

Promilleafgiftsfonden for landbrug

Notat: N-optag bestemt ud fra UAV-spektraldata - forsøg 070011919 og 070301818

23-10-2019

Philipp Trénel, ph.d., seniorkonsulent
phtr@teknologisk.dk, +45 72 20 33 92

Formål

- At undersøge, hvilke drone-indhentede spektralbånd (R, G, B, NIR og RedEdge) og heraf afledte vegetationsindekser udviser den bedste prædiktionssevne/sammenhæng til planteindeholdt kvælstof (N) målt i kgN/ha i vinterhvede i forsøgsserierne 070011919 og 070301818.
- Med særlig fokus på at undersøge, om alternative bølgelængder giver yderligere informationer om afgrødeudvikling og tendenser, end de bølgelængder som anvendes i dag til indekserne NDVI og NDRE.

Overordnede konklusioner

- Analysen udpeger ikke én VI som bedst, men viser en mere gradvis rangering af VI'er og denne rangering afhænger af kompleksiteten af den anvendte model, selvom prædiktionssevnen kun ringe er påvirket af at øge modelkompleksiteten ved at inddrage information om lokalitet eller lokalitet plus gødskningsstrategi, se Tabel 1.
- De fire højest-rangerende VI med bedste prædiktionsperformance i en model kun baseret på vegetationsindeks (model 1) er **GRVI**, **NDTI**, **NDRE_{sd}** og **MCARI**. Prædiktionssevnen er mellem 11.1 og 12.2% (CVP) svarende til 7.1 til 7.8 kgN/ha (RMSEP).
- Det er bemærkelsesværdigt, at både GRVI og NDTI er rent RGB-baserede indekser. NDRE_{sd} beskriver indenfor-parcel-variationen i NDRE, og det er umiddelbart ikke nemt at forklare, hvorfor prædiktionssevnen er så god for denne indeks. **MCARI** er den spektraldata-baserede indeks med bedste prædiktionsperformance (CVP = 12.2%, RMSEP = 7.8 kgN/ha).
- Prædiktionssevnen er relativt uafhængig af lokalitet (sammenligning af model 1 og model 2, f.eks. MCARI $RMSEP_{model1} = 7.8 \pm 0.5$ og $RMSEP_{model2} = 8.2 \pm 1.2$) og ligeledes uafhængig af lokalitet og gødskningsstrategi (sammenligning med model 3, f.eks. MCARI $RMSEP_{model3} = 8.1 \pm 2.4$).
- De tre bedste VI'er i hhv. model 2 og 3 er **redEdge_{mean}**, **MCARI** og **NDRE** i model 2 og **NDVI**, **G_{sd}**, og **TCARI/OSAVI**-ratioen i model 3.
- Det er bemærkelsesværdigt, at forventningen om en bedre testset-prædiktions sammenlignet med træningsset-prædiktions (sammenligning af RMSE og RMSEP) ikke er indfriet for model 1. Dette indikerer, at en prædiktions baseret på f.eks. **MCARI** forventes at være relativ robust og uafhængig af lokalitet og gødskningsstrategi.

Metoder

Forsøgsdesign

- Forsøgsbehandlinger, måleparametre og forsøgsdesign er beskrevet i Nordic Field Trial System her:
 - 070011919 (3 forsøg), se [her](#).
 - 070301818 (6 forsøg), se [her](#).
- Yderligere informationer fremgår af resultatsiden i Nordic Field Trial System under de enkelte forsøg.
- Data anvendt i nuværende kontekst består af:
 - Vækststadier (t); vækststadier er anvendt som tids-variabel t og afspejler biologisk tid. Vækststadiet-information er hentet fra NFTS og/eller leveret af Mette Kramer Langgard, SEGES, se bilag 1.
 - kgN/ha (y) til tidspunkt t . Arealer for planteklip fremgår af bilag 1.

- forsøgs-id ($trial$) bestående af PLANR og LBNR hentet fra NFTS.
- Behandling (trt) som kombination af tildelte mængder kvælstof (N)
- Dronedata til vækststadiet t bestående af spektralbåndene R, G, B, NIR og RedEdge. For hvert spektralbånd og parcel er der beregnet parcelvis gennemsnit ($mean$) og standardafvigelse (sd) af pixelværdier indenfor parcellen. Ud fra de parcelvise spektralbåndgennemsnit er der beregnet publicerede vegetationsindekser (VI), se tabel 2. Spektralbåndene ($mean$ og sd) og VI tilsammen er de 61 forklarende variabler x analyseret her. Vi kalder dem her også tilsammen for VI.

