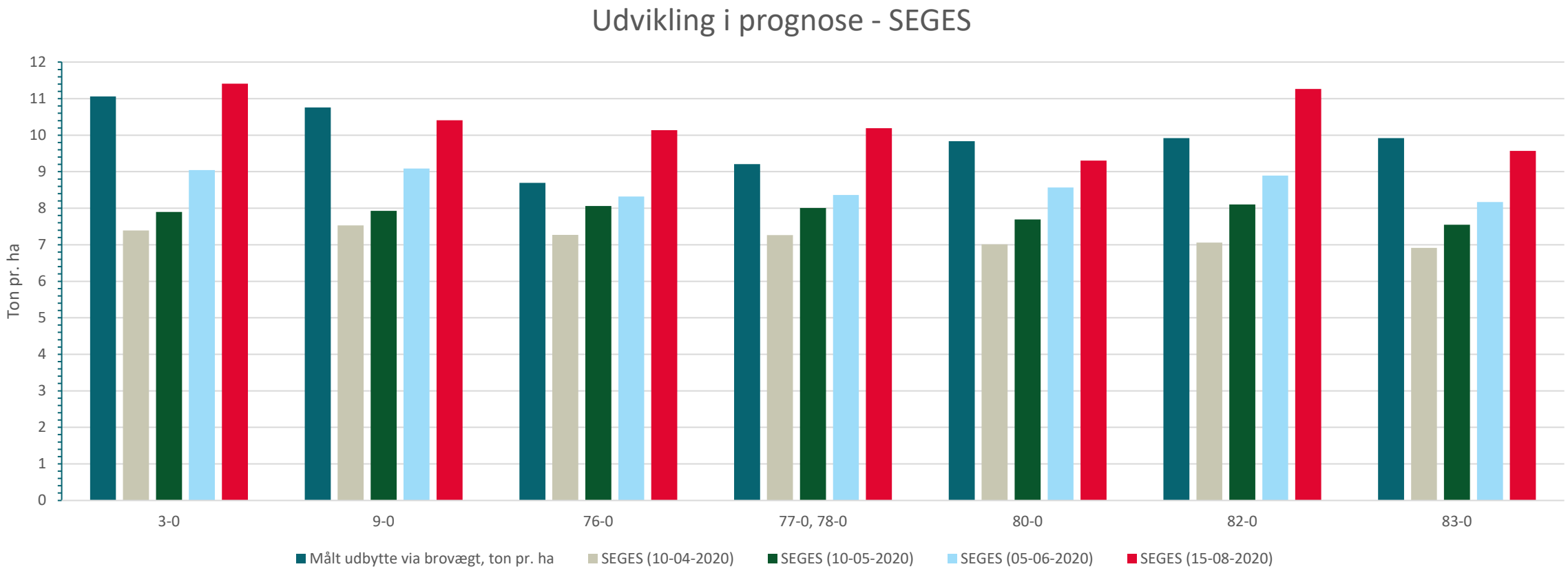


Resultater

Hvor godt rammer prognosen i forhold til det målte udbytte?

Forudsigelses dato	Alt data inkluderet (Satellit-, højde- og vejrdata)	Kun satellitdata inkluderet	
	Prognose nøjagtighed, t/ha (5x5 meter)	Prognose nøjagtighed, t/ha (5x5 meter)	Prognose nøjagtighed, t/ha (markniveau)
10. April	1,469	1,7296	0,8357
10. Maj	1,441	1,7144	0,9263
15. Juni	1,273	1,5869	0,8046
6. September	1,217	1,4252	0,6694

