
MODELLER I REMOTE SENSING AF VEGETATION

SPECIALE

KRISTINE BACHER

Roskilde Universitet
Geografi og matematik
August 2014

Vejledere: Eva Bøgh og Carsten Lunde Petersen

Abstract

The purpose of this thesis is to investigate how the spectral signal from an agricultural field is influenced by the vertical distribution of chlorophyll through the canopy. In particular it is examined whether a uniform vertical distribution of chlorophyll improves the results of two types of models using satellite-based remote sensing data.

Analysis of variance and simple linear regression is used to characterise the vertical chlorophyll profile in a canopy for field plots in five different European countries with a variety of crops. Out of a total of 68 fields, 28 fields are found to be characterised by a uniform vertical chlorophyll distribution while 16 fields have either a decreasing or increasing vertical distribution of chlorophyll. The vertical chlorophyll profile for 6 fields cannot be characterised by either of these models. 18 fields are not classified as observations lack homogeneity of variances.

Biophysical conditions can be estimated by an empirical-statistical model where spectral vegetation indices are related to field measurements of biophysical variables such as leaf area index (LAI) or leaf chlorophyll content. The empirical relation cannot be generalised as it depends on crop species, context and time. A physically-based model such as PROSAIL can be used to estimate biophysical variables. PROSAIL uses input regarding the geometry and architecture of the canopy as well as leaf pigment content to simulate the spectral signal. The Matlab code for PROSAIL was adjusted to compensate for an error in the implementation of the ellipsoidal leaf angle distribution.

Simulated reflectance data is used to evaluate the sensitivity of the vegetation indices. They are shown to be responsive of changes in leaf angle and LAI and the chlorophyll index in particular is shown to be sensitive of the leaf chlorophyll content. The empirical relation between the vegetation indices and LAI as well as $\text{Chl}_{\text{total}}$ is improved for fields with a uniform vertical chlorophyll profile but remains highly context-dependent.

PROSAIL is used to simulate the reflectance from a number of Danish wheat and barley fields from 2008 and 2009. The deviation between measured and simulated reflectance is evaluated in relation to LAI, plant type, landscape and vertical distribution of chlorophyll. It is shown that the deviation regarding red reflectance is influenced by the vertical chlorophyll profile however the simulation of reflectance of visible light is generally improved at low LAI. The simulation of near infrared reflectance is improved at higher LAI levels.

Resume

Formålet med dette speciale er at undersøge hvordan det spektrale signal fra en landbrugsmark påvirkes af den vertikale fordeling af klorofyl ned gennem plantestanden. Særligt undersøges det, om en ligefordeling af klorofylindhold forbedrer resultaterne for to typer af modeller, der anvender satellitbaseret remote sensing-data

Ensidet variansanalyse og simpel lineær regression anvendes til at karakterisere det vertikale klorofylprofil i en plantestand for landbrugsmarker fordelt i fem forskellige europæiske lande med forskellige plantetyper. Ud af 68 marker ses det, at 28 marker kan karakteriseres ved en uniform fordeling af klorofyl ned gennem plantestanden, mens 16 har en aftagende eller voksende vertikal fordeling af klorofyl. Det vertikale klorofylprofil for 6 marker kan hverken karakteriseres som uniformt eller lineært. 18 marker testes ikke, da data mangler varianshomogenitet.

Biofysiske forhold kan estimeres med en empirisk-statistisk model, hvor spektrale vegetationsindekser sammenholdes med feltmålinger af biofysiske variable som bladarealindeks (LAI) og klorofylindhold i bladene. Den empiriske sammenhæng er ikke generaliserbar, da den afhænger af plantetype, område og tidspunkt. En fysisk-baseret model som PROSAIL kan også anvendes til at estimere vegetationsforhold. PROSAIL anvender input om plantestandens geometri og arkitektur samt bladenes indhold af pigmenter som klorofyl til at simulere det spektrale signal.

Matlab-koden til PROSAIL blev justeret for at kompensere for en fejl i implementeringen af den ellipsoidale bladvinkelfordeling.

Simuleret reflektans fra PROSAIL anvendes til at vurdere vegetationsindeksernes følsomhed. Det ses, at de i høj grad responderer på ændringer i bladvinkel og LAI, og at særligt klorofylindekset er følsomt over for indholdet af klorofyl i bladene. Den empiriske sammenhæng mellem vegetationsindekserne og LAI og Chl_{total} forbedres for marker med et uniformt vertikalt klorofylprofil men er fortsat kontekstafhængig.

PROSAIL anvendes til at simulere reflektansen fra en række danske marker fra 2008 og 2009 med især hvede og byg. Afvigelserne mellem den målte og den simulerede reflektans vurderes i forhold til LAI, plantetype, landskab og vertikal fordeling af klorofyl. Det ses, at afvigelsen mellem målt og simuleret rød reflektans påvirkes af det vertikale klorofylprofil. Reflektans af synligt lys simuleres dog bedst ved lavt LAI. Den nærinfrarøde reflektans simuleres bedst ved højere LAI.

Indhold

1 Indledning	7
1.1 Problemstilling og problemformulering	9
1.1.1 Afgrænsning	9
1.2 Opbygning af rapporten	10
2 Teori	11
2.1 Sandsynlighedsregning og statistik	11
2.1.1 Kontinuerte fordelinger	11
2.1.2 Normalfordelingen i flere dimensioner	15
2.1.3 Den statistiske model og hypotesetest	16
2.1.4 Lineære normale modeller	18
2.1.5 Opsamling	34
2.2 Remote sensing	34
2.2.1 Begreber i remote sensing af vegetation	34
2.2.2 Spektrale vegetationsindekser	38
2.2.3 PROSAIL	45
3 Data	64
3.1 Landskaber	64
3.1.1 Områdebeskrivelse	66
3.2 Feltmålinger	68
3.2.1 Bladarealindeks og middelbladvinkel	68
3.2.2 Bladklorofyl	71
3.3 Satellitdata	73
4 Metode	74
4.1 Karakterisering af vertikale klorofylprofiler	74
4.2 Variation i datasættet	75
4.3 Empirisk–statistiske modeller	75
4.4 PROSAIL	76
4.4.1 Følsomhedsanalyse af PROSAIL	76
4.4.2 Vegetationsindeksernes følsomhed	76
4.4.3 Simulering af reflektans	77
5 Resultater	78
5.1 Landskabsvariation	78
5.2 Vertikale klorofylprofiler	79
5.3 Empirisk sammenhængen mellem overfladereflektans og blada- realindeks	82

5.4	Empirisk–statistiske modeller	85
5.4.1	Simple ratio	85
5.4.2	NDVI, GNDVI og EVI2	85
5.4.3	Klorofylindekset	87
5.5	Følsomhedsanalyse af PROSAIL	90
5.5.1	Følsomhed over for tør og våd jordbund	96
5.5.2	Resultat af følsomhedsanalysen	96
5.6	Vegetationsindeksernes følsomhed	99
5.7	Simulering af reflektans med PROSAIL	104
5.7.1	Målt og simuleret reflektans i forhold til profiler	104
5.7.2	Absolutte afvigelser	105
5.7.3	Opsamling	111
6	Diskussion	112
6.1	Det vertikale klorofylprofil	112
6.1.1	Sammenligning med Boegh et al. [2013]	113
6.1.2	Andre fremgangsmåder til karakteristik af klorofylprofil	114
6.2	Empirisk–statistiske modeller	116
6.2.1	Lineær regression	117
6.2.2	Ikke-lineær regression	118
6.2.3	Følsomhedsanalyse af vegetationsindekserne	119
6.3	Fysisk-baseret modellering	120
6.3.1	Følsomhedsanalyse	121
6.3.2	Simuleret og målt reflektans	123
7	Konklusion	125
	Appendiks	128
A	Ligningssystemet fra SAIL	129
B	Vertikale klorofylprofiler	136
C	Fraktildiagrammer	139
D	GNDVI og EVI2	141
E	Resultat af simuleringer	144

Forord

Dette er et integreret speciale mellem geografi og matematik. De grundlæggende principper bag remote sensing og de anvendte modeller introduceres i rapporten. For den matematiske del forudsættes det, at læseren er bekendt med lineær algebra og sandsynlighedsregning.

Jeg vil gerne takke mine vejledere Eva Bøgh og Carsten Lunde Petersen. Jeg vil også gerne takke Jørgen Larsen for bidrag omkring specialets anvendelse af statistiske metoder. Endelig vil jeg takke min mand Andreas for uundværlig støtte gennem specialearbejdet.

Simuleringer af den fysisk-baserede model er foretaget i Matlab [MATLAB[®], 2013]. R [R Development Core Team, 2013] er anvendt til statistisk databehandling og produktion af figurer. Quantum GIS [QGIS Development Team, 2013] er anvendt til at fremstille kort.

Rapporten er skrevet i L^AT_EX.

1. Indledning

Remote sensing er en metode til at kortlægge Jordens overflade ud fra den energi, der udstråles fra Jorden. En sensor monteret på en flyvemaskine, en satellit eller tættere på landjorden opfanger den reflekterede stråling fra landjorden, og målingerne kan anvendes til at konstruere et billede af landskabet. I optisk remote sensing har energien form af reflekteret sollys, mens termisk remote sensing beskæftiger sig med varmestråling [Richards og Jia, 2006].

Stråling fra solen er primært kortbølget stråling i form af synligt lys og nærinfrarød stråling. Når solens stråler passerer gennem atmosfæren på vej mod jordoverfladen, vil noget af strålingen blive spredt pga. skyer, vandindhold mv. Når strålingen rammer jordoverfladen, vil den absorberes eller reflekteres afhængig af det materiale, den rammer. I optisk remote sensing fra en satellit vil en sensor opfange den del af den reflekterede stråling, der kommer tilbage gennem atmosfæren. De optiske egenskaber ved vegetation er absorption af synligt lys og refleksion af nærinfrarød stråling. Derfor kan især kontrasten mellem refleksion af synligt lys og nærinfrarød stråling gøre det muligt at skelne planter fra bar jord, vand og andre forhold på landjorden. Potentielt set kan remote sensing observationer også bruges til at skelne mellem forskellige plantearter og vegetationens tilstand, eftersom det såkaldte spektrale signal fra planter afhænger af de konkrete forhold.

Det er især de grønne planters klorofylholdige blade, som giver vegetationen de karakteristiske optiske egenskaber. Remote sensing kan derfor potentielt anvendes til at estimere biofysiske forhold ved planter fx klorofylkoncentration og tæthed af blade. Sådanne estimationer har mange potentielle anvendelsesmuligheder inden for landbrug, hvor man gerne vil målrette brug af gødning til at give den bedst mulige udnyttelse af jorden med minimal miljøpåvirkning [Casa et al., 2010].

Tilgængelighed af kvælstof er essentiel for planter, da kvælstof er nødvendig for, at planterne kan danne proteiner. Der er også en nær sammenhæng mellem klorofyl og kvælstof, da klorofyl indeholder kvælstof. Optagelsen af andre næringsstoffer afhænger desuden af kvælstoftilgængeligheden. Planters indhold af kvælstof giver således en god indikation af deres tilstand. Hvis en plante har kvælstofmangel, vil farven blive gullig, og den vil have svært ved at omdanne sukker til protein. Hvis planten har et overskud af kvælstof, kan den få øget vækst og i nogle tilfælde blive skrøbelig. I nogle tilfælde kan kvælstof ophobes i bladene [Brady og Weil, 2008].

Bladkvælstof er knyttet til bladklorofyl i et artsspecifikt forhold. [Boegh et al., 2013] Når sollys rammer en plantestand, vil noget af lyset absorberes og andet vil gennemtrænge plantestanden. Ofte anvendes Beers lov til be-

skrive lysudslukningen som en eksponentielt aftagende funktion af dybden i plantestanden. Ifølge Boegh et al. [2013] er der teoretisk belæg for at antage, at planter vil fordele kvælstofressourcer proportionalt med lysgradienten i plantestanden.

Derfor kunne man formode, at planters klorofylindhold også varierer vertikalt gennem plantestanden. En plante som majs har en klokkeformet vertikal fordeling af blade og klorofylindhold [Ciganda et al., 2012; Winterhalter et al., 2012]. Den vertikale fordeling af klorofyl er dog ikke en faktor, der tages højde for i modeller som Beers lov [Ciganda et al., 2012]. Her antages udslukningskoefficienten at være konstant gennem det vertikale profil. Lysabsorption er dog netop styret af klorofylindholdet, hvorfor en vertikal variation i klorofylindholdet kan tænkes at have indflydelse på udslukningskoefficienten.

Boegh et al. [2013] undersøger og simulerer biofysiske forhold ved landbrugsmarker i forskellige europæiske landskaber vha. modellen REGFLEC. I den forbindelse er målinger relateret til bladklorofyl og bladkvælstof målt på afgrøder fra landbrugsmarker fordelt i flere europæiske lande. For de fleste afgrøder er målingerne foretaget ned igennem plantestanden for at give et udtryk for det vertikale profil. Det undersøges, om klorofylindholdet er ligefordelt eller om der er en vertikal gradient. Gradienten kan være voksende eller aftagende gennem plantestanden.

Der er to overordnede tilgange til at estimere biofysiske forhold i remote sensing. Den ene type modeller er baseret på den empirisk-statistiske sammenhæng mellem de spektrale vegetationsindekser og de biofysiske variable, der beskriver vegetationen. Vegetationsindekserne anvender forholdet mellem reflekteret synligt lys og nærinfrarød stråling og kan beskrive graden af vegetation i et område, men de er følsomme over for jordbundsforhold og atmosfærisk støj. Det empiriske forhold mellem vegetationsindekser og biofysiske variable vil altid være kontekstafhængigt og derfor forbundet med stor usikkerhed, hvis en funden sammenhæng skal overføres til en anden lokalitet [Houborg et al., 2007].

Den anden type modeller er de fysisk-baserede modeller. De findes i mange afskygninger og er ofte baseret på strålingstransportteori. Dermed tager de afsæt i fysiske lovmæssigheder og inkluderer et stort antal inputvariable, der beskriver bladenes optiske egenskaber ud fra indhold af diverse bestanddele og plantestandens struktur og geometri, til at simulere det spektrale signal. Hvis de inverteres, kan de beskrive de biofysiske forhold i et område ud fra det spektrale signal. Det kan dog ikke lade sig gøre udelukkende vha. multispektralt satellitdata. En ulempe ved strålingstransportmodellerne er nemlig, at antallet af inputvariable overstiger antallet af outputvariable, fordi en plantestand er et komplekst fysisk objekt med mange gensidigt relaterede variable. Derfor kan fysisk-baserede modeller generelt ikke inverteres uden at inkludere a priori viden, så nogle af variablene bortreduceres.

PROSAIL er et eksempel på en fysisk-baseret model, der sammenkobler

modellerne SAIL og PROSPECT. PROSAIL kan anvendes direkte til at simulere reflektans fra en plantestand baseret på inputparametre om strukturelle forhold og bladenes indhold af forskellige pigmenter, der har indflydelse på refleksion og absorption af stråling. PROSAIL kan også anvendes inverst til at estimere biofysiske variable ud fra reflektansspektret, hvis a priori viden anvendes. Modellen REGFLEC kombinerer spektrale vegetationsindekser med strålingstransportmodeller for at løse det inverse problem [Boegh et al., 2013]. PROSAIL indgår i REGFLEC sammen med en model for atmosfærisk korrektion.

1.1 Problemstilling og problemformulering

I det indsamlede datasæt fra Boegh et al. [2013] er der foretaget feltmålinger af klorofylindholdet vertikalt gennem plantestanden på markerne ligesom der foreligger satellitmålt reflektans. Derfor kan datasættet anvendes som grundlag for at undersøge sammenhængen mellem det vertikale klorofylprofil og det spektrale signal fra landbrugsmarker. På baggrund heraf opskrives følgende problemformulering:

- Hvilken betydning har det vertikale klorofylprofil på overfladereflektansen fra en landbrugsmark?

For at svare på problemformuleringen skal det vertikale klorofylprofil karakteriseres ud fra data. I den forbindelse anvendes statistiske metoder. Dernæst skal følgende spørgsmål besvares:

- I hvilken grad har det vertikale klorofylprofil indflydelse på den empiriske sammenhæng mellem spektrale vegetationsindekser og hhv. blada-realindeks og klorofylindhold i plantestanden?
- Hvilken indflydelse har det vertikale klorofylprofil på forskellen mellem simuleret og målt reflektans i den fysisk-baserede model PROSAIL?

1.1.1 Afgrænsning

Der er ikke indsamlet feltdata i forbindelse med specialet, da specialet ikke er eksperimentelt. I stedet anvendes det tilgængelige datasæt fra Boegh et al. [2013]. Datasættet er omfattende, fordi det er baseret på tidskrævende feltmålinger foretaget i fem forskellige lande. Der er dog ulemper forbundet ved at anvende feltdata indsamlet af andre og med andre beskrevne formål. Eksempelvis er datasættet fra Boegh et al. [2013] ikke fremstillet specifikt med henblik på at vurdere betydningen af det vertikale klorofylprofil i remote sensing-data.

Den valgte statistiske metode til karakterisering af det vertikale klorofylprofil hænger uløseligt sammen med datasættets udformning. Derfor er

ensidet variansanalyse og lineær regression valgt. Metoderne anvendes desuden sammen med ikke-lineær regression til at vurdere den empirisk-statistiske model og simuleringer fra den fysisk-baserede model i forhold til data.