Model

- Der er anvendt functional data analysis (FDA, Kokoszka & Reimherr 2017) til at analysere en function-on-function regression af responsfunktionen $\text{kgN/ha}(t)$ på den i 'te VI-funktion, $x_i(t)$ repræsenteret som deres functional principal components scores. Fordelen ved anvendelse af functional principal components (fPCA, se Fig. 5) er, at den tillader at modellere data observeret på et uregelmæssigt tids-grid og med manglende dataobservationer, som tilfældet her.
- Der er brugt følgende tre modeller med stigende kompleksitet:

- (1) $\log(\text{kgN/ha})(t) \sim \beta_0 + FPC(x_i(t)) + e(t)$
- (2) $\log(\text{kgN/ha})(t) \sim \beta_0 + b_{\text{trial}} + FPC(x_i(t)) + e(t)$
- (3) $\log(\text{kgN/ha})(t) \sim \beta_0 + b_{\text{trial}} + \beta_{1\text{trt}} + FPC(x_i(t)) + e(t)$

hvor $b \sim N(0, \sigma^2_{\text{trial}})$ (dvs. tilfældig effekt for trial) og $e \sim N(0, \sigma^2_e)$ og responsvariablen er modelleret på log-transformeret skala.

- Model 1 ser bort fra information vedr. lokalitet og forsøgsbehandling og beskriver kun sammenhængen mellem planteindeholdt N-mængde og dronedata. Model 2 og 3 undersøger, hvor stor indflydelse viden om lokalitet og lokalitet + forsøgsbehandling har på evnen med at prædiktere kgN/ha ud fra spektraldata.
- Alle $i = 1, \dots, N = 63$ VI'er er analyseret jf. modellerne.
- Modelperformance er vurderet på trainingsdatasettet vha. forklaringsgraden R^2 , RMSE på tilbagetransformeret skala for gns. N-mængde og CV (Coefficient of Variance; angivet i % af gns. N-mængde) og på testdata genereret vha. 10-times repeated 5% random holdout sampling vha. RMSEP på tilbagetransformeret skala for gns. N-mængde og CVP (Coefficient of Variance for Prediction).
- Der er gennemført standard model kontrol.
- Alle analyser er gennemført i R, særligt med pakke `refund` (Goldsmith et al. 2018)

Resultater

- Resultater fremgår af Tabel 1 og Fig. 1 til 4.

Forslag til videre analyser

- 1) At undersøge, om prædiktionen kan forbedres ved at vægte de enkelte VI'ernes prædiktion med vægtningen værende en funktion af vækststadiene t .
- 2) At gentage analysen som en scalar-on-function FDA til prædiktion af N-udbytte.

Tabel 2: De 61 anvendte vegetationsindekser (VI) inkl. spektralbånd. Der er også anvendt rent RGB-baserede VI.