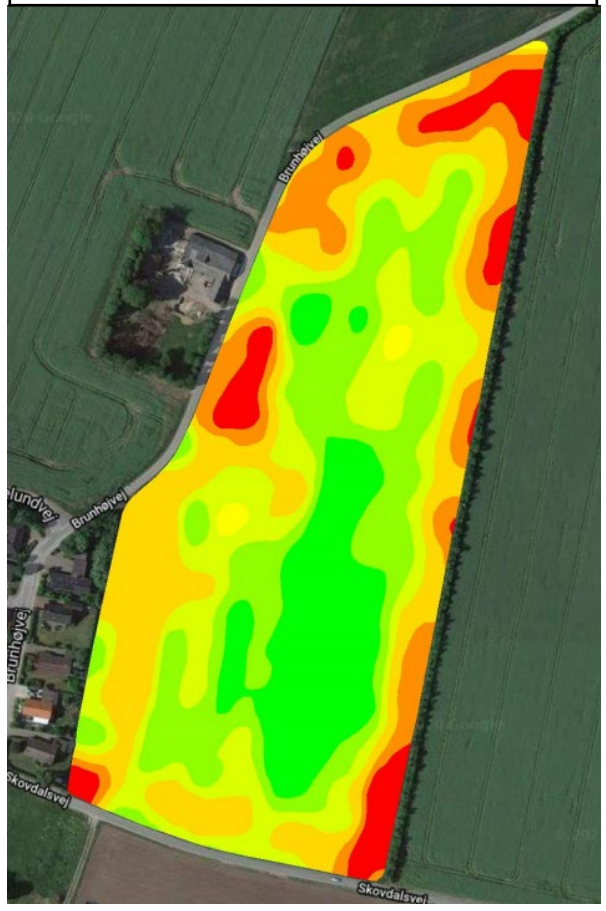
Test af udbytteprognosen mod målt udbytter