PROSAIL anvendes direkte til at estimere reflektansspektret ud fra en række inputparametre. Det estimerede reflektansspektrum sammenlignes derefter med det målte. At PROSAIL ikke anvendes inverst skyldes, at modelinversionen ikke er en simpel procedure og generelt kræver komplekse løsningsstrategier og a priori viden. Derfor vurderes PROSAIL i forhold til den direkte simulering af reflektans ud fra implementeringen i Matlab, der er tilgængelig i Jacquemoud et al. [2013].

PROSAIL består af to modeller, PROSPECT og SAIL. PROSPECT leverer input til SAIL om bladenes spektrale signal, mens SAIL simulerer reflektansen fra plantestanden ud fra input om solen og sensorens vinkler og plantestandens arkitektur. PROSPECT er udviklet af Jacquemoud og Baret [1990] behandler bladenes optiske egenskaber dvs. reflektans og transmittans af elektromagnetisk stråling som funktion af bladets indhold af pigmenter. Modellerne vil ikke blive udsat for en fuldstændig matematisk gennemgang, da det vil være for omfattende. Behandlingen af PROSPECT vil fokusere på parametrenes betydning i simuleringen af reflektans og transmittans for blad. SAIL er udviklet af Verhoef [1984, 1985] med senere tilføjelser. Behandlingen af SAIL fokuserer på principperne i den originale model og de senere tilføjelser kommenteres. Det er kun udvalgte dele af den originale model, der behandles grundigt.

Gennemgangen af PROSAIL gør det muligt at foretage en følsomhedsanalyse, hvor parametrenes betydning for resultatet vurderes. I den forbindelse vurderes PROSAILs resultater kun i forhold til de frekvensbånd, der findes målte værdier at sammenligne med. Dvs. at PROSAILs følsomhed vurderes i et sæt afgrænsede intervaller.

1.2 Opbygning af rapporten

Næste kapitel beskriver den anvendte teori. Kapitlet falder i to dele af forskellig karakter. Den første del omfatter den sandsynlighedsteoretiske baggrund for de udvalgte statistiske metoder. Anden del af kapitlet beskriver de to typer af modeller i remote sensing af vegetation. Efter en indledende begrebsafklaring forklares først den empirisk-statistiske model i form af en række udvalgte vegetationsindekser og deres konceptuelle baggrund og dernæst den fysisk-baserede model PROSAIL. PROSAIL består som nævnt af to delmodeller, PROSPECT og SAIL, der gennemgås hver for sig. Datakapitlet beskriver de relevante dele af datasættet fra Boegh et al. [2013] samt de metoder, data er indsamlet ud fra. Der indgår også en områdebeskrivelse af landskaberne, særligt det danske landskab baseret på forskellige andre kilder. Herefter følger metodekapitel og resultatkapitel. Resultaterne bliver efterfølgende diskuteret.

2. Teori

Teorikapitlet falder i to dele. Det første afsnit gennemgår den matematiske teori bag de statistiske metoder, der anvendes i databehandlingen. Det andet afsnit gennemgår begreber inden for remote sensing af vegetation og introducerer de to forskellige modeltilgange, de empirisk–statistiske modeller og den fysisk-baserede model.

2.1 Sandsynlighedsregning og statistik

Det følgende afsnit er inspireret af gennemgangene i Larsen [2006a,b] og Tjur [2002b]. Formålet med afsnittet er at danne en baggrund for de statistiske tests, der udføres. Derfor tages der afsæt i sandsynlighedsregning og en række sandsynlighedsfordelinger præsenteres. Derefter præsenteres den statistiske model samt hypotesetest. Endelig introduceres lineære normalfordelingsmodeller generelt og lineær regression samt ensidet variansanalyse i særdeleshed.

2.1.1 Kontinuerte fordelinger

Formålet med dette afsnit er at indføre en række definitioner og en enkelt sætning, der skal danne baggrund for at anvende sandsynlighedstæthedsfunktioner for forskellige konkrete fordelinger.

Delafsnittet er inspireret af Larsen [2006b] og Dudley [2002]. Udgangspunktet er sandsynlighedsregning og det følgende handler om kontinuerte fordelinger.

Et sandsynlighedsrum består af et målbart rum og et sandsynlighedsmål. Et målbart rum kan skrives som (Ω, \mathcal{F}) , hvor Ω er et udfaldsrum og \mathcal{F} er en mængde af delmængder af Ω . En delmængde af udfaldsrummet kaldes en hændelse.

For at definere et sandsynlighedsrum for overtællelige mængder, skal mængden \mathcal{F} af alle delmængder af Ω skal være en σ -algebra.:

Definition 2.1 (σ -algebra). *En mængde \mathcal{F} bestående af delmængder af en ikke-tom mængde Ω kaldes en σ -algebra på Ω hvis følgende tre betingelser er opfyldt:*

1. $\Omega \in \mathcal{F}$
2. \mathcal{F} er afsluttet over for komplementærmængdedannelse. Dvs., at hvis en mængde $C \in \mathcal{F}$, så vil komplementærmængden til C , $C^c \in \mathcal{F}$.

3. \mathcal{F} er afsluttet over for tællelig foreningsmængdedannelse. Dvs., at hvis A_1, A_2, \dots er en følge af hændelser i \mathcal{F} , så er $\cup_{n=1}^{\infty} A_n$ også i \mathcal{F} .

Af definitionen følger, at den tomme mængde $\emptyset \in \mathcal{F}$. Hvis udfaldsrummet Ω er en overtællelig mængde, så vil mængden \mathcal{F} typisk være en ægte delmængde af potensmængden 2^Ω , som er mængden af alle delmængder af Ω . Dvs., at \mathcal{F} ikke vil bestå af hele potensmængden.

Sandsynlighedsmålet noteres P . Man siger, at $P(A)$ er sandsynligheden for, at hændelsen A indtræffer. P er en afbildning fra \mathcal{F} ind i de reelle tal \mathbb{R} , dvs. $P : \mathcal{F} \rightarrow \mathbb{R}$, om hvilken der gælder følgende:

1. P er normeret, dvs. $P(\Omega) = 1$
2. P er ikke-negativt, dvs. $P(A) \geq 0$ for alle $A \in \mathcal{F}$.
3. P er σ -additivt, dvs. hvis A_1, A_2, \dots er en følge af parvist disjunkte hændelser fra \mathcal{F} så er $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

På baggrund af definitionen af et sandsynlighedsrum, kan en stokastisk variabel defineres:

Definition 2.2. Hvis et sandsynlighedsrum udgøres af (Ω, \mathcal{F}, P) , så er en stokastisk variabel X på (Ω, \mathcal{F}, P) en afbildning af Ω ind i \mathbb{R} , dvs. $X : \Omega \rightarrow \mathbb{R}$. En n -dimensional stokastisk variabel afbilder Ω ind i \mathbb{R}^n .

I denne sammenhæng kan udfaldsrummet Ω udgøres af de reelle tal \mathbb{R} eller intervaller i \mathbb{R} . For det n -dimensionale tilfælde kan Ω udgøres af \mathbb{R}^n eller delmængder af \mathbb{R}^n .

Definition 2.3 (Sandsynlighedstæthedsfunktion). En sandsynlighedstæthedsfunktion på \mathbb{R} er en integrabel funktion $f : \mathbb{R} \rightarrow [0, \infty[$ der integrerer til 1, dvs.

$$\int_{\mathbb{R}} f(x) dx = 1$$

Mere generelt er en sandsynlighedstæthedsfunktion på \mathbb{R}^n en integrabel funktion $f : \mathbb{R}^n \rightarrow [0, \infty[$ der integrerer til 1, dvs.

$$\int_{\mathbb{R}^n} f(x) dx = 1$$

Definition 2.4 (Kontinuert fordeling). En kontinuert sandsynlighedsfunktion er en sandsynlighedsfordeling, som har en sandsynlighedstæthedsfunktion. Fordelingsfunktionen for en sandsynlighedstæthedsfunktion f på \mathbb{R} er givet ved den kontinuerte funktion

$$F(x) = \int_{-\infty}^x f(u) du \tag{2.1}$$

Definitionen for en kontinuert fordeling angives ikke i den flerdimensionale udgave, da den i rapporten kun anvendes i forbindelse med fordelinger på \mathbb{R} . I det følgende defineres en række kontinuerte fordelinger, som er relevante i forbindelse med den statistiske model.

En sætning om fordelingen af en transformeret stokastisk variabel præsenteres her, da den skal anvendes senere.

Sætning 2.1 (Fordeling af en transformeret stokastisk variabel). Antag, at X er en reel stokastisk variabel, at I er en åben delmængde af \mathbb{R} , så $P(X \in I) = 1$, og $t : I \rightarrow \mathbb{R}$ er en kontinuert differentiabel funktion defineret på I med $t' \neq 0$ overalt på I . Hvis X har tæthedsfunktion f_X , så har $Y = t(X)$ følgende tæthedsfunktion:

$$f_Y(y) = f_X(x)|t'(x)|^{-1} = f_X(x)|(t^{-1})'(x)|$$

Hvor $x = t^{-1}(y)$ og $y \in t(I)$.

Bevis. Da t er kontinuert differentiabel, er t' kontinuert. Endvidere antages det, at t' er forskellig fra nul, dvs. at t er enten strengt voksende eller strengt aftagende. Hvis det antages, at t er strengt voksende, så vil $Y \leq b \Leftrightarrow X \leq t^{-1}(b)$. Fordelingsfunktionen F_Y kan skrives op:

$$\begin{aligned} F_Y(b) &= P(Y \leq b) \\ &= P(X \leq t^{-1}(b)) \\ &= \int_{-\infty}^{t^{-1}(b)} f_X(x) dx \end{aligned}$$

I integralet udføres substitution med $x = t^{-1}(y)$:

$$= \int_{-\infty}^b f_X(t^{-1}(y))(t^{-1})'(y) dy$$

Funktionen $f_Y(y) = f_X(t^{-1}(y))(t^{-1})'(y)$ har dermed egenskaben, at $F_Y(y) = \int_{-\infty}^y f_Y(u) du$. Dvs. at Y har tæthedsfunktionen f_Y . \square

En anden sætning om tæthedsfunktionen for summen af to stokastiske variable indføres her for at kunne anvende den i et senere afsnit.

Sætning 2.2. Hvis de stokastiske variable X_1 og X_2 er uafhængige og med tæthedsfunktion f_1 og f_2 , så har $Y = X_1 + X_2$ tæthedsfunktionen:

$$f(y) = \int_{\mathbb{R}} f_1(x)f_2(y-x)dx, y \in \mathbb{R}$$

Beviset for sætningen er undladt, da det anvender en flerdimensional udgave af sætning 2.1, hvilket vurderes til at være for omfattende at inddrage.

Eksempler på kontinuerte fordelinger

Formålet med afsnittet er at præsentere en række fordelinger, hvis egenskaber anvendes i de lineære normalfordelingsmodeller.

Gammafordelingen gør brug af gammafunktionen $\Gamma(t)$, som er givet ved

$$\Gamma(t) = \int_0^{\infty} x^{t-1} \exp(-x) dx$$

Definition 2.5 (Gammafordelingen). *Gammafordelingen er givet ved sandsynlighedstæthedsfunktionen*

$$f(x) = \begin{cases} \frac{1}{\Gamma(k)\beta^k} x^{k-1} \exp(-x/\beta) & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

Hvor $k > 0$ er formparameter og $\beta > 0$ er skalaparameter.

χ^2 -fordelingen er et eksempel på en gammafordeling:

Definition 2.6 (χ^2 -fordeling). *χ^2 -fordelingen med n frihedsgrader er en gammafordeling med formparameter $n/2$ og skalaparameter 2 og er således givet ved sandsynlighedstæthedsfunktionen*

$$f(x) = \frac{1}{\Gamma(n/2)2^{n/2}} x^{n/2-1} \exp(-\frac{1}{2}x), x > 0 \quad (2.2)$$

Definition 2.7 (F-fordelingen). *F-fordelingen med f_t og f_n frihedsgrader er den kontinuerte fordeling på $]0; \infty[$ med tæthedsfunktion*

$$f(x) = \frac{\Gamma(\frac{f_t+f_n}{2})}{\Gamma(\frac{f_t}{2})\Gamma(\frac{f_n}{2})} f_t^{f_t/2} f_n^{f_n/2} \frac{x^{(f_t-2)/2}}{(f_n + f_t x)^{(f_t+f_n)/2}}$$

Ifølge Tjur [2002a] kan F-fordelingen fortolkes som fordelingen af forholdet mellem to uafhængige χ^2 -fordelte stokastiske variable med henholdsvis f_t og f_n frihedsgrader – på nær en skalaparameter. f_t og f_n svarer til antallet af frihedsgrader i henholdsvis tæller og nævner i forholdet.

Definition 2.8 (Standardnormalfordelingen). *Standardnormalfordelingen er givet ved sandsynlighedstæthedsfunktionen*

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2), x \in \mathbb{R}$$

Definition 2.9 (Normalfordelingen). *Normalfordelingen er givet ved sandsynlighedstæthedsfunktionen*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right), x \in \mathbb{R}$$

Hvor μ er positionsparameter og σ^2 kvadratisk skalaparameter.

Standardnormalfordelingen er en normalfordeling med $\mu = 0$ og $\sigma^2 = 1$.

2.1.2 Normalfordelingen i flere dimensioner

Normalfordelingen kan også defineres i n dimensioner, hvilket er relevant for at anvende normalfordelingsmodeller. For at udvide definitionen til n dimensioner, skal der tages højde for antagelser om varianser, hvorfor definitionen falder i flere skridt.

En n -dimensional stokastisk variabel \mathbf{X} kan opfattes som en vektor i \mathbb{R}^n hvis koordinater er de endimensionale stokastiske variable X_1, X_2, \dots, X_n . Middelværdien af \mathbf{X} er vektoren $E\mathbf{X}$ bestående af middelværdierne EX_1, EX_2, \dots, EX_n . Variansen af \mathbf{X} er en symmetrisk positivt semidefinit $n \times n$ -matrix:

$$\text{Var}(\mathbf{X}) = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_2, X_1) & \cdots & \text{Cov}(X_n, X_1) \\ \text{Cov}(X_1, X_2) & \text{Var}(X_2) & \cdots & \text{Cov}(X_n, X_2) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_1, X_n) & \text{Cov}(X_2, X_n) & \cdots & \text{Var}(X_n) \end{pmatrix} \quad (2.3)$$

Variansmatricen er symmetrisk, da $\text{Cov}(X, Y) = \text{Cov}(Y, X)$. At den er positivt semidefinit betyder, at det for alle \mathbf{x} gælder at $\langle \mathbf{x}, \Sigma \mathbf{x} \rangle \geq 0$. Hvis ulighedstegnet er skarpt, kaldes matricen positivt definit.

Definition 2.10 (n -dim. standardnormalfordeling). *Den n -dimensionale standardnormalfordeling er den n -dimensionale kontinuerte fordeling med sandsynlighedstæthedsfunktion*

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \|\mathbf{x}\|^2\right), \mathbf{x} \in \mathbb{R}^n$$

Det følger af definitionen, at en stokastisk variabel \mathbf{X} er standardnormalfordelt hvis og kun hvis X_1, X_2, \dots, X_n er standardnormalfordelte.

Definition 2.11 (Regulær normalfordeling). *Antag at $\boldsymbol{\mu} \in \mathbb{R}^n$, dvs. $\boldsymbol{\mu}$ er en n -dimensional vektor, og at Σ er en positiv definit lineær afbildning af \mathbb{R}^n ind i sig selv, dvs. en positiv definit $n \times n$ -matrix. Så er den n -dimensionale regulære normalfordeling med middelværdi $\boldsymbol{\mu}$ og varians Σ den n -dimensionale kontinuerte fordeling hvis tæthedsfunktion er givet ved:*

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\det \Sigma|^{1/2}} \exp(-1/2(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})), \text{ hvor } \mathbf{x} \in \mathbb{R}^n \quad (2.4)$$

Hvis X_1, X_2, \dots, X_n er uafhængige og endimensionelt normalfordelte, så X_i har middelværdi μ_i og varians σ^2 , dvs. at alle X_i 'er har ens varians, så er variansmatricen Σ givet ved $\sigma^2 I$, dvs. matricen:

$$\sigma^2 I = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix} \quad (2.5)$$

Determinanten af matricen $\sigma^2 I$ er så givet ved:

$$\det(\sigma^2 I) = \prod_{i=1}^n \sigma^2 = (\sigma^2)^n \quad (2.6)$$

Den inverse variansmatrix bliver:

$$\sigma^2 I = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma^2} \end{pmatrix} \quad (2.7)$$

Hvis man indsætter denne variansmatrix i tæthedsfunktionen, kan man omskrive udtrykket:

$$(\mathbf{x} - \boldsymbol{\mu})'(\sigma^2 I)^{-1}(\mathbf{x} - \boldsymbol{\mu}) = \frac{\|\mathbf{x} - \boldsymbol{\mu}\|^2}{\sigma^2} \quad (2.8)$$

Således kan tæthedsfunktionen for den regulære normalfordeling under de nævnte forudsætninger omskrives til følgende:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}} \exp(-1/2(\frac{\|\mathbf{x} - \boldsymbol{\mu}\|^2}{\sigma^2})), \text{ hvor } \mathbf{x} \in \mathbb{R}^n \quad (2.9)$$

Den generelle normalfordeling i n dimensioner defineres på følgende vis:

Definition 2.12 (Den n -dimensionale normalfordeling). *Antag at $\boldsymbol{\mu} \in \mathbb{R}^n$, dvs. $\boldsymbol{\mu}$ er en n -dimensional vektor, og at Σ er en positiv semidefinit lineær afbildning af \mathbb{R}^n ind i sig selv, dvs. en positiv semidefinit $n \times n$ -matrix. Rangem af Σ er p . Den n -dimensionale normalfordeling med middelværdi $\boldsymbol{\mu}$ og varians Σ er fordelingen af $\boldsymbol{\mu} + B\mathbf{U}$, hvor \mathbf{U} er p -dimensionelt standardnormalfordelt og B er en injektiv lineær afbildning af \mathbb{R}^p ind i \mathbb{R}^n .*

2.1.3 Den statistiske model og hypotesetest

I dette afsnit beskrives den statistiske model generelt. Opstilling af en statistisk model har til formål at undersøge et talmateriale gennem en række antagelser. Talmaterialet kaldes en observation og kan betegnes \mathbf{x} . Observationen tilhører et observationsrum \mathcal{X} fx \mathbb{R}^n . I en statistisk model antages det, at \mathbf{x} er en observeret værdi af en stokastisk variabel \mathbf{X} . Man vil gerne undersøge fordelingen af \mathbf{X} ud fra observationen \mathbf{x} . I den statistiske model indgår nogle parametre til at beskrive denne fordeling.