<p>Bands: R_{mean} (R), G_{mean} (G), B_{mean} (B), NIR_{mean} (NIR), $\text{redEdge}_{\text{mean}}$ (redEdge), R_{sd}, G_{sd}, B_{sd}, NIR_{sd}, $\text{redEdge}_{\text{sd}}$</p> <p>Spectral-band-based indices:</p> <ul style="list-style-type: none"> • Normalised Difference Vegetation Index, Rouse 1974 $NDVI = (NIR - R) / (NIR + R)$ • Normalized Difference Red Edge Index $NDRE = (NIR - \text{redEdge}) / (NIR + \text{redEdge})$ • Corrected Transformed Vegetation Index, Perry 1984 $CTVI = (NDVI + 0.5) / \sqrt{\text{abs}(NDVI + 0.5)}$ • Difference Vegetation Index, Richardson 1977 $DVI = s * NIR - R$, hvor $s = 1$ • Enhanced Vegetation Index, Huete 1999 $EVI = g * ((NIR - R) / (NIR + C_1 * R - C_2 * B + L_{\text{evi}}))$, hvor $g = 2.5$, $C_1 = 6$, $C_2 = 7.5$ og $L_{\text{evi}} = 1$ • Two-band Enhanced Vegetation Index, Jiang 2008 $EVI2 = g * (NIR - R) / (NIR + 2.4 * R + 1)$ • Green Normalised Difference Vegetation Index, Gitelson 1998 $GNDVI = (NIR - G) / (NIR + G)$ • Modified Chlorophyll Absorption Ratio Index, Daugherty 2000 $MCARI = ((\text{redEdge} - R) - (\text{redEdge} - G)) * (\text{redEdge} / R)$ • Modified Soil Adjusted Vegetation Index, Qi 1994 $MSAVI = NIR + 0.5 - (0.5 * \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - (2 * R))})$ • Modified Soil Adjusted Vegetation Index 2, Qi 1994 $MSAVI2 = (2 * (NIR + 1) - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - R)}) / 2$ • Normalised Difference Water Index, McFeeters 1996 $NDWI = (G - NIR) / (G + NIR)$ • Normalized difference spectral index, Zhao et al. 2018 $NDSI1 = (\text{redEdge} - G) / (\text{redEdge} + G)$ $NDSI2 = (\text{redEdge} - B) / (\text{redEdge} + B)$ $NDSI3 = (\text{redEdge} - \text{avg}(G, B)) / (\text{redEdge} + \text{avg}(G, B))$, hvor $\text{avg}(\cdot)$ står for gennemsnittet • NG $NG = NIR / G$ • Normalised Ratio Vegetation Index, Baret 1991 $NRVI = (R / NIR - 1) / (R / NIR + 1)$ • Optimized soil-adjusted vegetation index, Rondeaux et al. 1996 $OSAVI = (1 + 0.16 * ((NIR - R) / (NIR + R + 0.16)))$ • RDVI $RDVI = (NIR - R) / (NIR + R)^2$ • Ratio spectral index (RSI), Zhao et al. 2018 $RSI1 = \text{redEdge} / B$; $RSI2 = \text{redEdge} / G$; $RSI3 = \text{redEdge} / \text{avg}(B, G)$ • Ratio Vegetation Index $RVI = R / NIR$ • Soil Adjusted Vegetation Index, Huete 1988 $SAVI = (NIR - R) * (1 + L) / (NIR + R + L)$, hvor $L = 0.5$ • Simple Ratio Vegetation Index, Birth 1968 $SR = NIR / R$ • Thiam's Transformed Vegetation Index Thiam 1997 $TTVI = \sqrt{\text{abs}((NIR - R) / (NIR + R) + 0.5)}$ • Transformed Chlorophyll Absorption in Reflectance Index, Haboudane et al. 2002 $TCARI = 3 * ((\text{redEdge} - R) - 0.2 * (\text{redEdge} - G)) * (\text{redEdge} / R)$ • TCARI / OSAVI ratio, Haboudane et al. 2002 $TCARI\text{-}OSAVI\text{-ratio} = TCARI / OSAVI$ • Transformed Vegetation Index, Deering 1975 $VI = \sqrt{(NIR - R) / (NIR + R) + 0.5}$

RGB-based indices:

- Brightness Index - Average soil reflectance magnitude
 $BI = \sqrt{(R^2 + G^2 + B^2)} / 3$
- Soil Colour Index
 $SCI = (R - G) / (R + G)$
- Green leaf index, Louhaichi et al. (2001)
 $GLI = (2 * G - R - B) / (2 * G + R + B)$
- Primary colours Hue Index
 $HI = (2 * R - G - B) / (G - B)$
- Normalized difference turbidity index Water
 $NDTI = (R - G) / (R + G)$
- Normalized green red difference index (sometimes GRVI) Tucker (1979)
 $NGRDI = (G - R) / (G + R)$
- Spectral Slope Saturation Index
 $SI = (R - B) / (R + B)$
- A Visible Atmospherically Resistant Index (VARI)
 $VARI = (G - R) / (G + R - B)$
- VVI
 $VVI = (1 - (R - 30) / (R + 30)) * (1 - (G - 50) / (G + 50)) * (1 - (B - 1) / (B + 1))$
- GLAI
 $GLAI = (25 * (G - R) / (G + R - B) + 1.25)$
- Green-Red Vegetation Index
 $GRVI = (G - R) / (G + R)$
- Coloration Index
 $CI = (R - B) / R$
- Overall Hue Index
 $HUE = \text{atan}(2 * (R - G - B) / 30.5 * (G - B))$
- Overall Saturation Index
 $SAT = (\max(R, G, B) - \min(R, G, B)) / \max(R, G, B)$