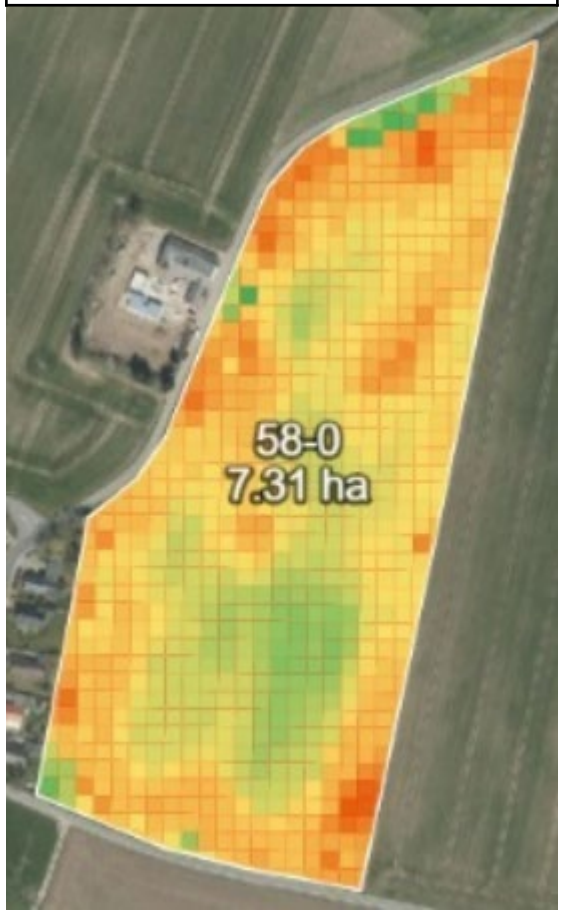


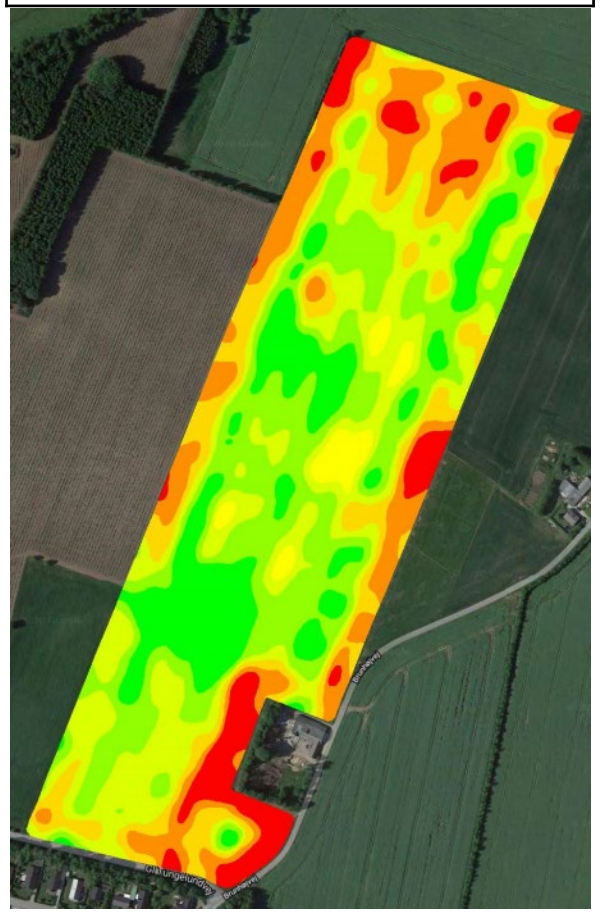
SEGES INNOVATION

Feedback på modellen i CropManager

Mejetærsker	Udbytteprognose	Mejetærsker	Udbytteprognose







Forskning på området

Yield prediction in spring barley from spectral reflectance and weather data using machine learning (ikke publiceret endnu)

af Carsten T. Petersen , Mette Kramer Langgaard, Søren D. Petersen 2022

Vårbyg fra 2010-17 + 2020 (JB6, drænet)

Behandlinger (4 gentagelser):

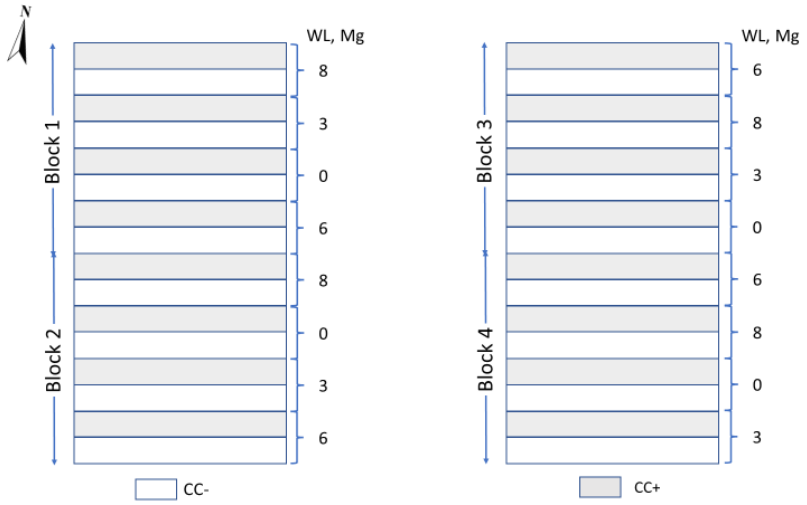
- Jordpakning (hjuloverkørsel med 0, 3, 6 og 8 ton) før såning af vårbyg (WL)
- 2013-2016: Olieræddike efterafgrøde + 30 kg N pr. ha (CC+) eller uden (CC-)

Forsøgsdesign:

- Randomiseret blok design (blok 1-4)

Udbyttevariation:

- 4,2-9,3 ton pr. ha



Observationer: 4(år;2010-13) x 4(WL) x 4(Blok) + 5(år; 2014-17,2020) x 4(WL) x 2(CC) x 4(Blok) = 224

SEGES INNOVATION

Resultater (upubliceret)