De ukendte parametre betegnes $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ og siges at tilhøre parameterummet Θ , der ofte er en åben delmængde af \mathbb{R}^d . Hvis man eksempelvis antager, at en stokastisk variabel \mathbf{X} af hvilken der foreligger en observation \mathbf{x} er normalfordelt vil $\boldsymbol{\theta}$ bestå af parametrene μ og σ^2 .

For hver værdi af $\boldsymbol{\theta}$ findes et sandsynlighedsmål P_θ på \mathcal{X} . For mindst én værdi af $\boldsymbol{\theta} \in \Theta$ er det rigtigt, at P_θ er fordelingen af \mathbf{X} . I eksemplet fra før, hvor

\mathbf{X} antages normalfordelt, kan modellen formuleres på en sådan måde, at det vil findes et sæt af værdier af μ og σ^2 , som beskriver fordelingen af \mathbf{X} .

Dvs., at en generel statistisk model består af en mængde $\{P_\theta \mid \theta \in \Theta\}$ af sandsynlighedsfordelinger på \mathcal{X} . Observationen \mathbf{x} er fremkommet som en værdi af den stokastiske variabel \mathbf{X} med fordelingen P_θ , som ikke er kendt.

En modelfunktion $f(\mathbf{x}, \theta)$ kan opskrives. For den rigtige værdi af θ , angiver f sandsynlighedstæthedsfunktionen for den stokastiske variabel. Dvs. $f : \mathcal{X} \times \Theta \rightarrow [0, \infty[$.

På samme måde opskrives likelihoodfunktionen $L(\theta) = f(\mathbf{x}, \theta)$. Likelihoodfunktionen er sandsynlighedstæthedsfunktionen som funktion af den ukendte parameter θ hvor \mathbf{x} er fast, dvs. $L : \Theta \rightarrow [0, \infty[$. Likelihoodfunktionen er altså sandsynligheden for det observerede udfald af den stokastiske variabel som funktion af den ukendte parameter. Den ukendte parameter bestemmes som den værdi af θ , der maksimaliserer likelihoodfunktionen. Denne metode kaldes maksimum-likelihood-metoden. For denne værdi $\hat{\theta}$ skal der gælde følgende:

$$L(\hat{\theta}) \geq L(\theta) \text{ for alle } \theta \in \Theta$$

Hvis denne værdi $\hat{\theta}$ findes, kaldes den for maksimaliseringsestimatorens. Maksimaliseringsestimatorens bestemmes ved at foretage en funktionsundersøgelse af likelihoodfunktionen eller log-likelihoodfunktionen, dvs. logaritmen til likelihoodfunktionen. Det er ofte nemmere at bestemme maksimum for log-likelihoodfunktionen. $\hat{\theta}$ er givet ved løsningen til likelihoodligningerne:

$$\frac{d}{d\theta_j} L(\theta) = 0, \text{ hvor } j = 1, 2, \dots, d$$

En hypotese er en formodning om, at observationen \mathbf{x} beskrives lige så godt eller næsten lige så godt ved en reduceret model. Hvis man har en statistisk model med modelfunktion $f(\mathbf{x}, \theta)$, hvor $\mathbf{x} \in \mathcal{X}$ og $\theta \in \Theta$, så er en statistisk hypotese H_0 en påstand om, at den sande parameterværdi $\theta \in \Theta_0$, hvor $\Theta_0 \subset \Theta$. Man skriver $H_0 : \theta \in \Theta_0$.

Et statistisk test går ud på at undersøge overensstemmelsen mellem hypotese og observationer. Overensstemmelsen kan testes gennem kvotientteststørrelsen Q , hvor $\hat{\theta}$ angiver maksimaliseringsestimatorens i den reducerede model, mens $\hat{\theta}$ som før er maksimaliseringsestimatorens i grundmodellen:

$$Q = \frac{L(\hat{\theta})}{L(\hat{\theta})}$$

Kvotientteststørrelsen er et tal mellem 0 og 1, fordi tælleren aldrig kan være større end nævneren. Det skyldes, at tælleren består af maksimum for samme funktion som nævneren, men over en delmængde af definitionsmængden for nævneren. Hvis Q er tæt på nul, vil man forkaste hypotesen. Det svarer nemlig til, at sandsynligheden for observationen i den reducerede model er lille. Hvis Q er tæt på 1, vil man ikke forkaste hypotesen, eftersom den

reducerede model giver en næsten lige så god sandsynlighed for observationen som grundmodellen.

Kvotientteststørrelsen kan også skrives som $-2\log Q = 2(\log L(\hat{\theta}) - \log L(\hat{\theta}_0))$. $-2\log Q$ er en monotont aftagende funktion af Q , som er approksimativt χ^2 -fordelt. Antallet af frihedsgrader for χ^2 -fordelingen svarer til at antallet af bortreducerede parametre under hypotesen. Dvs., hvis antallet af parametre i grundmodellen er d , og antallet af parametre under hypotesen er d_0 , er antallet af frihedsgrader $d - d_0$.

Knyttet til teststørrelsen er testsandsynligheden p . p angiver sandsynligheden for - forudsat at hypotesen er rigtig - at få en værdi af den stokastiske variabel \mathbf{X} , der stemmer dårligere overens med hypotesen end den foreliggende observation \mathbf{x} . Dvs., at hvis man udfører et lignende eksperiment, så angiver p sandsynligheden for at få en teststørrelse, der giver anledning til en ligeså stor eller større tvivl om hypotesen, selv om hypotesen er den rigtige.

Man taler om statistisk signifikans, når testsandsynligheden p er større end et på forhånd fastsat signifikansniveau α . Ofte vil α være 1 pct., 5 pct. og i nogle tilfælde 10 pct. α angiver sandsynligheden for at forkaste hypotesen, selv om den er rigtig. Hvis og kun hvis testsandsynligheden er mindre end α , vil man forkaste hypotesen. Hvis testsandsynligheden er større end α , vil man afvise at forkaste hypotesen.

2.1.4 Lineære normale modeller

I dette afsnit præsenteres lineære normale modeller generelt for derefter at kunne præsentere to konkrete eksempler på lineære normale modeller.

Lineære normale modeller eller lineære normalfordelingsmodeller er en klasse af modeller, som behandler observationer af uafhængige normalfordelte stokastiske variable med samme varians og tester hypoteser om middelværdiparametrene.

For at vise sætningen om den lineære normale model introduceres først nogle sætninger, der skal anvendes i beviset.

Sætning 2.3 (Fordeling af aX , hvis X er gammafordelt). Hvis den stokastiske variabel X er gammafordelt med formparameter k og skalaparameter β og hvis $a > 0$, så er $Y = aX$ gammafordelt med formparameter k og skalaparameter $a\beta$.

Bevis. For at vise sætningen anvendes sætning 2.1. Først opskrives tæthedsfunktionen for X :

$$f_X(x) = \frac{1}{\Gamma(\frac{k}{2})\beta^k} x^{k-1} \exp(-x/\beta)$$

I dette tilfælde er $y = t(x) = ax$ og $x = t^{-1}(y) = \frac{1}{a}y$. Derfor er f_Y på følgende

form:

$$\begin{aligned}
 f_Y(y) &= f_X(x)|(t^{-1})'(y)| \\
 &= \frac{1}{\Gamma(\frac{k}{2})\beta^k} x^{k-1} \exp(-x/\beta) \frac{1}{a} \\
 &= \frac{1}{\Gamma(\frac{k}{2})\beta^k} \frac{1}{a} \left(\frac{y}{a}\right)^{k-1} \exp\left(-\frac{x}{a\beta}\right) \\
 &= \frac{1}{\Gamma(\frac{k}{2})(a\beta)^k} y^{k-1} \exp\left(-\frac{y}{a\beta}\right)
 \end{aligned}$$

□

Sætningen kan bruges til at omskrive en χ^2 -fordeling til en gammafordeling, da χ^2 -fordelingen er en gammafordeling. Hvis X er χ^2 -fordelt med f frihedsgrader og $a > 0$, så er aX gammafordelt med formparameter $\frac{f}{2}$ og skalaparameter $2a$, da fordelingen for X også kan udtrykkes som en gammafordeling med formparameter $\frac{f}{2}$ og skalaparameter 2 .

Sætning 2.4 (Fordeling af summen af to gammafordelinger). Hvis X_1 og X_2 er uafhængige gammafordelte stokastiske variable med formparametre k_1 og k_2 og samme skalaparameter β , så er $X_1 + X_2$ gammafordelt med formparameter $k_1 + k_2$ og skalaparameter β .

Bevis. Tæthedsfunktionen for summen Y af to uafhængige stokastiske variable X_1 og X_2 med tæthedsfunktioner f_1 og f_2 kan, når $y > 0$, generelt skrives som ifølge sætning 2.2:

$$\begin{aligned}
 f(y) &= \int_{\mathbb{R}} f_1(x)f_2(y-x)dx \\
 f(y) &= \int_0^y \frac{1}{\Gamma(k_1)\beta^{k_1}} x^{k_1-1} \exp(-x/\beta) \frac{1}{\Gamma(k_2)\beta^{k_2}} (y-x)^{k_2-1} \exp(-(y-x)/\beta) dx \\
 &= \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} \exp(-y/\beta) \int_0^y x^{k_1-1} (y-x)^{k_2-1} dx
 \end{aligned}$$

Ved at indføre substitutionen $u = x/y$ omskrives integralet:

$$\begin{aligned}
 f(y) &= \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} \exp(-y/\beta) \int_0^1 (uy)^{k_1-1} (y-uy)^{k_2-1} y du \\
 &= \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} \int_0^1 u^{k_1-1} (1-u)^{k_2-1} du \exp(-y/\beta) y^{k_1+k_2-1}
 \end{aligned}$$

Integralet, der indgår i ovenstående udtryk, er den såkaldte betafunktion, som er givet ved:

$$B(k_1, k_2) = \int_0^1 u^{k_1-1} (1-u)^{k_2-1} du, \text{ hvor } k_1, k_2 > 0$$

Betafunktionen afhænger af u og ikke af y , og derfor kan udtrykket for $f(y)$ skrives:

$$f(y) = \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} B(k_1, k_2) \exp(-y/\beta) y^{k_1+k_2-1}$$

Tæthedsfunktionen kan altså skrives som en konstant multipliceret med $\exp(-y/\beta)y^{k_1+k_2-1}$. For at få integralet af f til at summere til 1, skal der gælde følgende om konstanten:

$$\begin{aligned} \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} B(k_1, k_2) \int_0^\infty x^{k_1+k_2-1} \exp(-x/\beta) dx &= 1 \\ \Leftrightarrow \frac{1}{\Gamma(k_1)\Gamma(k_2)\beta^{k_1+k_2}} B(k_1, k_2) &= \frac{1}{\int_0^\infty x^{k_1+k_2-1} \exp(-x/\beta) dx} \end{aligned}$$

Det gælder om gammafunktionen ifølge Larsen [2006b], at

$$\int_0^\infty x^{k_1+k_2-1} \exp(-x/\beta) dx = \Gamma(k_1 + k_2) \beta^{k_1+k_2}$$

For at normaliseringskonstanten kan få integralet til at summere til 1, må der derfor gælde om betafunktionen, at

$$B(k_1, k_2) = \frac{\Gamma(k_1)\Gamma(k_2)}{\Gamma(k_1 + k_2)}$$

□

Sætningen anvendes til at vise følgende sætning:

Sætning 2.5 (Fordeling af sum af kvadrater af normalfordelte stokastiske variable). Hvis X_1, X_2, \dots, X_n er uafhængige standardnormalfordelte stokastiske variable, så er fordelingen af $Y = X_1^2 + X_2^2 + \dots + X_n^2$ en χ^2 -fordeling med n frihedsgrader.

Bevis. Først undersøges tilfældet $n = 1$. For $x > 0$ gælder følgende:

$$\begin{aligned} P(X_1^2 \leq x) &= P(-\sqrt{x} \leq X_1 \leq \sqrt{x}) \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} \phi(u) du \end{aligned}$$

Da $\phi(x)$ er en lige funktion (dvs. den er symmetrisk omkring y -aksen) er udtrykket lig:

$$\begin{aligned} &= 2 \int_0^{\sqrt{x}} \phi(u) du \\ &= 2 \int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du \end{aligned}$$

Ved at substituere med $t = u^2$ fås følgende:

$$\begin{aligned} &= 2 \int_0^x \frac{1}{\sqrt{2\pi}} \frac{1}{2t^{\frac{1}{2}}} \exp\left(-\frac{1}{2}t\right) dt \\ &= \int_0^x \frac{1}{\sqrt{2\pi}} t^{-\frac{1}{2}} \exp\left(-\frac{1}{2}t\right) dt \end{aligned}$$

Dvs., at tæthedsfunktionen for X_1^2 er:

$$f(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x\right)$$

Tæthedsfunktionen for χ^2 -fordelingen med en frihedsgrad er jf. definition 2.6:

$$f(x) = \frac{1}{\Gamma(1/2)\sqrt{2}} x^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x\right)$$

De to tæthedsfunktioner vil begge integrere til 1. Derfor må normeringskonstanterne være lig hinanden. Heraf følger, at $\Gamma(1/2) = \sqrt{\pi}$.

Sætning 2.4 siger, at summen af to uafhængige gammafordelinger også er gammafordelt med samme skalaparameter og med summen af formlparametrene som formlparameter. Da χ^2 -fordelingen er en gammafordeling, er sætningen vist. \square

χ^2 -fordelingen opstår altså som summen af kvadraterne af n standardnormalfordelte stokastiske variable. Hvis de n stokastiske variable er normalfordelte med middelværdi 0 og ens varianser σ^2 , så kaldes fordelingen af summen af deres kvadrater for en χ^2 -fordeling med skalaparameter σ^2 .

\mathbb{R}^n kan opspaltes i ortogonale underrum. Det noteres $\mathbb{R}^n = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots \oplus \mathcal{L}_n$. I den næste sætning vises det, hvordan orthogonalprojektioner af en n -dimensionalt normalfordelt stokastisk variabel er fordelt under en sådan opspaltning. Det er endvidere nødvendigt at præsentere et korollar, der dog ikke føres bevis for:

Korollar 2.1. Hvis \mathbf{X} er n -dimensionalt standardnormalfordelt og $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ er en ortonormal basis for \mathbb{R}^n , så er koordinaterne for \mathbf{X} , X_1, X_2, \dots, X_n i denne basis uafhængige og endimensionalt standardnormalfordelte.

Sætning 2.6 (Spaltningssætningen). Antag at \mathbf{X} er en n -dimensional normalfordelt stokastisk variabel med middelværdi $\mathbf{0}$ og varians $\sigma^2 I$.

Hvis $\mathbb{R}^n = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots \oplus \mathcal{L}_k$ er en ortogonal opspaltning, og $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_k$ er de tilhørende ortogonale projektioner, så er de stokastiske variable $\mathbf{p}_1 \mathbf{X}, \mathbf{p}_2 \mathbf{X}, \dots, \mathbf{p}_k \mathbf{X}$ uafhængige.

$\mathbf{p}_j \mathbf{X}$ er n -dimensionalt normalfordelt med middelværdi $\mathbf{0}$ og varians $\sigma^2 \mathbf{p}_j$.
 $\|\mathbf{p}_j \mathbf{X}\|^2$ er χ^2 -fordelt med skalaparameter σ^2 og $f_j = \dim \mathcal{L}_j$ frihedsgrader.

Bevis. Det er nok at se på tilfældet $\sigma^2 = 1$.

Man kan vælge en ortonormal basis for \mathbb{R}^n bestående af $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, hvor hver basisvektor ligger i et af underrummene L_i .

De stokastiske variable $X_i = \langle \mathbf{X}, \mathbf{e}_i \rangle, i = 1, 2, \dots, n$ er uafhængige og endimensionalt standardnormalfordelte. Det følger af korollar 2.1.