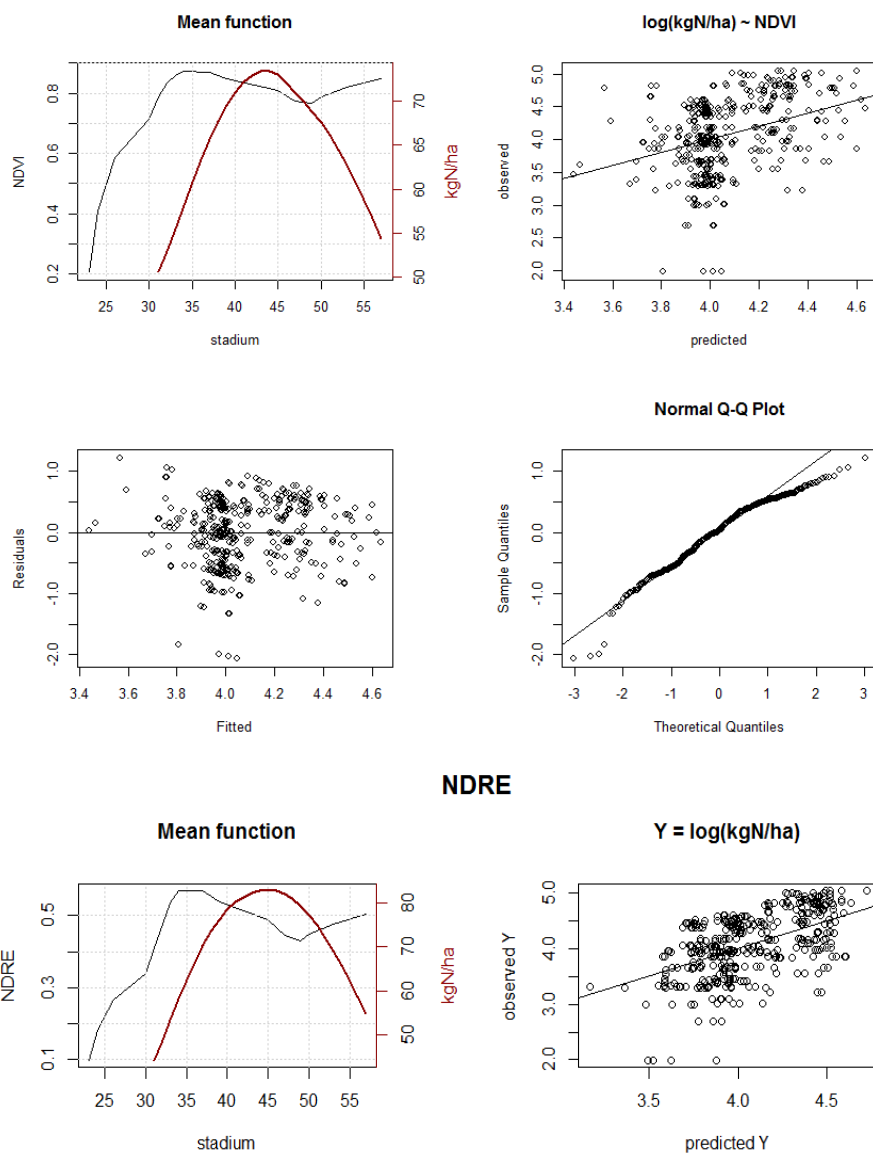


Fig. 1: **NDVI** og **NDRE**. Øverste og nederste venstre paneler: De i FDA-modellen estimerede $kgNA/ha(t)$ og $x(t)$ mean-funktioner. Øverste og nederste højre paneler: observeret vs. prædikeret. Midt-panelet viser modelkontrol for NDVI.