- RVI (RVI_{max}, $SL_{max}SL_{min}$ og Int)
- Akkumuleret nedbør i april, maj og juni (P_4, P_5 og $P_{6,}$ P_{30})
- Akkumuleret evapotranspiration (ET_4, $ET_{5,}$ ET_6 og ET_{30})
- Hjul vægt (WL)
- Efterafgrøde (CC)

Kilde: Oversigten over landsforsøg 2013

Procedure no.	Number of foldings	Performance of AML-model		Performance of GLM-model	
		MAE	SDE	MAE	SDE
		Mg/ha			
1	10	0.33	0.04	0.63	0.13
2	10	0.22	0.05	0.49	0.10
3	10	0.24	0.05	0.45	0.06
4	4	0.25	0.02	0.44	0.02
5	9	0.38	0.17	0.93	0.35

Tabel. Mean absolute error (MAE) for the overall best cross-validated AML (Automatic Machine Learning) and GLM (Generalized Linear Model) models in procedures 1-5, and standard deviation (SDE) between MAEs obtained with the different foldings of data used when developing the models (see Table 2).

SEGES INNOVATION

Eksperimenter

Int is the difference between RVI of the plot and a dry soil surface before crop emergence integrated daily between May 1 and the date when predictions were made

RVI_{max} is the maximum RVI observed in the plot to this date

SL_{max} and SL_{min} are the maximal and minimal slopes of the RVI vs. time curve, respectively (day^{-1}) seen so far

P_4 P_5 and P_6 accumulated precipitation during April, May, June

ET4, ET5, ET6 = accumulated reference evapotranspiration

P_{30}, ET_{30}: the most recent 30 days.

WL (Wheel Load; four levels))

CC (Catch crop; two levels);

Block (experimental Block; four levels).

Procedure no.	Independent variables* / Dates within years for modeling or yield prediction/ Investigated years	Distribution into training and test data sets
1 (vejrparameter er ikke inkluderet)	Int, RVI_{max}, SL_{max}, SL_{min}, WL, CC, Block / June 30 / All years	Ninety percent of data for training, 10 percent for testing (ten foldings). Equal representation (as far as possible) of treatments, blocks, and years in the two data sets, otherwise random selection
2 (alle variable)	Int, RVI_{max}, SL_{max}, SL_{min}, WL, CC, Block, P_4, P_5, P_6, ET_4, ET_5, ET_6 / June 30 / All years	
3 (se næste figure)	Int, RVI_{max}, P_6, ET_6 / June 30 /All years	
4 (Blok)		Data from three blocks (seventy-five percent of data) were used for training, and data from the fourth block (twenty-five percent of data) were used for testing (four foldings). Equal representation (as far as possible) of treatments and years in the two data sets, otherwise random selection.
5 (år)		Data from 8 out of 9 years for training, last year used for testing (nine foldings). Equal representation (as far as possible) of treatments and blocks in the two data sets, otherwise random selection
6 (10-30 juni)	Int, RVI_{max}, P_{30}, ET_{30} / All dates between June 10 and June 30/ All years	
7 (antal år)	Int, RVI_{max}, P_6, ET_6 / June 30/ Variable sequences of years	Data from all single years in sequence were used once for testing, and the other data (one to eight years) were used for training (variable number of foldings between 2 and 9). Sequences starting in 2010 (moving forward in time) and in 2020 (moving backward), respectively. Equal representation (as far as possible) of treatments and blocks in the two data sets, otherwise random selection

Videreudvikling af prognosen

Videreudvikling på modellen

- Større datagrundlag
- Forbedre datagrundlag (oprensning)
- Nye parametre
  - Sort
  - Aktuel fordampning, potentiel fordampning, vandbalanceunderskud
  - JB
  - Sædskifte
- Feedback system – udbytter fra grovvaren

År	Udbyttedata	
	Antal marker	Hektar
2021	171	3.611
2020	41	438
2019	51	619
2018	51	647
2017	64	826
2016	21	190
Total	**400**	**6.352**

SEGES INNOVATION


Tak for opmærksomheden