Projektionen $\mathbf{p}_j \mathbf{X}$ af \mathbf{X} ned på \mathcal{L}_j er summen af led på formen $X_i \mathbf{e}_i$ for hvilke det gælder, at $\mathbf{e}_i \in \mathcal{L}_j$. Dvs., at der er $\dim \mathcal{L}_j = f_j$ led. Den lineære afbildning \mathbf{p}_j er en projektion fra \mathbb{R}^n i \mathcal{L}_j , og den transponerede \mathbf{p}_j' er så en afbildning fra \mathcal{L}_j ind i \mathbb{R}^n . $\mathbf{p}_j \mathbf{p}_j' = I$, da det gælder for en ortogonalprojektion at den transponerede er lig den inverse. $\mathbf{p}_j \mathbf{X}$ er derfor n -dimensionelt normalfordelt med middelværdi $\mathbf{0}$ og variansmatrix I jf. definition 2.12.

$\mathbf{p}_1 \mathbf{X}, \mathbf{p}_2 \mathbf{X}, \dots, \mathbf{p}_k \mathbf{X}$ er uafhængige, fordi de hver især er funktioner af forskellige X_i -er.

$\|\mathbf{p}_j \mathbf{X}\|^2$ beregnes som summen af de X_i^2 -er for hvilke $\mathbf{e}_i \in \mathcal{L}_j$. Dvs., at der er i alt $f_j = \dim \mathcal{L}_j$ X_i^2 -er. Ifølge sætning 2.5 er summen af f_j X_i^2 -er, hvor X_i -erne er standardnormalfordelte, χ^2 -fordelt med f_j frihedsgrader.

Hvis σ^2 er forskellig fra 1, får χ^2 -fordelingen dette tal som skalaparameter. \square

Sætningen skal som nævnt anvendes til at vise sætningen om den lineære normale model. I den lineære normale model er \mathbf{y} en observation af den n -dimensionalt normalfordelte stokastiske variabel \mathbf{Y} med middelværdivektor $\boldsymbol{\mu}$ og variansmatrix $\sigma^2 I$.

Sætning 2.7 (Den lineære normale model). I den generelle lineære normale model gælder følgende:

1. Middelværdivektoren $\boldsymbol{\mu}$ estimeres ved $\hat{\boldsymbol{\mu}} = \mathbf{p}\mathbf{y}$, hvor \mathbf{p} er ortogonalprojektion ind i middelværdirummet.
2. Estimatoren $\hat{\boldsymbol{\mu}} = \mathbf{p}\mathbf{y}$ er n -dimensionalt normalfordelt med middelværdivektor $\boldsymbol{\mu}$ og varians $\sigma^2 \mathbf{p}$.

3. Variansen σ^2 estimeres centralt ved $s^2 = \frac{1}{n - \dim \mathcal{L}} \|\mathbf{y} - \mathbf{p}\mathbf{y}\|^2$. Maksimaliseringsestimatorens for σ^2 er $\hat{\sigma}^2 = \frac{1}{n} \|\mathbf{y} - \mathbf{p}\mathbf{y}\|^2$.
4. Estimatorens s^2 er χ^2 -fordelt med skalaparameter $\frac{\sigma^2}{n - \dim \mathcal{L}}$ og $n - \dim \mathcal{L}$ frihedsgrader.
5. De to estimatorens $\hat{\boldsymbol{\mu}}$ og s^2 er stokastisk uafhængige.

At variansen estimeres centralt vil sige, at middelværdien af det centrale estimat er netop variansen. Det opnås i normalfordelingsmodeller ved at dividere residualkvadratsummen med $(n - \dim \mathcal{L})$, hvor $\dim \mathcal{L}$ antallet af estimerede middelværdiparametre, i stedet for at dividere med n . $(n - \dim \mathcal{L})$ kaldes antallet af frihedsgrader for variansskønnet.

Strategien til at bevise sætning 2.7 er at anvende sætning 2.6 på den n -dimensionale stokastiske variabel $\mathbf{X} = \mathbf{Y} - \boldsymbol{\mu}$. \mathbf{X} har middelværdivektor $\mathbf{0}$ og variansmatrix $\sigma^2 I$.

Bevis. \mathbb{R}^n opspaltes i to ortogonale underrum, der betegnes \mathcal{L} og \mathcal{L}^\perp .

For at estimere middelværdivektoren $\boldsymbol{\mu}$ og variansen σ^2 opskrives likelihoodfunktionen. Det antages, at $\boldsymbol{\mu}$ er et punkt i underrummet \mathcal{L} af $\mathcal{V} = \mathbb{R}^n$, og at $\sigma^2 > 0$. Parameterrummet er dermed givet ved $\mathcal{L} \times]0; \infty[$.

Likelihoodfunktionen, der svarer til observationen \mathbf{y} , kan opskrives vha. definitionen på en regulær normalfordeling og antagelsen om variansmatrixen:

$$L(\boldsymbol{\mu}, \sigma^2) = \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \exp(-1/2 (\frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{\sigma^2})), \text{ hvor } \boldsymbol{\mu} \in \mathcal{L}, \sigma^2 > 0 \quad (2.10)$$

En vektor $\mathbf{z} \in \mathbb{R}^n$ kan udtrykkes på følgende vis vha. orthogonalprojektionens \mathbf{p} :

$$\mathbf{z} = (\mathbf{z} - \mathbf{p}\mathbf{z}) + \mathbf{p}\mathbf{z} \quad (2.11)$$

Det vil så gælde, at $(\mathbf{z} - \mathbf{p}\mathbf{z}) \perp \mathbf{p}\mathbf{z}$, da $\mathbf{z} - \mathbf{p}\mathbf{z} \in \mathcal{L}^\perp$ og $\mathbf{p}\mathbf{z} \in \mathcal{L}$. De tre vektorer i ligning 2.11 udspænder en retvinklet trekant, hvorfor Pythagoras' sætning kan anvendes:

$$\|\mathbf{z}\|^2 = \|\mathbf{z} - \mathbf{p}\mathbf{z}\|^2 + \|\mathbf{p}\mathbf{z}\|^2 \quad (2.12)$$

Hvis man i stedet for vektoren \mathbf{z} betragter vektoren $\mathbf{y} - \boldsymbol{\mu}$, så vil relationen i ligning 2.12 også gælde. Orthogonalprojektionens af $\boldsymbol{\mu}$ er blot lig $\boldsymbol{\mu}$, $\boldsymbol{\mu} \in \mathcal{L}$ jf. modellens antagelse. Derfor kan ligning 2.12 skrives:

$$\begin{aligned} \|\mathbf{y} - \boldsymbol{\mu}\|^2 &= \|\mathbf{y} - \boldsymbol{\mu} - \mathbf{p}(\mathbf{y} - \boldsymbol{\mu})\|^2 + \|\mathbf{p}(\mathbf{y} - \boldsymbol{\mu})\|^2 \\ &= \|\mathbf{y} - \mathbf{p}\mathbf{y}\|^2 + \|\mathbf{p}\mathbf{y} - \boldsymbol{\mu}\|^2 \end{aligned}$$

I likelihoodfunktionen $L(\boldsymbol{\mu}, \sigma^2)$ indgår netop udtrykket $\|\mathbf{y} - \boldsymbol{\mu}\|^2$, så det følger, at $L(\boldsymbol{\mu}, \sigma^2) \leq L(\mathbf{p}\mathbf{y}, \sigma^2)$ for alle σ^2 . Dermed vil likelihoodfunktionen antage sit maksimum, når $\boldsymbol{\mu} = \mathbf{p}\mathbf{y}$, altså estimeres middelværdivektoren ved ortogonalprojektion af observationsvektoren \mathbf{y} ned i \mathcal{L} . Hermed er det første punkt i sætningen bevist.

I forhold til det andet punkt i sætningen fortæller sætning 2.6, at $\mathbf{p}\mathbf{X}$ er n -dimensionalt normalfordelt med middelværdivektor $\mathbf{0}$ og variansmatrix $\sigma^2\mathbf{p}$. Dvs., at $\mathbf{p}\mathbf{y} - \boldsymbol{\mu}$ er n -dimensionalt normalfordelt med middelværdivektor $\mathbf{0}$ eller $\boldsymbol{\mu} = \mathbf{p}\mathbf{y}$ er n -dimensionalt normalfordelt med middelværdivektor $\boldsymbol{\mu}$ og variansmatrix $\sigma^2\mathbf{p}$.

Maksimaliseringsestimateret $\hat{\sigma}^2$ findes ved at opskrive og differentiere $\log L$ og sætte udtrykket lig nul:

$$\begin{aligned}\log L(\boldsymbol{\mu}, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} n \log(\sigma^2) - \frac{1}{2} \frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{\sigma^2} \\ &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \left(n \log(\sigma^2) + \frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{\sigma^2} \right) \\ \frac{d \log L(\boldsymbol{\mu}, \sigma^2)}{d \sigma^2} &= -\frac{1}{2} \left(n \frac{1}{\sigma^2} - \frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{(\sigma^2)^2} \right) \\ n \frac{1}{\sigma^2} - \frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{(\sigma^2)^2} &= 0 \\ n \frac{1}{\sigma^2} &= \frac{\|\mathbf{y} - \boldsymbol{\mu}\|^2}{(\sigma^2)^2} \\ \sigma^2 &= \frac{1}{n} \|\mathbf{y} - \boldsymbol{\mu}\|^2\end{aligned}$$

Dvs., at maksimaliseringsestimateret $\hat{\sigma}^2$ er $\frac{1}{n} \|\mathbf{y} - \boldsymbol{\mu}\|^2$. Det centrale estimat s^2 findes ved at dividere med $(n - \dim \mathcal{L})$ i stedet for n , dvs.

$$s^2 = \frac{1}{n - \dim \mathcal{L}} \|\mathbf{y} - \boldsymbol{\mu}\|^2$$

Det følger af sætning 2.6, at $\|\mathbf{y} - \boldsymbol{\mu}\|^2$ er χ^2 -fordelt med skalaparameter σ^2 og $\dim \mathcal{L}^\perp$ frihedsgrader. Endvidere er $\dim \mathcal{L}^\perp = n - \dim \mathcal{L}$. Derfor er s^2 χ^2 -fordelt med skalaparameter $\frac{\sigma^2}{n - \dim \mathcal{L}}$ og $n - \dim \mathcal{L}$ frihedsgrader.

Det sidste, der skal vises, er at $\hat{\boldsymbol{\mu}}$ og s^2 er stokastisk uafhængige. Som det var tilfældet i beviset for sætning 2.6 gælder det om den valgte ortonormale basis for \mathbb{R}^n at hver enhedsvektor $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ligger i et af de to underrum \mathcal{L} og \mathcal{L}^\perp . De to estimater tilhører hver sit underrum idet $\hat{\boldsymbol{\mu}} \in \mathcal{L}$ og s^2 afhænger af $\mathbf{y} - \boldsymbol{\mu} \in \mathcal{L}^\perp$. Dvs., at de er stokastisk uafhængige. \square

Sætning 2.7 gælder for alle lineære normalfordelingsmodeller og i de kommende afsnit introduceres to eksempler. Først introduceres hypotesetest i en lineær normalfordelingsmodel.

En hypotese i en sådan model handler om middelværdierne. Dvs., at en hypotese kan formuleres på følgende vis:

$$H_0 : \boldsymbol{\mu} \in \mathcal{L}_0, \text{ hvor } \mathcal{L}_0 \subseteq \mathcal{L}$$

Hypotesen går altså ud på, at estimatet for middelværdien findes i et under- rum af middelværdirummet i grundmodellen. Maksimaliseringsestimaterne under hypotesen er så givet ved $\mathfrak{p}_0\mathbf{y}$ og $\frac{1}{n} \|\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2$. Kvotientteststørrelsen Q skrives op:

$$Q = \frac{L(\mathfrak{p}_0\mathbf{y}, \frac{1}{n} \|\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2)}{L(\mathfrak{p}\mathbf{y}, \frac{1}{n} \|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2)}$$

Ved at indsætte likelihoodfunktionerne kan udtrykket kvotientteststørrelsen reduceres til følgende:

$$\begin{aligned} Q &= \left(\frac{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2} \right)^{n/2} \\ &= \left(\frac{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2 + \|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2} \right)^{n/2} \end{aligned}$$

Den sidste omskrivning følger af, at $\mathbf{y} - \mathfrak{p}\mathbf{y} \perp \mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}$ eftersom $\mathfrak{p}\mathbf{y} - P_0\mathbf{y} \in \mathcal{L}$. Q omskrives yderligere:

$$\begin{aligned} Q &= \left(\frac{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2 + \|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2} \right)^{n/2} \\ &= \left(\frac{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2 + \|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2} \right)^{-n/2} \\ &= \left(1 + \frac{\|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2} \right)^{-n/2} \\ &= \left(1 + \frac{\dim \mathcal{L} - \dim \mathcal{L}_0}{n - \dim \mathcal{L}} \frac{n - \dim \mathcal{L}}{\dim \mathcal{L} - \dim \mathcal{L}_0} \frac{\|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2}{\|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2} \right)^{-n/2} \\ &= \left(1 + \frac{\dim \mathcal{L} - \dim \mathcal{L}_0}{n - \dim \mathcal{L}} \frac{\frac{1}{\dim \mathcal{L} - \dim \mathcal{L}_0} \|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2}{\frac{1}{n - \dim \mathcal{L}} \|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2} \right)^{-n/2} \end{aligned}$$

Størrelsen $\frac{\frac{1}{\dim \mathcal{L} - \dim \mathcal{L}_0} \|\mathfrak{p}\mathbf{y} - \mathfrak{p}_0\mathbf{y}\|^2}{\frac{1}{n - \dim \mathcal{L}} \|\mathbf{y} - \mathfrak{p}\mathbf{y}\|^2}$ kaldes F . Q er en aftagende funktion af F , og hypotesen skal derfor forkastes for store værdier af F .

Ifølge sætning 2.6 er tælleren og nævneren i F stokastisk uafhængige og χ^2 -fordelte med skalaparametre $\frac{\sigma^2}{\dim \mathcal{L} - \dim \mathcal{L}_0}$ og $\frac{\sigma^2}{n - \dim \mathcal{L}}$ og henholdsvis $\dim \mathcal{L} - \dim \mathcal{L}_0$ og $n - \dim \mathcal{L}$ frihedsgrader.

F er F -fordelt med $\dim \mathcal{L} - \dim \mathcal{L}_0$ og $n - \dim \mathcal{L}$ frihedsgrader jf. definition 2.7.

Ensidet variansanalyse

Ensidet variansanalyse er et eksempel på en lineær normal model, som behandler observationer, som er inddelt i grupper på en måde som nedenstående tabel illustrerer.

Antallet af grupper betegnes k . En observation noteres y_{ij} og er det j -te element i den i -te gruppe. Den i -te gruppe består af n_i observationer, og $n = \sum_{i=1}^k n_i$ er det samlede antal observationer i alle k grupper. Der er ikke nødvendigvis samme antal observationer i hver gruppe.

Observationer						
Gruppe 1	y_{11}	y_{12}	\cdots	y_{1j}	\cdots	y_{1n_1}
Gruppe 2	y_{21}	y_{22}	\cdots	y_{2j}	\cdots	y_{2n_2}
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
Gruppe i	y_{i1}	y_{i2}	\cdots	y_{ij}	\cdots	y_{in_i}
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
Gruppe k	y_{k1}	y_{k2}	\cdots	y_{kj}	\cdots	y_{kn_k}

Grupperne er defineret ud fra en kvalitativ faktor. I ensidet variansanalyse er der én faktor, der beskriver grupperne. Det kan fx være forskelle i mængden af tilført gødning på forskellige landbrugsmarker. I tosidet variansanalyse, der ikke gennemgås her, vil man have observationer, der er kendetegnet ved to forskellige faktorer, fx gødningsniveau på den ene side, og jordbundstype på den anden side, og man vil undersøge kombinerede effekter af de to faktorer.

Grunden til, at disciplinen kaldes variansanalyse, er, at formålet er at undersøge variationen mellem grupperne i forhold til variationen inden for grupperne.

Det antages, at forskellen mellem observationerne inden for en gruppe er tilfældig, mens der er en systematisk forskel mellem grupperne.

I den i -te gruppe bestående af observationer $y_{i1}, y_{i2}, \dots, y_{in_i}$ antages det, at y_{ij} 'erne er observerede værdier af de stokastiske variable Y_{ij} , der er normalfordelte med middelværdi μ_i og varians σ^2 .