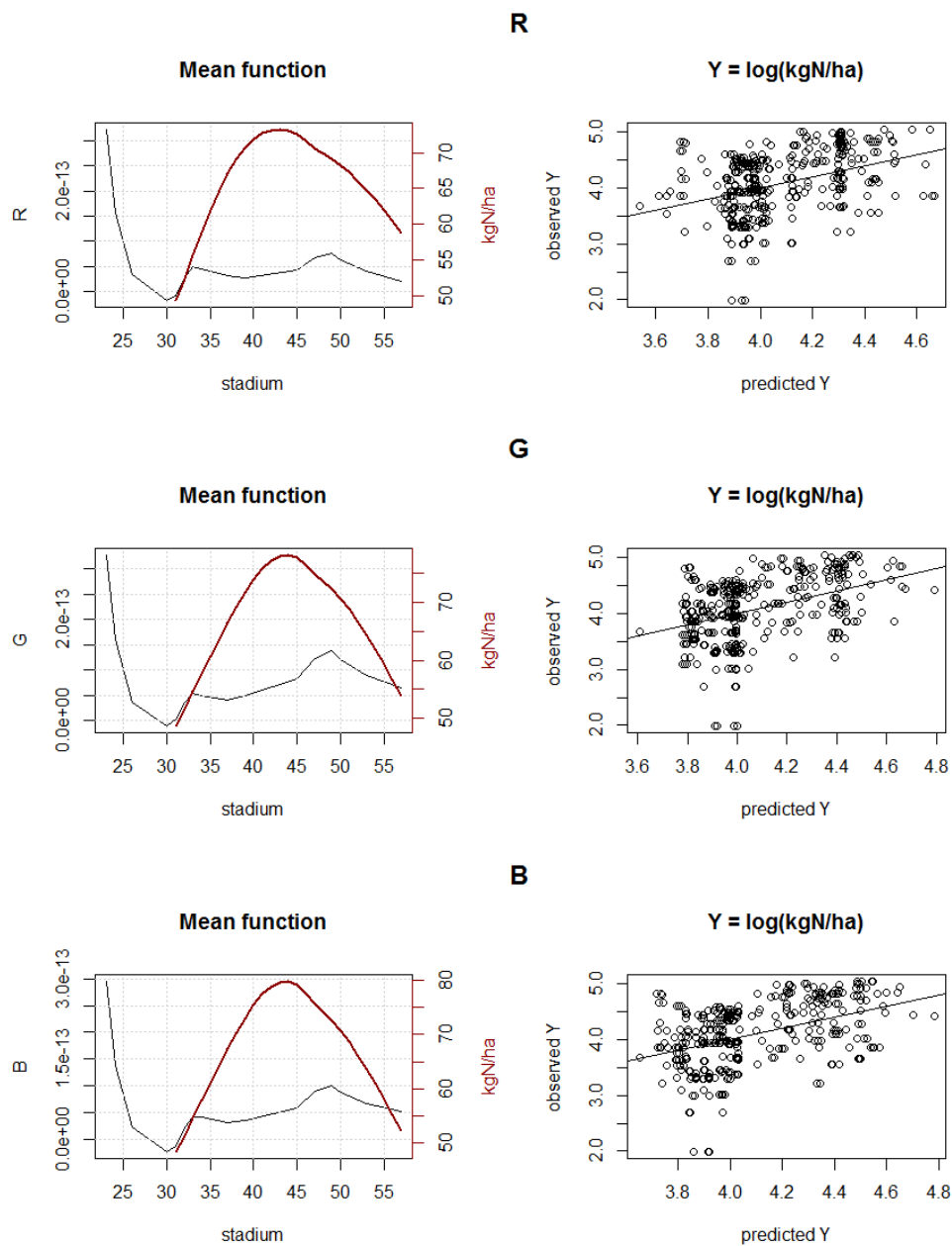


Fig. 2: **R**, **G** og **B**. Venstre paneler: De i FDA-modellen estimerede $kgNA/ha(t)$ og $x(t)$ mean-funktioner. Højre paneler: observeret vs. prædikeret.