For at lave ensidet variansanalyse skal man estimere middelværdierne $\mu_1, \mu_2, \dots, \mu_k$ og variansen σ^2 . Derefter tester man hypotesen H_0 om, at middelværdierne er ens:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

For at bestemme maksimaliseringsestimateret $\hat{\boldsymbol{\mu}}$ i den ensidede variansanalysemodel ser man på likelihoodfunktionen (ligning 2.10) for fast σ^2 . Da $\boldsymbol{\mu}$ kun indgår i den negative eksponent, vil likelihoodfunktionen antage sit maksimum for den mindst mulige værdi af eksponenten. For fast σ^2 betyder det, at $\|\mathbf{y} - \boldsymbol{\mu}\|^2$ skal minimeres:

$$\begin{aligned} \|\mathbf{y} - \boldsymbol{\mu}\|^2 &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \mu_i)^2 \\ &= \sum_{j=1}^{n_1} (y_{1j} - \mu_1)^2 + \sum_{j=2}^{n_2} (y_{2j} - \mu_2)^2 + \dots + \sum_{j=1}^{n_k} (y_{kj} - \mu_k)^2 \end{aligned}$$

Opskrivningen viser, at minimering af udtrykket gøres ved at minimere hver enkelt sum. Hvis man differentierer den i -te sum og sætter den lig nul fås følgende maksimaliseringsestimater:

$$\begin{aligned} 2 \sum_{j=1}^{n_i} (y_{ij} - \mu_i) &= 0 \\ \sum_{j=1}^{n_i} y_{ij} - n_i \mu_i &= 0 \\ \mu_i &= \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} \end{aligned}$$

Dvs., at elementerne i vektoren $\hat{\boldsymbol{\mu}}$ er de enkelte gruppegennemsnit.

Maksimaliseringsestimateret $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$ og det centrale estimat findes ved at anvende $n - \dim \mathcal{L}$ frihedsgrader i stedet for n i nævneren.

$$s_0^2 = \frac{1}{n - k} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

Antallet af frihedsgrader er $n - k$, da dimensionen af \mathcal{L} i dette tilfælde er antallet af grupper k .

Hypotesen H_0 kan omskrives til $H_0 : \boldsymbol{\mu} \in \mathcal{L}_0$. Hypotesen siger, at middelværdierne kan estimeres ved en vektor med samme tal μ på alle pladser. Estimateret for $\boldsymbol{\mu}$ under hypotesen findes ved ortogonalprojektion af \mathbf{y} på \mathcal{L}_0 , som er givet ved vektoren $P_0 \mathbf{y}$ med totalgennemsnittet \bar{y} på alle pladser.

Teststørrelsen F er givet ved følgende:

$$\frac{\frac{1}{\dim \mathcal{L} - \dim \mathcal{L}_0} \|P \mathbf{y} - p_0 \mathbf{y}\|^2}{\frac{1}{n - \dim \mathcal{L}} \|\mathbf{y} - p \mathbf{y}\|^2}$$

Nævneren i brøken svarer til den centrale variansestimater i grundmodellen (s_0^2), mens tælleren i brøken svarer til det centrale variansestimater under hypotesen, nemlig $s_1^2 = \frac{1}{k-1} \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$.

Hvis hypotesen accepteres, så estimeres $\boldsymbol{\mu}$ som nævnt ved vektoren $P_0 \mathbf{y}$ med totalgennemsnittet på alle pladser. Variansen estimeres centralt ved følgende udtryk:

$$\begin{aligned} s_{01}^2 &= \frac{1}{n - \dim \mathcal{L}_0} \|\mathbf{y} - P_0 \mathbf{y}\|^2 \\ &= \frac{1}{n-1} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 \end{aligned}$$

Ækvivalent med udtrykket er følgende:

$$\begin{aligned} s_{01}^2 &= \frac{1}{n - \dim L_0} (\|\mathbf{y} - P\mathbf{y}\|^2 + \|P\mathbf{y} - P_0\mathbf{y}\|^2) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2 \right) \end{aligned}$$

Bartletts test for varianshomogenitet

I den ensidede variansanalyse er det en forudsætning, at grupperne har ens varianser. Det kaldes også varianshomogenitet eller homoskedasticitet. Hvis observationerne ikke har varianshomogenitet er der tale om heteroskedasticitet. Bartletts test, der gennemgås i det følgende, har til formål at undersøge, om der er ens varianser. Det er en forudsætning, at der er tale om en normalfordelingsmodel for at bruge Bartletts test.

Grundmodellen går ud på, at hver enkelt gruppe y_{ij} af observationer antages at være beskrevet ved egen middelværdi μ_i og egen varians σ_i^2 . Hypotesen om ens varianser kan så formuleres på følgende vis:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

Observationerne y_{ij} er som nævnt inddelt i k grupper. Den i -te gruppe består af n_i observationer. I gruppe nr. i opfattes medlemmerne $y_{i1}, y_{i2}, \dots, y_{in_i}$ som observerede værdier af de stokastiske variable $Y_{i1}, Y_{i2}, \dots, Y_{in_i}$, om hvilket det antages, at de er indbyrdes uafhængige og identisk normalfordelte med middelværdi μ_i og varians σ_i^2 .

Estimatet af σ_i^2 er den stokastiske variabel s_i^2 , som jf. sætning 2.7 vil være χ^2 -fordelt med skalaparameter σ_i^2/f_i og f_i frihedsgrader, hvor $f_i = n_i - 1$. Det svarer jf. sætning 2.3 til, at s_i^2 er gammafordelt med forparameter $f_i/2$ og skalaparameter $2\sigma_i^2/f_i$.

Likelihoodfunktionen i grundmodellen opskrives som et produkt af likelihoodfunktioner for alle grupperne ud fra gammafordelingen:

$$L(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2) = \prod_{i=1}^k \frac{1}{\Gamma(f_i/2)(2\sigma_i^2/f_i)^{f_i/2}} (s_i^2)^{f_i/2-1} \exp\left(-\frac{s_i^2}{\frac{2\sigma_i^2}{f_i}}\right)$$

For at finde maksimum for likelihoodfunktionen, omskrives den:

$$\begin{aligned} L(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2) &= \prod_{i=1}^k \left(\frac{1}{\Gamma(f_i/2)(2/f_i)^{f_i/2}} (s_i^2)^{f_i/2-1} \right) \cdot \prod_{i=1}^k (\sigma_i^2)^{-f_i/2} \exp\left(-\frac{f_i}{2} \frac{s_i^2}{\sigma_i^2}\right) \\ &= \prod_{i=1}^k \left(\frac{1}{\Gamma(f_i/2)(2/f_i)^{f_i/2}} (s_i^2)^{f_i/2-1} \right) \cdot \left(\prod_{i=1}^k (\sigma_i^2)^{-f_i/2} \right) \exp\left(-\frac{1}{2} \sum_{i=1}^k f_i \frac{s_i^2}{\sigma_i^2}\right) \end{aligned}$$

For at bestemme maksimaliseringsestimaterne bestemmes log-likelihood-funktionen ved at tage den naturlige logaritme til det fundne udtryk. Her ses der bort fra konstantleddet, hvor σ_i^2 ikke indgår, da det er irrelevant i forhold til bestemmelse af maksimaliseringsestimaterne.

$$\begin{aligned} \log L(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2) &= -\frac{1}{2} \sum_{i=1}^k f_i \log(\sigma_i^2) - \frac{1}{2} \sum_{i=1}^k f_i \frac{s_i^2}{\sigma_i^2} \\ &= -\frac{1}{2} \sum_{i=1}^k f_i \left(\log(\sigma_i^2) + \frac{s_i^2}{\sigma_i^2} \right) \end{aligned}$$

Log-likelihoodfunktionen skal differentieres med hensyn til σ_i^2 . Den i -te afledte ser således ud:

$$\frac{\partial}{\partial \sigma_i^2} \log L(\sigma_1^2, \sigma_2^2, \dots, \sigma_i^2, \dots, \sigma_k^2) = -\frac{1}{2} f_i \left(\frac{1}{\sigma_i^2} - \frac{s_i^2}{(\sigma_i^2)^2} \right)$$

Alle de afledte med hensyn til $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ antager denne form. Den differentierede log-likelihoodfunktion sættes lig nul:

$$\begin{aligned} -\frac{1}{2} f_i \left(\frac{1}{\sigma_i^2} - \frac{s_i^2}{(\sigma_i^2)^2} \right) &= 0 \\ \Leftrightarrow \sigma_i^2 &= s_i^2 \end{aligned}$$

Altså er maksimaliseringsestimaterne for variansen σ_i^2 givet ved s_i^2 , hvilket gælder for $i = 1, 2, \dots, k$.

Likelihoodfunktionen under hypotesen $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ skal nu opskrives. Det foregår ligesom før, dog er σ^2 den samme for alle grupper. Resultatet af opskrivningen giver følgende log-likelihoodfunktion differentieret med hensyn til σ^2 :

$$\frac{d}{d\sigma^2} \log L(\sigma^2, \sigma^2, \dots, \sigma^2) = -\frac{1}{2} \left(\sum_{i=1}^k f_i \right) \frac{1}{\sigma^2} + \frac{1}{2} \left(\sum_{i=1}^k f_i s_i^2 \right) \frac{1}{(\sigma^2)^2}$$

Maksimaliseringsestimaterne findes ved at sætte log-likelihoodfunktionen lig nul:

$$\left(\sum_{i=1}^k f_i\right) \frac{1}{\sigma^2} = \left(\sum_{i=1}^k f_i s_i^2\right) \frac{1}{(\sigma^2)^2}$$

$$\hat{\sigma}^2 = \frac{1}{\sum_{i=1}^k f_i} \sum_{i=1}^k f_i s_i^2$$

Det fundne maksimaliseringsestimater for σ^2 kaldes s_0^2 og kan også skrives som $\frac{1}{n-k} \sum_{i=1}^k f_i s_i^2$.

Det næste skridt er at bestemme teststørrelsen vha. kvotienttestet.

$$Q = \frac{L(s_0^2, s_0^2, \dots, s_0^2)}{L(s_1^2, s_2^2, \dots, s_k^2)}$$

Teststørrelsen $-2\log Q$ er så givet ved:

$$-2\log Q = -2(\log L(s_0^2, s_0^2, \dots, s_0^2) - \log L(s_1^2, s_2^2, \dots, s_k^2))$$

$$\begin{aligned} -2\log Q &= -2\left(-\frac{1}{2} \sum_{i=1}^k f_i \left(\log(s_0^2) + \frac{s_i^2}{s_0^2}\right) + \frac{1}{2} \sum_{i=1}^k f_i \left(\log(s_i^2) + \frac{s_i^2}{s_i^2}\right)\right) \\ &= -\left(\sum_{i=1}^k f_i (\log(s_i^2) + 1) - \sum_{i=1}^k f_i \left(\log(s_0^2) + \frac{s_i^2}{s_0^2}\right)\right) \\ &= -\left(\sum_{i=1}^k f_i \log(s_i^2) + \sum_{i=1}^k f_i - \sum_{i=1}^k f_i \log(s_0^2) - \sum_{i=1}^k f_i \frac{s_i^2}{s_0^2}\right) \end{aligned}$$

I det sidste led i ovenstående udtryk indgår s_0^2 , som er bestemt tidligere. Hvis $\frac{1}{s_0^2}$ sættes uden for sumtegnet og udtrykket for s_0^2 indsættes fås følgende:

$$\begin{aligned} &= -\left(\sum_{i=1}^k f_i (\log(s_i^2) - \log(s_0^2)) + \sum_{i=1}^k f_i - \frac{\sum_{i=1}^k f_i}{\sum_{i=1}^k f_i s_i^2} \sum_{i=1}^k f_i s_i^2\right) \\ &= -\left(\sum_{i=1}^k f_i \log\left(\frac{s_i^2}{s_0^2}\right) + \sum_{i=1}^k f_i - \sum_{i=1}^k f_i\right) \\ &= -\sum_{i=1}^k f_i \log\left(\frac{s_i^2}{s_0^2}\right) \end{aligned}$$

Resultatet kaldes Bartlett's teststørrelse:

$$B = -\sum_{i=1}^k f_i \log\left(\frac{s_i^2}{s_0^2}\right)$$

I stedet for B anvendes ofte Bartletts korrigerede teststørrelse fra Bartlett [1937], som antager formen:

$$B_{\text{korr}} = \frac{(n-k)\log(s_0^2) - \sum_{i=1}^k f_i \log(s_i^2)}{1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^k \frac{1}{f_i} - \frac{1}{n-k} \right)}$$

Det skyldes ifølge Tjur [2002b], at det kan vises, at B_{korr} er bedre approksimeret ved en χ^2 -fordeling med $k-1$ frihedsgrader end B . Hvis der er et stort antal observationer i grupperne, vil B og B_{korr} approksimativt være ens, da det sidste led i nævneren vil blive relativt lille. Ifølge Bartlett [1937] er B_{korr} en bedre approksimation, der korrigerer signifikansniveauer for meget små observationssæt.

Hvis teststørrelsen er stor, vil testsandsynligheden være lille, og hypotesen om ens varianser forkastes. I så fald er der ikke varianshomogenitet, og man kan i princippet ikke udføre ensidet variansanalyse. Ifølge Tjur [2002b] vil man dog for store datamaterialer, hvor hypotesen om ens varianser forkastes uden, at der er store relative forskelle på varianser, ofte udføre ensidet variansanalyse alligevel med opmærksomheden rettet mod en øget unøjagtighed.

Simpel lineær regression

Simpel lineær regression er et eksempel på en lineær normal model. Man har et antal sammenhørende værdier (x_i, y_i) , hvor $i = 1, 2, \dots, n$. y -værdierne er observerede værdier af de stokastiske variable Y_1, Y_2, \dots, Y_n . x -værdierne er baggrundsvARIABLE og er ikke stokastiske. Dvs., at x -værdierne er kendte eller målt uden måleusikkerhed. y -værdierne kaldes også responsvariable eller afhængige variable, da de i modellen afhænger af x -værdierne.

Det antages, at y -ernes variation omkring middelværdien kan beskrives ved en normalfordeling med samme varians σ^2 for alle observationer. Modellen går ud på at beskrive middelværdistrukturen ved $EY_i = \alpha + \beta x_i$.

For at estimere parametrene opskrives likelihoodfunktionen ud fra ligning 2.10:

$$L(\alpha, \beta, \sigma^2) = \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \prod_{i=1}^n \exp\left(-\frac{1}{2} \frac{(y_i - (\alpha + \beta x_i))^2}{\sigma^2}\right)$$

Som tidligere minimeres likelihoodfunktionen ved at minimere residualkvadratsummen $\sum_{i=1}^n (y_i - (\alpha + \beta x_i))^2$. Først opspaltes udtrykket:

$$\sum_{i=1}^n (y_i - (\alpha + \beta x_i))^2 = \sum_{i=1}^n ((y_i - \bar{y}) + (\bar{y} - (\alpha + \beta \bar{x}) - \beta(x_i - \bar{x}))^2$$

Kvadratet på den treledede størrelse bliver så:

$$\begin{aligned} \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - (\alpha + \beta\bar{x}))^2 + \beta^2 \sum_{i=1}^n (x_i - \bar{x})^2 + 2(\bar{y} - (\alpha + \beta\bar{x})) \sum_{i=1}^n (y_i - \bar{y}) \\ - 2\beta \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) - 2\beta(\bar{y} - (\alpha + \beta\bar{x})) \sum_{i=1}^n (x_i - \bar{x}) \end{aligned}$$

To af leddene indeholder enten $\sum_{i=1}^n (y_i - \bar{y})$ eller $\sum_{i=1}^n (x_i - \bar{x})$, der summerer til 0. Altså er udtrykket lig:

$$\sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - (\alpha + \beta\bar{x}))^2 + \beta^2 \sum_{i=1}^n (x_i - \bar{x})^2 - 2\beta \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$$

Det kan ses ud fra omskrivningen af residualkvadratsummen, at der kun er et led, der indeholder α . Dette led er mindst, når $\alpha = \bar{y} - \beta\bar{x}$. De resterende tre led kan betragtes som en andengradsfunktion af β . Koefficienten til andengradsleddet er større end nul, hvorfor funktionen er konveks, og der kan bestemmes et minimum.

Minimum findes ved at sætte den afledte i forhold til β lig nul. Løsningen angiver estimatet for β . Estimerterne for α og β er således:

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ \hat{\alpha} &= \bar{y} - \hat{\beta}\bar{x} \end{aligned}$$

Det centrale variansestimat med $n - 2$ frihedsgrader er:

$$s_{02}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - (\hat{\alpha} + \hat{\beta}x_i))^2$$

Modelkontrol i en simpel lineær regressionsmodel kan udføres ved at indlejre modellen som en hypotese i en større model. Hvis der til hver x -værdi er flere tilhørende y -værdier kan denne større model være ensidet variansanalyse, hvor man opfatter y -værdierne til et x som en gruppe, der er karakteriseret ved middelværdien. Variansanalysemodellen kan skrives som $Y_{ij} \sim \mathcal{N}(\mu_i, \sigma^2)$, hvor der er k grupper med samme varians σ^2 , og hvor μ_i er middelværdien for den i -te gruppe. Den lineære regressionsmodel kan skrives som $Y_{ij} \sim \mathcal{N}(\alpha + \beta x_i, \sigma^2)$.

Den hypotese, der testes, er $H_2: \mu_i = \alpha + \beta x_i$. Den testes vha. en kvotient F mellem to centrale variansestimater s_2^2 og s_0^2 :

$$s_2^2 = \frac{1}{k-2} \sum_{i=1}^k n_i (\bar{y}_i - (\hat{\alpha} + \hat{\beta}x_i))^2$$

$$s_0^2 = \frac{1}{n-k} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

De to centrale variansestimater fremkommer ved at opsplitte residualkvadratsummen på følgende vis:

$$\begin{aligned} & \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - (\hat{\alpha} + \hat{\beta}x_i))^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^k n_i (\bar{y}_i - (\hat{\alpha} + \hat{\beta}x_i))^2 \end{aligned}$$

Teststørrelsen er:

$$F = \frac{s_2^2}{s_0^2}$$

F er F -fordelt med $k-2$ og $n-k$ frihedsgrader.