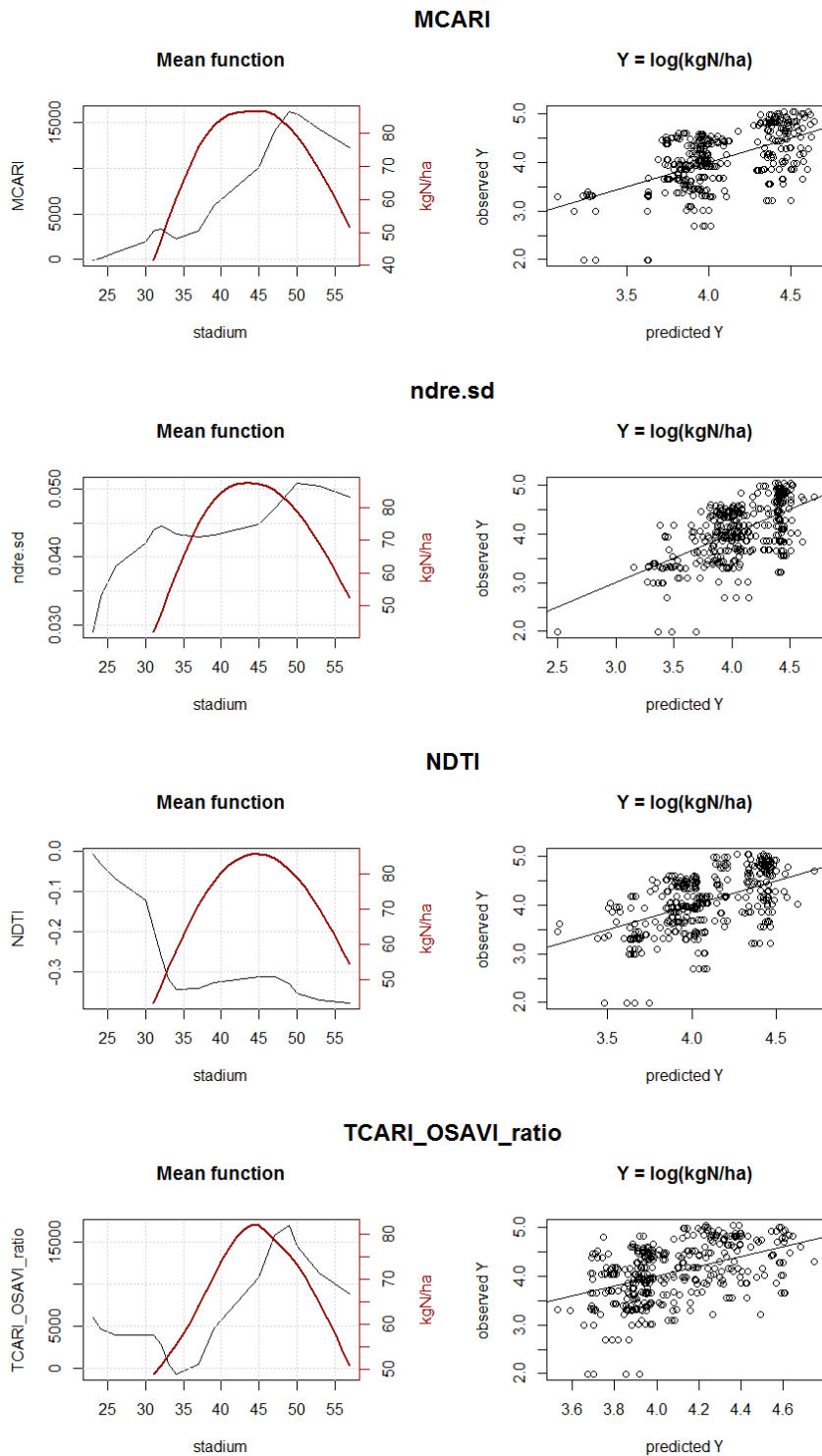


Fig. 3: **MCARI**, **NDRE_{sd}**, **NDTI** og **TCARI/OSAVI-ratio**. Venstre paneler: De i FDA-modellen estimerede $kgNA/ha(t)$ og $x(t)$ mean-funktioner. Højre paneler: observeret vs. prædikeret.

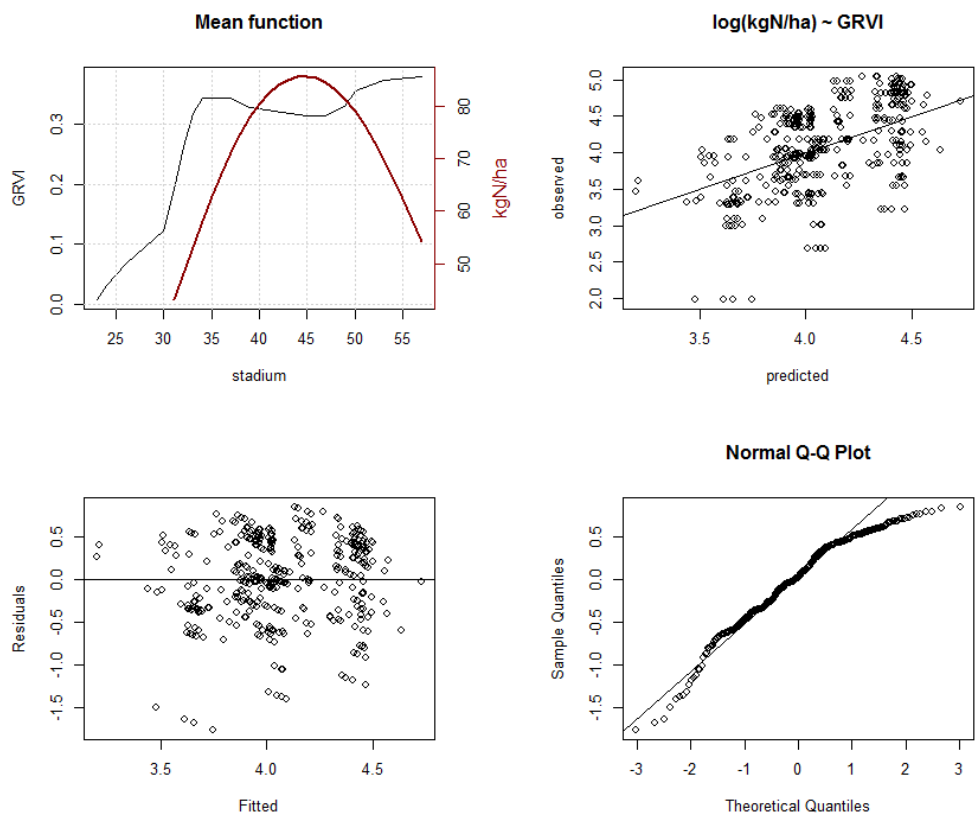


Fig. 4: **GRVI**. Øverste panel til venstre: De i FDA-modellen estimerede $kgN/ha(t)$ og $x(t)$ mean-funktioner. Øverste til højre paneler: observeret vs. prædikeret. Nederste panel viser modelkontrol.

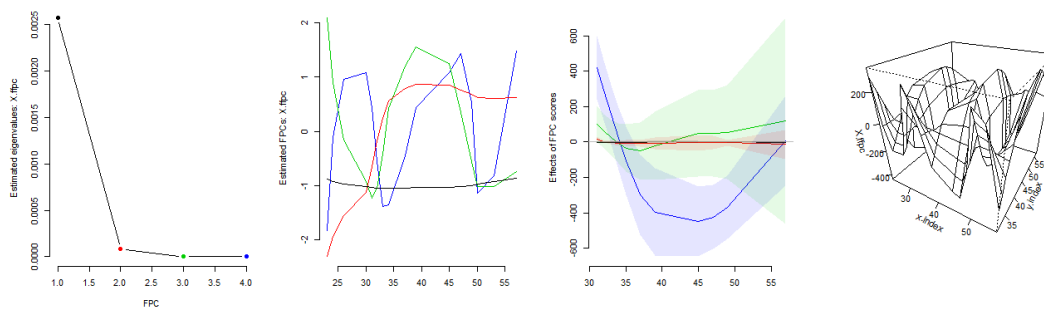


Fig. 5: **GRVI**. Plots fra fPCA analysen af GRVI.

Referencer

Kokoszka, P. & Reimherr, M (2017). Introduction to Functional Data Analysis. CRC Press, NY.

Jeff Goldsmith, Fabian Scheipl, Lei Huang, Julia Wrobel, Jonathan Gellar, Jaroslaw Harezlak, Mathew W. McLean, Bruce Swihart, Luo Xiao, Ciprian Crainiceanu and Philip T. Reiss (2018). refund: Regression with Functional Data. R package version 0.1-17. <https://CRAN.R-project.org/package=refund>