Determinationskoefficienten

Et ofte anvendt mål i forbindelse med den lineære regressionsmodel er determinationskoefficienten R^2 .

Determinationskoefficienten R^2 anvendes ofte som modelkontrol i lineær regression. R^2 kan beregnes ved en af følgende to formler:

$$\begin{aligned} R^2 &= \frac{(\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{y}))^2}{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2} \\ R^2 &= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \end{aligned}$$

Den sidste formel udtrykker hvor stor en del af den samlede variation omkring totalgennemsnittet \bar{y} , der beskrives af modellen. Derfor kaldes R^2 undertiden for forklaringsgraden og en lineær regressionsmodel med en R^2 -værdi på fx 0.85 siges at forklare 85% af variationen i data. Et berømt konstrueret datasæt af F. J. Anscombe fra 1973 viser dog, at man kan opnå samme R^2 -værdi for vidt forskellige situationer. Derfor kan det være problematisk at lade R^2 -værdien fungere alene som modelkontrol.

Fraktildiagrammer

Hvis man antager, at y -værdierne er observerede værdier af en normalfordelt stokastisk variabel Y med middelværdi μ og varians σ^2 , så kan man beregne de teoretiske fraktiler og sammenligne dem med de empiriske fraktiler. Resultatet kaldes et fraktildiagram og anvendes til at vurdere, om data er normalfordelt. Hvis data er normalfordelt skal punkterne ligge på en ret linje.

2.1.5 Opsamling

I det foregående afsnit blev baggrunden for lineære normalfordelingsmodeller beskrevet. Ensidet variansanalyse og simpel lineær regression er to eksempler på sådanne, som skal anvendes i behandlingen af datasættet samt til at vurdere modellernes resultater.

2.2 Remote sensing

Afsnittet gennemgår de to typer af modeller, der er nævnt i indledningen. Den empirisk-statistiske model introduceres ud fra udvalgte spektrale vegetationsindekser og den fysisk-baserede model PROSAIL præsenteres. Indledningsvis gennemgås en række relevante begreber.

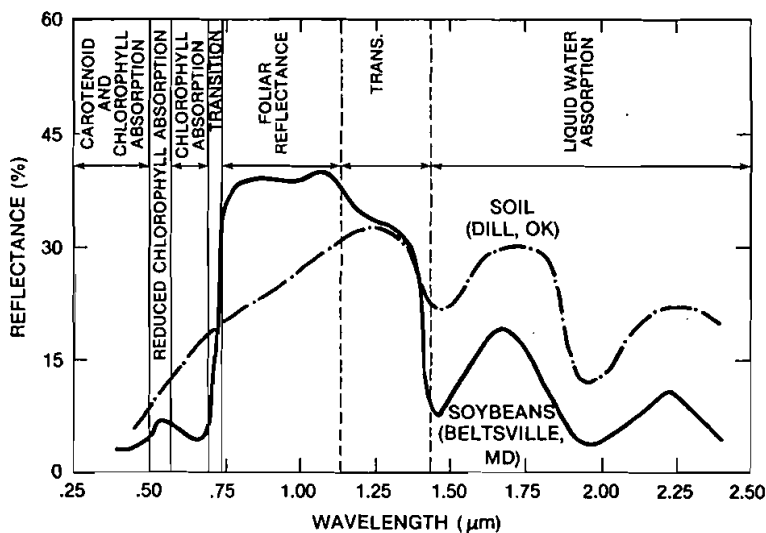
2.2.1 Begreber i remote sensing af vegetation

For at forstå de to forskellige typer af modeller, der anvendes til estimation af vegetationsforhold med remote sensing-data, introduceres her nogle overordnede begreber og koncepter. Først og fremmest præsenteres det spektrale signal fra planter. Derefter præsenteres Beers lov, der anvendes i mange remote sensing-analyser. Til slut gives en oversigt over vigtige radiometriske størrelser og deres enheder.

Reflektans fra planter

De grønne planter med klorofylholdige blade absorberer lys fra solen til at lave fotosyntetisk aktivitet. Bladene absorberer rødt og blå lys og reflekterer en del af det grønne lys. På figur 2.1 kan man se, hvilke egenskaber ved de grønne blade, der reflekterer stråling i den relevante del af det elektromagnetiske spektrum fra 400 nm til 2500 nm. Man kan se, at reflektansen af det blå lys er lav pga. bladenes indhold af klorofyl og karotenoider. Omkring det grønne lys reflekteres en større andel af strålingen som følge af egenskaber ved klorofyl. Det røde lys bliver i højere grad absorberet som følge af klorofyl, og som nævnt er der en høj reflektans i det nærinfrarøde spektrale interval. I den melleminfrarøde del af spektret sker der nogle dyk, der skyldes vandindholdet i planten [Richards og Jia, 2006]. På figuren ses også jordens reflektans, som er monotont voksende over det synlige lys og den nærinfrarøde stråling frem til de melleminfrarøde bølgelængder, hvor jordens fugtindhold giver anledning til nogle dyk i reflektansen.

Den spektrale signatur fra planter er altså kendetegnet ved, at de grønne, klorofylholdige blade reflekterer en stor del af den nærinfrarøde stråling, mens særligt rødt og blå lys absorberes. Netop den spektrale signatur er central i remote sensing, da den kan anvendes til at kortlægge vegetation.



Figur 2.1: Figuren viser, hvordan grønne blade og jord vil reflektere stråling afhængig af bølgelængden. På figuren kan man se, hvilken biofysisk egenskab ved planten, der kontrollerer reflektansen i det givne spektrale interval. [Tucker og Sellers, 1986]

Planters reflektansegenskaber afhænger af, hvor grønne de er. På figur 2.2 kan man se, at løvfældende træer reflekterer mest stråling i de grønne og nærinfrarøde dele af spektret. Akvatiske planter har en mindre karakteristisk kurve, mens nåletræer udsender en lavere andel af både grøn og nærinfrarød stråling. Samtidig kan man se, at de planter, der reflekterer mest stråling i den nærinfrarøde del af spektret, absorberer mest stråling i den røde del af spektret.

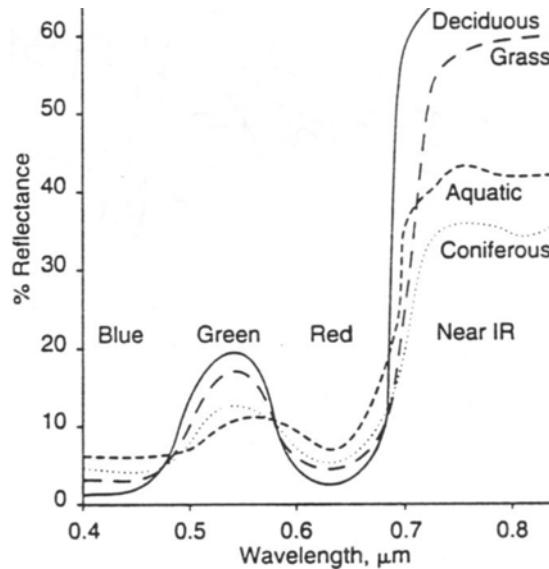
Bladarealindekset, der forkortes LAI fra engelsk *leaf area index*, er en dimensionsløs størrelse, der angiver mængden af blade i forhold til underlaget. Man kan betragte enheden som areal af blade (m^2) pr. enhed grundareal (også m^2).

Over en vækstsæson udvikler LAI sig fra 0, dvs. ingen blade, til en maksimal værdi. Typiske maksimale værdier af LAI er 6-8 for løvfældende skove og 2-4 for enårige afgrøder som korn og kartofler. [Beadle, 1993]

Beers lov

Beers lov anvendes ofte i remote sensing-analyser til at beskrive udslukningen af stråling i en plantestand.

Beers lov beskriver, hvordan en lysstråle I absorberes som funktion af dybden L i et homogent absorberende medie. Den kaldes også Lambert-Beers lov og er udviklet til at beskrive, hvordan elektromagnetisk stråling svækkes, når



Figur 2.2: Spektral reflektans for forskellige typer af vegetation [*International Center for Remote Sensing Education, 2011*].

det strømmer gennem væsker, der kun absorberer og ikke spreder strålingen. Ændringen dI/dL i lysintensiteten med dybden L afhænger af udslukningskoefficienten a og lysintensiteten I :

$$\frac{dI}{dL} = -aI$$

Løsningen på differentialligningen har formen:

$$I = I_0 \exp(-aL)$$

Her er I_0 lysintensiteten før gennemstrømningen, og I er lysintensiteten i dybden L [Den Store Danske, 2014b].

Monsi og Saeki bruger i 1953 Beers lov som en model for udslukningen af stråling i en plantestand [Rosenberg, 1974].

$$I = I_0 \exp(-kL)$$

Her er I_0 lysintensiteten over plantestanden. L er et mål for dybden i plantestanden, og er altså i dette tilfælde en form for akkumuleret LAI. I er dermed lysintensiteten til dybden LAI. k er udslukningskoefficienten som før.

Modellen antager, at plantestanden er fuldstændig homogent. Planter vil dog aldrig have en homogen struktur. Modellen antager også, at plantestanden udelukkende absorberer stråling. I stedet sker der både absorption, spredning og refleksion af stråling i en plantestand. Endvidere vil k ikke være en konstant, men snarere en funktion af en lang række forhold.

Navn	Symbol	Størrelse	Enhed
Energi	–	–	J
Effekt	–	Energi pr. tid	W=J/s
Irradians	E	Indstrålingseffekt pr. areal	W/m ²
Emittans	M	Udstrålingseffekt pr. areal	W/m ²
Radians	L	Lysstyrke	W/m ² /sr
Reflektans	ρ	Udstråling ift. indstråling	–

Tabel 2.1: Tabellen viser en række udvalgte radiometriske størrelser jf. Schaepman-Strub et al. [2006].

Radiometriske størrelser

I det følgende beskrives en række radiometriske størrelser som angivet i tabel 2.1. Denne gennemgang er inspireret af Schaepman-Strub et al. [2006].

En flux er den rate, hvormed energi overføres fra en flade til en anden. Irradians og emittans er eksempler på fluxe. Irradians eller bestrålingsstyrke angiver den indfaldende flux på en overflade pr. overfladearealenhed. Irradiansen er et mål for hvor meget stråling, der rammer overfladen. Enheden er W/m². På samme måde kan man definere emittans, dvs. hvor meget stråling, der forlader overfladen.

Radiansen er et mål for, hvor lys en genstand ser ud. Radiansen afhænger af afstanden til genstanden. Her angives derfor effekt pr. arealenhed vinkelret på strålen i forhold til rumvinklen. Enheden for radians er W/m²/sr, hvor sr er steradian, altså enheden for rumvinklen.

Reflektansen ρ er en dimensionsløs størrelse, der fremkommer ved forholdet mellem emittans og irradians. Man kan også tale om spektral reflektans angivet ρ_λ , der henviser til reflektansen for en givet bølgelængde λ .

$$\rho = \frac{M}{E} \text{ og } \rho_\lambda = \frac{M_\lambda}{E_\lambda}$$

En flade kaldes lambertsk, hvis den har samme radians i alle retninger. Hvis en flade er lambertsk vil følgende forhold gælde om reflektansen [Bird, 1991]:

$$\rho_\lambda = \frac{\pi L_\lambda}{E_\lambda}$$

Dvs., at emittansen kan skrives som π multipliceret med radiansen. Hvis man fastholder antagelsen om, at fladen er lambertsk og i øvrigt ser bort fra atmosfæriske effekter, kan man netop ifølge Bird [1991] opskrive forholdet:

$$M_\lambda = \pi L_\lambda$$

2.2.2 Spektrale vegetationsindekser

Afsnittets formål er at beskrive den teoretiske og konceptuelle baggrund for de spektrale vegetationsindekser, der anvendes i undersøgelsen. Der findes en lang række vegetationsindekser med forskellige egenskaber. Fælles for vegetationsindekserne er, at de er dimensionsløse størrelser baseret på forholdet mellem reflektans af synligt lys og nærinfrarød stråling. Vegetationsindekserne kan også være baseret på linearkombinationer af disse forhold. De kan anvendes til at beskrive en relativ vækst i grøn vegetation i et område [Huete et al., 1994]. I rapporten er der udvalgt fem indekser, som alle er baseret på rød eller grøn reflektans i kombination med nærinfrarød reflektans, eftersom der ikke er data til at understøtte indekser baseret på eksempelvis blå eller *red edge* reflektans. De udvalgte vegetationsindekser er SR, NDVI, EVI2, GNDVI og CI, som bliver gennemgået i de følgende afsnit i nævnte rækkefølge. De samme vegetationsindekser indgår i Boegh et al. [2013], hvorfor resultaterne senere kan sammenlignes.

SR og NDVI

SR (*simple ratio*) og NDVI (*normalized difference vegetation index*) er givet ved følgende udtryk:

$$SR = \frac{\rho_{NIR}}{\rho_{rød}} \qquad NDVI = \frac{\rho_{NIR} - \rho_{rød}}{\rho_{NIR} + \rho_{rød}}$$

hvor $\rho_{rød}$ og ρ_{NIR} er rød og nærinfrarød reflektans.

Spektrale vegetationsindekser som SR og NDVI anvender forholdet mellem refleksion af stråling i røde og nærinfrarøde bølgelængder til at give en indikation af vegetationen i et givent område. Som figur 2.2 viser, er netop forholdet mellem refleksion af rød og infrarød stråling karakteristisk for de forskellige typer vegetation. Ulempen ved SR er, at brøkens værdi ved små mængder rød stråling bliver meget stor. Derfor virker SR dårligt i områder med tæt vegetation [Huete et al., 1999]. NDVI er en normalisering af SR og antager værdier mellem -1 og 1 . På figur 2.1 kan man se, at ρ_{NIR} for vand vil være 0 . I det tilfælde vil NDVI blive -1 . For bar jord vil den røde og nærinfrarøde være tilnærmelsesvis lig hinanden, og NDVI vil antage værdien 0 . Hvis den nærinfrarøde reflektans overstiger den røde reflektans, som det er tilfældet ved vegetation, vil NDVI antage positive værdier. Ifølge Huete et al. [1999] vil NDVI typisk være mellem 0 og 0.8 for vegetation.

Da NDVI er normalisering af SR, kan NDVI udtrykkes vha. SR:

$$\begin{aligned}
 NDVI &= \frac{\rho_{NIR} - \rho_{rød}}{\rho_{NIR} + \rho_{rød}} \\
 &= \frac{\frac{\rho_{NIR}}{\rho_{rød}} - 1}{\frac{\rho_{NIR}}{\rho_{rød}} + 1} \\
 &= \frac{SR - 1}{SR + 1}
 \end{aligned}$$

Som nævnt er det oplagt, at der er en empirisk sammenhæng mellem vegetationsindekser og biofysiske variable som LAI. Ifølge Huete et al. [1999] sker måtningen typisk ved værdier af LAI over 2 eller 3.

SAVI, EVI og EVI2

EVI er en forkortelse for *enhanced vegetation index* og som navnet antyder, er vegetationsindekset baseret på en hensigt om at forbedre NDVI. Både SR og NDVI har begrænsninger i form af følsomhed over for refleksion fra jordbunden og atmosfærisk støj.

Huete [1988] indfører *soil-adjusted vegetation index* (SAVI) for at forbedre påvirkningen fra jordbunden på de spektrale signaler.

Jordbundens spektrale signal påvirkes af en række faktorer. En væsentlig faktor er jordens tekstur, dvs. de relative andele af ler, silt og sand. Ler består af jordpartikler under 0.002 mm i diameter, silt af partikler mellem 0.002 og 0.05 mm, mens sand består af partikler mellem 0.05 og 2.0 mm. Størrelsen på jordpartiklerne er afgørende for jordbundens indhold af vand. Jo større partikler, des større hulrum mellem partiklerne. Lerjord vil være tæt pakket, og vand vil drænes og fordampes i mindre grad end ved en sandjord. En sandjord vil derimod kunne afgive en stor del fugt ved lufttørring. Når jorden bliver våd, falder reflektansen af synligt lys. Det ses ved, at jordbunden bliver mørkere i farven. En allerede fugtig jord vil dog ikke blive tilsvarende mørkere af at blive våd. Teksturen påvirker også overfladen. Mindre partikler vil give en glattere overflade, hvilket vil øge reflektansen. Indholdet af organisk materiale påvirker også reflektansen. Generelt vil reflektansen falde ved stigende indhold af organisk materiale. Visuelt ser man, at jorde med højt indhold af organisk materiale ser mørke ud. Graden af dekompostering giver dog en farveforskel. Endelig har indholdet af jernoxid indflydelse på reflektansen. Jernoxid vil give jorden en rødlig farve og en lavere reflektans af synligt lys [Hoffer, 1978].

Når rødt lys rammer et beplantet område, vil det absorberes stærkt af de øverste grønne blade i plantestanden. Jordbunden bliver kun ramt af rødt lys de steder, hvor den er blottet.

Den nærinfrarøde stråling bliver i høj grad reflekteret af planterne. Det betyder også, at den bliver spredt og transmitteret ned på jordbunden og hen på andre planter. Jf. figur 2.1 vil jordbunden reflektere en stor del af den

nærinfrarøde stråling tilbage til sensoren. På den måde forstyrrer jordbunden signalet fra vegetationen, da både jordbund og vegetation reflekterer nærinfrarød stråling. Resultatet er støj i vegetationsindekset, da man ikke kan skelne om den nærinfrarøde stråling reflekteres af jord eller vegetation.

Ifølge Huete [1988] er problemet størst ved et vegetationsdække på cirka 50 %. Hvis der er en høj vegetationstæthed, vil signalet fra jordbunden ikke være stærkt nok til at forstyrre vegetationsindekset. Hvis der er en lav vegetationstæthed, vil der ikke være nok planter til at sprede og transmittere nærinfrarød stråling, og jordbunden vil ikke skabe støj. Hvis vegetationsdækket er omkring 50 % vil jordbunden til gengæld kunne skabe et spektralt signal, der ligner vegetation.

SAVI er således justeret for jordbundens indflydelse og er udviklet ud fra NDVI ved at tilføje en parameter l :

$$\text{SAVI} = (1 + l) \cdot \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{\rho_{\text{NIR}} + \rho_{\text{rød}} + l}$$

På figur 2.3 er isolinjer fra vegetationsindekserne NDVI og PVI illustreret. Isolinjer er linjer langs hvilke vegetationsindekset er konstant. PVI er et andet eksempel på et vegetationsindeks og behandles ikke yderligere her. For NDVI er isolinjerne kendetegnet ved, at de alle går igennem punktet $(0, 0)$, men har forskellig hældning. Isolinjen $\text{NDVI} = 0$ kan kaldes "jord-linjen", da det er netop disse forhold mellem rød og nærinfrarød stråling, der ifølge NDVI svarer til bar jord.

Hvis man har et område med delvis vegetation svarende til $\text{NDVI} = 0.33$ vil det ligge et sted på isolinjen med hældning 2. Punktet A er et sådant sted. Hvis jordbunden bliver våd, vil reflektansen ændres. Hvis NDVI ikke var følsom over for jordbundens tilstand, burde punktet A i denne situation flytte sig til B , dvs. bevæge sig langs isolinjen, da NDVI bevares på 0.33. Eksperimentel data viser ifølge Huete [1988], at NDVI for våd jord bliver højere end forventet. På figuren svarer det til, at punktet A flyttes til placering D , der ligger over isolinjen for $\text{NDVI}=0.33$.

Hvis man lader en linje gå gennem A og D vil linjen ikke gå gennem $(0, 0)$, men i stedet skære y -aksen. Denne linje vil skære isolinjen for bar jord ($\text{NDVI}=0$) i et punkt E i 3. kvadrant. Koordinatsættet for E er $(-l_1, -l_2)$.

Forskydningen af $(0, 0)$ til $(-l_1, -l_2)$ sættes ind i formelen for NDVI:

$$\frac{(\rho_{\text{NIR}} + l_2) - (\rho_{\text{rød}} + l_1)}{(\rho_{\text{NIR}} + l_2) + (\rho_{\text{rød}} + l_1)}$$

Hældningen for isolinjen for jord ($\text{NDVI}=0$) er tæt på 1, hvorfor Huete [1988] vælger $l_1 = l_2$ og $L = l_1 + l_2$, så udtrykket bliver:

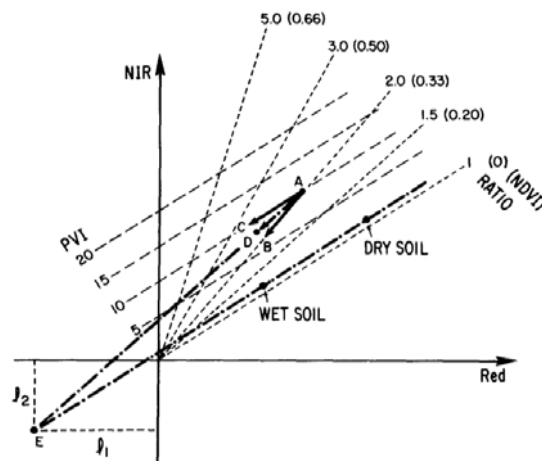
$$\frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{\rho_{\text{NIR}} + \rho_{\text{rød}} + L}$$

Som nævnt antager NDVI værdier mellem -1 og 1 , og denne begrænsning er en fordel i forhold til fx SR, der kan antage meget store værdier. Derfor tilføjes en faktor for at fastholde, at det nye indeks har en øvre og nedre grænse:

$$\text{SAVI} = (1 + L) \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{\rho_{\text{NIR}} + \rho_{\text{rød}} + L}$$

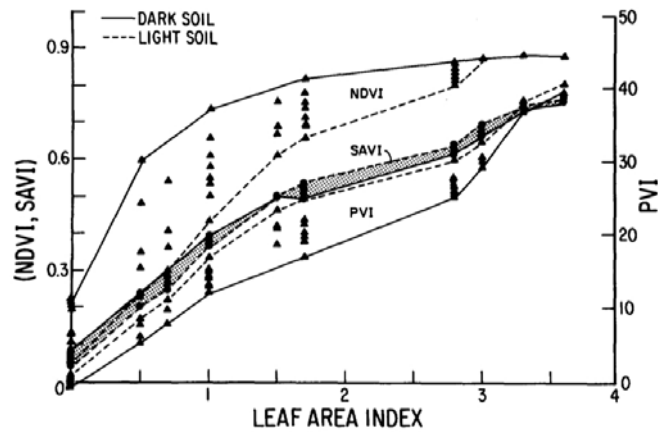
På den måde konstrueres SAVI ud fra NDVI.

Huete [1988] foretager en empirisk bestemmelse af L . Det viser sig, at $L = 1$ ved LAI tæt på 0 næsten fjerner effekten fra jorden. Ved LAI på 1 skal $L = 0.75$, da højere L giver effekt fra jorden. Hvis LAI er højere, skal L være tilsvarende lavere. Det konkluderes dog, at for alle værdier af L mellem 0.25 og 1, giver SAVI en forbedring i forhold til NDVI. SAVI har et mere lineært forhold til LAI, mens forholdet mellem NDVI og LAI er meget afhængigt af jordbunden. På figur 2.4 ses forholdene mellem henholdsvis NDVI, SAVI og LAI for både mørk og lys jord. Man kan se, at NDVI ved mørk jord har en stærk tendens til at blive mættet ved højere værdier af LAI. Forholdet mellem NDVI og LAI ved lys jord er mere lineært. Det tyder netop på den stærke afhængighed af jordbunden. SAVI har til gengæld et mere lineært forhold til LAI for både mørk og lys jord.



Figur 2.3: Figuren viser isolinjer langs hvilke vegetationsindekserne NDVI og PVI er konstante [Huete, 1988].

Det viser sig, at SAVI har et mere lineært forhold til LAI både for mørk og lys jord i forhold til NDVI. Efter udviklingen af SAVI er der foreslået en række modifikationer heraf [Jiang et al., 2008]. Et andet problem med NDVI er påvirkningen fra aerosoler. Her er der også udviklet alternative vegetationsindekser bl.a. ARVI (*atmospherically resistant vegetation index*).



Figur 2.4: Figuren viser, hvordan NDVI og SAVI forholder sig til LAI. Figur fra Huete [1988].

ARVI bruger forskellen mellem rød og blå reflektans til at rense signalet i de røde bølgelængder for effekten fra aerosoler [Jiang et al., 2008].

EVI er en forkortelse for *enhanced vegetation index*. EVI er en modifikation af NDVI, som er udviklet ud fra både SAVI og ARVI. Dermed indeholder EVI en faktor L , der korrigerer for jordbundens indfyldelse, og den blå reflektans er inkluderet for at fjerne atmosfærisk støj [Jiang et al., 2008].

$$\begin{aligned} \text{EVI} &= G \cdot \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{L + \rho_{\text{NIR}} + C_1 \rho_{\text{rød}} - C_2 \rho_{\text{blå}}} \\ &= 2,5 \cdot \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{1 + \rho_{\text{NIR}} + 6\rho_{\text{rød}} - 7,5\rho_{\text{blå}}} \end{aligned}$$

I anden linje ses de typisk anvendte parameterværdier for EVIs konstanter. Parametrene er empirisk bestemt [Huete et al., 1999].

Ifølge Boegh et al. [2002]; Houborg et al. [2007] har EVI en øget følsomhed over for store værdier af LAI. EVIs egenskaber minder således om SAVIs.

Der er udviklet en udgave af EVI, EVI2, som ikke inkluderer det blå bånd. Formålet med det blå bånd i EVI at fjerne støj fra aerosoler. Det blå bånd bidrager derfor ikke med ny viden om vegetationen, da denne information allerede er indeholdt i det røde bånd. På den baggrund udvikler Jiang et al. [2008] EVI2 ud fra EVI. EVI2 er fortsat tæt relateret til EVI.

$$\text{EVI2} = 2,5 \cdot \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{1 + \rho_{\text{NIR}} + 2,4\rho_{\text{rød}}}$$

I Boegh et al. [2013] anvendes følgende variant af EVI2, som er identisk med SAVI, for $L = 1$.

$$\text{EVI2} = 2 \cdot \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{1 + \rho_{\text{NIR}} + \rho_{\text{rød}}}$$

Det grønne NDVI

Det grønne NDVI (GNDVI) er udviklet af Gitelson et al. [1996] og anvender som navnet antyder reflektansen i det grønne frekvensbånd frem for det røde:

$$\text{GNDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{grøn}}}{\rho_{\text{NIR}} + \rho_{\text{grøn}}}$$

Ifølge Gitelson et al. [1996] er korrelationen mellem reflektans og klorofylkoncentration bedre for bølgelængden 550 nm, som er midt i den grønne del af det synlige lysspektrum, end for bølgelængden 675 nm i den røde del af spektret. Det oprindelige NDVI har en tendens til at blive mættet for meget lave klorofylkoncentrationer, hvor det GNDVI forventeligt er mere følsomt over for klorofyl. Gitelson et al. [1996] finder, at GNDVI er korreleret med plantestandens klorofylindhold.

Klorofylindekset

Klorofylindekset er et spektralt vegetationsindeks med det formål at estimere det totale klorofylindhold i en plantestand og er udviklet af Gitelson et al. [2005].

Modellen tager udgangspunkt i den såkaldte "uendelige reflektans", R_{∞} . Man betragter reflektansen fra et materiale bestående af et antal tynde homogene lag. Jo flere lag, desto højere reflektans. R_{∞} er den reflektans, hvor tilføjelse af flere lag ikke giver udslag i en betydeligt større reflektans. Parallelt hermed kan R_{∞} betragtes som grænseværdien af reflektansen i en given bølgelængde, hvor en forøgelse af bladets tykkelse ikke længere påvirker reflektansen. Gitelson et al. [2003] angiver følgende funktionsudtryk vedrørende R_{∞} :

$$f(R_{\infty}) = \frac{1 - R_{\infty}^2}{2R_{\infty}} = \frac{k}{s}$$

Her er k absorptionskoefficienten og s spredningskoefficienten for det pågældende materiale. Udtrykket kan også ifølge Gitelson et al. [2003] opskrives vha. målt reflektans R og transmittans T :

$$f(R_{\infty}) = \frac{1 + R^2 - T^2}{2R} - 1$$

Tælleren bestemmes empirisk til at være tæt på 1, hvorfor $f(R_{\infty})$ tilnærmelsesvist er proportional med den hyperbolske funktion $\frac{1}{R}$. Dvs., at der tilnærmelsesvis gælder, at:

$$\frac{1}{R} \propto \frac{k}{s} \quad (2.13)$$

Hvis absorptionskoefficienten er stor i forhold til spredningskoefficienten, så vil reflektansen altså være lav. Omvendt vil en lille absorptionskoefficient i

forhold til spredningskoefficienten give udslag i en høj reflektans. Forholdet er afhængigt af bølgelængden jf. figur 2.1 på side 35. I det synlige lys vil reflektansen være lav som følge af bladpigmenternes absorption, mens reflektansen er høj i den nærinfrarøde del af spektret, hvor bladene reflekterer strålingen.

Gitelson et al. [2005] bruger dette resultat til at udvikle et klorofylfølsomt vegetationsindeks. Først deles k op i to bidrag: k_{chl} er absorptionskoefficienten for klorofyl, og k_0 er absorptionskoefficienten for øvrige pigmenter. Strategien er at isolere k_{chl} i forholdet i ligning 2.13. Først udvælges to bølgelængder λ_1 og λ_2 , hvorom det skal gælde, at den klorofylstyrede absorption er høj i λ_1 og forsvindende lille i λ_2 . Samtidig skal absorptionen som følge af andre pigmenter end klorofyl være tilnærmelsesvis ens i de to bølgelængder. Så betragtes differensen mellem de reciprokke reflektanser i de to bølgelængder:

$$\frac{1}{R(\lambda_1)} - \frac{1}{R(\lambda_2)} \propto \frac{k_{chl}(\lambda_1) + k_0(\lambda_1)}{s} - \frac{k_{chl}(\lambda_2) + k_0(\lambda_2)}{s} \approx \frac{k_{chl}(\lambda_1)}{s}$$

Det antages desuden, at spredningskoefficienten s er den samme i hele spektret. For at isolere $k_{chl}(\lambda_1)$ anvendes endnu en bølgelængde λ_3 , for hvilken det skal gælde, at reflektansen er helt styret af s :

$$\left(\frac{1}{R(\lambda_1)} - \frac{1}{R(\lambda_2)} \right) R(\lambda_3) \propto k_{chl}$$

Gitelson et al. [2005] bestemmer λ_1 , λ_2 og λ_3 empirisk, og finder, at $\lambda_2 = \lambda_3 = 800\text{nm}$, dvs. i den nærinfrarøde del af spektret, mens der er to muligheder for λ_1 , der kan vælges til ca. 550 nm, dvs. grønt lys, eller mellem 700-730 nm, som er rødt lys på kanten af det nærinfrarøde. Klorofylindekset CI , der bruger det grønne lys, kan så opskrives:

$$CI = \frac{\rho_{\text{NIR}}}{\rho_{\text{grøn}}} - 1$$

Det andet klorofylindeks anvender det røde lys tæt på de nærinfrarøde bølgelængder. På engelsk betegnes dette frekvensbånd som *red edge*.

Forhold mellem vegetationsindekser og bladarealindeks

Ifølge Baret og Guyot [1991] vil vegetationsindekserne nærme sig et mætningspunkt for værdier af LAI større end en tærskelværdi, som kan være mellem 2 og 6 afhængig af vegetationsindeks, afgrødetype og andre forhold.

Baret og Guyot [1991] beskriver det asymptotiske forhold mellem de spektrale vegetationsindekser og bladarealindekset med følgende "semi-empiriske model" baseret på Beers lov (se afsnit 2.2.1):

$$VI = VI_{\text{max}} + (VI_{\text{min}} - VI_{\text{max}}) \exp(-K \cdot LAI)$$

Her er VI et spektralt vegetationsindeks. VI_{\max} er grænseværdien for vegetationsindekset, når LAI går mod uendelig. Ifølge Baret og Guyot [1991] er det i praksis værdier af LAI over 8. VI_{\min} er værdien af vegetationsindekset for bar jord. Når LAI er 0, vil udtrykket for vegetationsindekset blive netop VI_{\min} . K er udslukningskoefficienten som i Beers lov og afhænger af vegetationsindekset.

2.2.3 PROSAIL

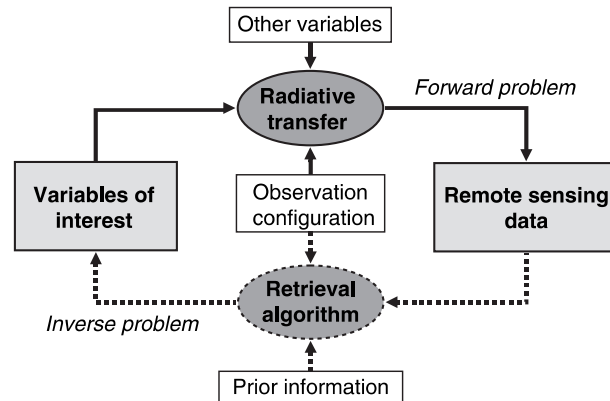
I dette afsnit præsenteres modellen PROSAIL, som beskrives i Jacquemoud et al. [2009]. PROSAIL er en strålingstransportmodel for en plantestand. Den består af to modeller, SAIL (*Scattering from Arbitrarily Inclined Leaves*) og PROSPECT. SAIL er en model for plantestandens spredning og refleksion af stråling, mens PROSPECT er en model for bladenes optiske egenskaber. SAIL er udviklet af Verhoef [1984, 1985] og PROSPECT af Jacquemoud og Baret [1990]. I begyndelsen af 1990'erne blev modellerne koblet sammen under navnet PROSAIL [Jacquemoud et al., 2009]. PROSAIL er implementeret i MATLAB [Jacquemoud et al., 2013].

Strålingstransportmodeller kaldes på engelsk *radiative transfer models*. Disse modeller er baseret på fysiske love og beskriver interaktionen mellem stråling og planter i plantestanden, dvs., hvordan elektromagnetisk stråling absorberes, reflekteres og spredes i plantestanden. Modellerne tager højde for indstrålingsvinkler, mængden af blade, fordelingen af bladenes vinkler, plantestandens arkitektur, jordbundens reflektans og bladenes biokemiske bestanddele. Der er altså et stort antal parametre, som har indflydelse på den spektrale reflektans, og som modellerne tager højde for. Den spektrale reflektans måles dog typisk multispektralt, hvorfor antallet af målinger af reflektans er mindre end antallet af parametre, der bestemmer det [Goel, 1988]. I nedenstående citat er denne problematik beskrevet:

”Estimating biophysical variables from remote sensing measurements is often an under-determined problem: the number of unknowns is generally larger than the number of independent radiometric information remotely sampled by sensors.” [Baret og Buis, 2008]

Ifølge Goel [1988] var forskningen i strålingstransportmodeller de første år fokuseret på at forstå det direkte problem for at kunne løse det inverse problem.

På figur 2.5 er sammenhængen mellem det direkte problem (fuldt optrukne linjer) og det inverse problem (stiplede linjer) illustreret. I det direkte problem har man nogle variable af interesse fx bladarealindeks eller klorofylindhold. Sammen med andre variable er de input i en strålingstransportmodel og resultatet er simuleret remote sensing data. I det inverse problem har man remote



Figur 2.5: Figuren illustrerer det direkte og inverse problem i strålingstransportmodeller. [Baret og Buis, 2008]

sensing data og gennem modellen simuleres de interessante variable. Der kræves også a priori viden, når modellen inverteres. Det skyldes jf. ovennævnte citat, at antallet af ubekendte overstiger antallet af uafhængige inputvariable i inversionen.

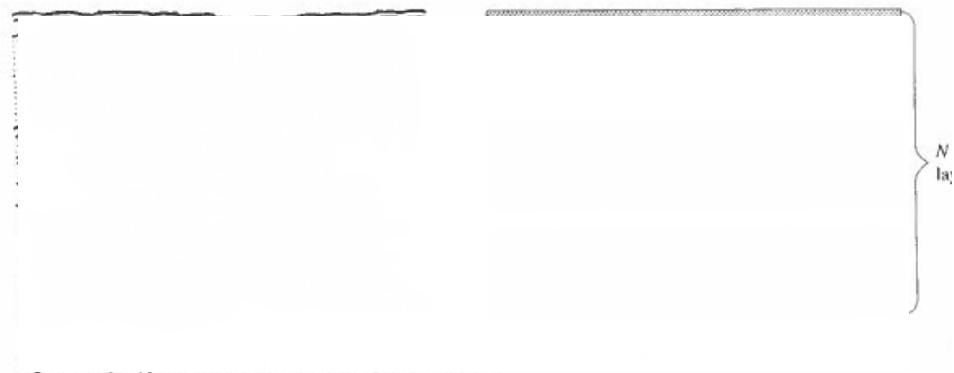
PROSPECT

PROSPECT er en model for bladets sammensætning af forskellige biokemiske komponenter, hvis påvirkning af reflektans og transmittans af elektromagnetisk stråling beskrives.

Et blad består af overfladevæv (epidermis), grundvæv (bladkød eller mesofyl) og ledningsvæv. I mesofyllet indeholder cellerne grønkorn (kloroplaster), hvor fotosyntesen finder sted. Ledningsvævet sørger dels for at transportere vand til bladet og for at transportere kulhydrater produceret i fotosyntesen ud til planten. Generelt er blade grønne, men afhængig af indholdet af carotenoider kan de også være røde eller gule. Indholdet af carotenoider gør det muligt for planten at udnytte en større del af det synlige lysspektrum i fotosyntesen [Den Store Danske, 2014a].

På figur 2.6 på side 47 til venstre ses en illustration af et tværsnit gennem et blad, hvor lagdelingen er tydelig. Yderst er epidermis. Inde i bladet er mesofyllet delt op i et lag med tætsiddende celler indeholdende kloroplaster og et lag med hulrum mellem cellerne.

PROSPECT er beskrevet i Jacquemoud og Baret [1990] og Baret og Fourty [1997]. Strålernes absorption og refleksion i et blad afhænger af bladets optiske egenskaber, der er en funktion af kemiske og fysiske karakteristika ved bladet. Baseret på information om pigmentindhold i bladene simulerer PROSPECT reflektans og transmittans for det elektromagnetiske spektrum mellem 400 og 2500 nm.

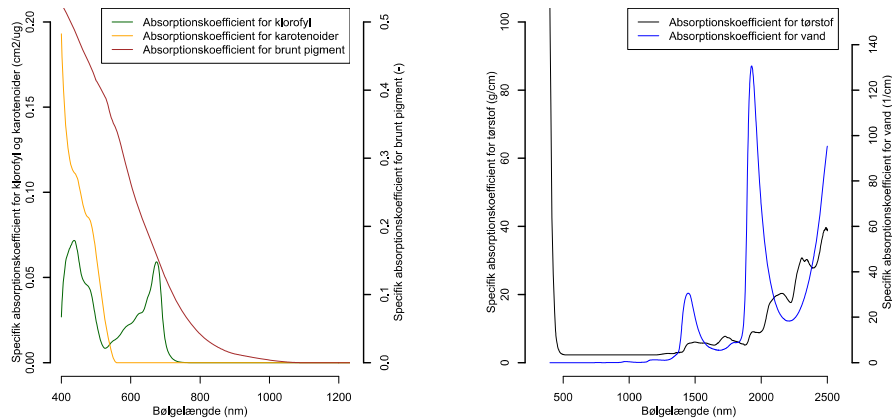


Figur 2.6: Figuren viser en illustration af et tværsnit gennem et typisk blad (til venstre) og PROSPECTs model for et blad ud fra et antal lag [Baret og Fourty, 1997].

PROSPECT er en generalisering af en model udviklet for et såkaldt kompakt blad til at gælde blade, der kan opdeles i flere homogene lag. På figur 2.6 på side 47 sammenlignes et tværsnit gennem et typisk blad med PROSPECTs lagdeling af bladet.

Modellen for et kompakt blad går ud på at beskrive bladets optiske egenskaber med to optiske konstanter, det effektive brydningsindeks n og den effektive absorptionskoefficient k . Brydningsindekset n er et tal, der beskriver, hvordan lys udbredes gennem et medium, og for et blad vil det effektive brydningsindeks afhænge af bølgelængden.

I PROSPECT antages det at, at bladet kan beskrives som en stak af N homogene lag, hvor der kan ske interaktion mellem fotoner og bestanddele i bladet. Den globale absorptionskoefficient k , der afhænger af bladets indhold af absorberende molekyler. Absorberende molekyler i et blad er klorofyl, vand, protein mm. På figur 2.7a på side 48 og figur 2.7a på side 48 ses de specifikke absorptionskoefficienter for klorofyl, carotenoider, brunt pigment, vand- og tørstofindhold, som anvendes i PROSPECT-5. Absorptionskoefficienterne er eksperimentelt bestemt i Feret et al. [2008]. På figur 2.1 på side 35 kan man se, hvilket pigmenter der er styrende for bladets reflektans i forskellige bølgelængder. Carotenoider er styrende for absorption af lys med bølgelængde under 500 nm, hvilket også illustreres af den specifikke absorptionskoefficient på figur 2.7a på side 48. Klorofyls absorption er stærkest ved ca. 450 nm og mellem 600–700 nm, hvilket stemmer godt overens med figur 2.1 på side 35. Brunt pigment påvirker reflektansen mest ved lave bølgelængder. Brunt pigment har indflydelse på reflektansen, når bladets alder [Houborg og Anderson, 2009]. Vandindholdets absorption af stråling har først indflydelse ved bølgelængder over 1400 nm ifølge figur 2.1 på side 35, hvilket også kan ses ud fra den specifikke absorptionskoefficient på figur 2.7a på side 48. Ab-



(a) Absorption ved klorofyl, carotenoider og brunt pigment

(b) Absorption ved vand og tørstof

Figur 2.7: Figuren til venstre viser den specifikke absorptionskoefficient for klorofyl (a+b), carotenoider og brunt pigment for bølgelængderne 400-1200 nm som anvendes i PROSPECT-5. Ved højere bølgelængder er der ingen absorption. Figuren til højre viser de specifikke absorptionskoefficienter for vand og tørstof for bølgelængderne 400-2500 nm som anvendes i PROSPECT-5. Indflydelsen er særlig stor i den melleminfrarøde del af spektret.

sorption pga. tørstof sker ved bølgelængder under ca. 450 nm og derudover især i den melleminfrarøde del af spektret. Ifølge Feret et al. [2008] skyldes absorptionspeaket i det blå bånd indholdet af brune pigmenter, hvis betydning skjules af kloroplaster ved højere bølgelængder.

Den globale absorptionskoefficient k er givet ved en sum af specifikke absorptionskoefficienter:

$$k(\lambda) = \sum_i K_i(\lambda)C_i$$

Her er C_i indholdet af bladkomponent i pr. enhed bladareal, mens $K_i(\lambda)$ er den specifikke absorptionskoefficient for den pågældende bestanddel. λ er bølgelængden. Absorptionskoefficienterne er eksperimentelt bestemt og valideret for forskellige typer af blade oprindeligt i Jacquemoud og Baret [1990] for det effektive brydningsindeks, vandindhold, klorofyl og mesofylstrukturparameteren og senere forbedret bl.a. i Feret et al. [2008], hvor absorptionskoefficienter for carotenoider og brunt pigment også bestemmes. Tørstofindholdet er også inkluderet i senere udgaver af PROSPECT [Feret et al., 2008].

I tabel 2.2 ses en oversigt over inputparametre til PROSPECT-5. I det følgende undersøges indholdet af klorofyl og carotenoider og værdien af mesofyl-

Symbol	Navn	Enhed
Chl_{a+b}	Klorofylindhold	$\mu g/cm^2$
C_{ar}	Indhold af carotenoider	$\mu g/cm^2$
C_{brown}	Indhold af brunt pigment	–
C_w	Vandindhold	cm
C_m	Tørstofindhold	g/cm^2
N	Mesofylstrukturparameter	–

Tabel 2.2: Oversigt over parametre i PROSPECT-5 [Feret et al., 2008].

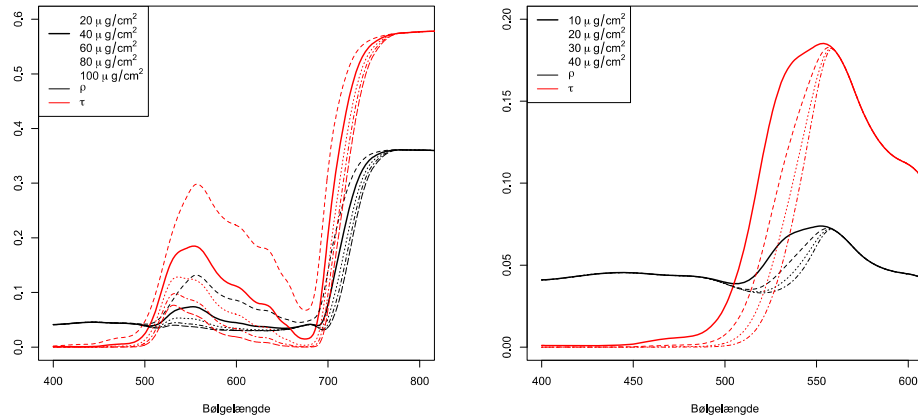
strukturparameterens betydning for PROSPECTs simuleringer af reflektans og transmittans.

Indhold af klorofyl og carotenoider For at illustrere betydningen af indholdet af bladklorofyl i PROSPECT er reflektans og transmittans ved varierende klorofylindhold simuleret ved fast indhold af carotenoider. Resultatet ses på figur 2.8a på side 50. Indholdet af carotenoider er ligeledes varieret ved fast indhold af klorofylindhold. Resultatet ses på figur 2.8b på side 50.

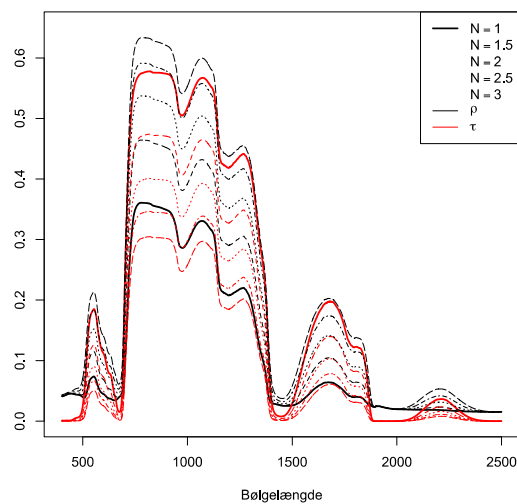
I PROSPECT har klorofylindholdet en betydning for reflektansen (ρ) og transmittansen (τ) ved bølgelængder mellem ca. 500–750 nm. Reflektansen er størst ved lavt klorofylindhold. For klorofylindhold over $60 \mu g/cm^2$ bliver ρ næsten konstant mellem 500–700 nm. Det tyder på, at reflektansen opnår et minimum, hvor yderligere tilføjelse af klorofyl ikke længere påvirker reflektansen. τ er ligeledes størst ved lavest klorofylindhold. Over 750 nm er summen af ρ og τ tæt på 1 svarende til, at der ikke længere sker absorption af stråling i bladet. Indholdet af carotenoider påvirker reflektansen (ρ) og transmittansen (τ) mellem 400–600 nm i PROSPECT. På figur 2.8b på side 50 ses det, at ρ falder ved stigende indhold af carotenoider mellem 500–550 nm. Samtidig falder τ ved stigende indhold af carotenoider mellem 450–550 nm.

Mesofylstrukturparameteren Parameteren N kaldes for mesofylstrukturparameteren og beskriver antallet af lag i bladet. Ifølge Jacquemoud og Baret [1990] er mesofylstrukturparameteren N ikke direkte målbar. Selv om N beskrives som antallet af lag inde i bladet, kan den antage ikke-heltalsværdier. N er relateret til cellernes placering i bladet. N mellem 1–1.5 svarer til en kompakt mesophyll-struktur, der ses hos blade, der er monocotyledone. N mellem 1.5–2.5 ses hos blade, der er dicotyledone. ”Mono-” og ”di-” hentyder til, at der for disse planter er enten ét eller to kim i et frø. Større værdier af N over 2.5 ses hos aldrende blade, hvis indre struktur går i opløsning.

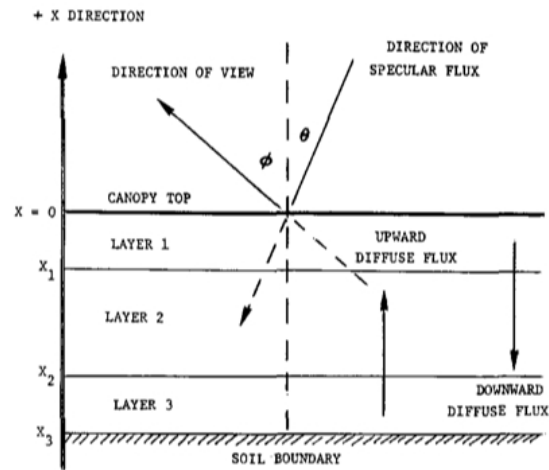
På figur 2.9 på side 50 ses den simulerede reflektans (ρ) og transmittans (τ) i PROSPECT ved forskellige værdier af N i hele spektret fra 400–2500 nm. Jo større værdi af N , desto større bliver ρ , og desto mindre bliver τ . Ved værdier af N over 1.5 bliver ρ større end τ .



Figur 2.8: Figuren til venstre viser simuleret reflektans (ρ) og transmittans (τ) ved varieret klorofylindhold med PROSPECT-5. Indholdet af carotenoider er $10 \mu\text{g}/\text{cm}^2$, vandindholdet 0.01 cm , tørstofindholdet $0.0055 \text{ g}/\text{cm}^2$, N er 1, og bladet indeholder ikke brunt pigment. Figuren til højre viser simuleret reflektans (ρ) og transmittans (τ) ved varieret indhold af carotenoider med PROSPECT-5. Klorofylindholdet er $40 \mu\text{g}/\text{cm}^2$, vandindholdet 0.01 cm , tørstofindholdet $0.0055 \text{ g}/\text{cm}^2$, N er 1, og bladet indeholder ikke brunt pigment.



Figur 2.9: Figuren viser simuleret reflektans (ρ) og transmittans (τ) ved varierende værdi af mesofylstrukturparameteren N med PROSPECT-5. Klorofylindholdet er $40 \mu\text{g}/\text{cm}^2$, indholdet af carotenoider er $10 \mu\text{g}/\text{cm}^2$, vandindholdet er 0.01 cm , tørstofindholdet $0.0055 \text{ g}/\text{cm}^2$, og bladet indeholder ikke brunt pigment.



Figur 2.10: Illustration af Suits' model Suits [1972].

SAIL

SAIL står for *Scattering of Arbitrarily Inclined Leaves* og er udviklet af Verhoef [1984]. SAIL er en udvidelse af Suits' model [Suits, 1972], der er illustreret på figur 2.10 på side 51. Suits' model er baseret på AGR-modellen [Allen et al., 1970], som er baseret på Kubelka-Munk-teorien fra 1931. Kubelka-Munk-teorien handler om spredning og udslukning af lys i ideelle diffuserende medier og beskriver to diffuse fluxe: En opadrettet E_+ og en nedadrettet E_- . I Allen et al. [1970] overføres Kubelka-Munk-teorien til at beskrive spredning og udslukning af stråling i en plantestand, og der tilføjes en direkte flux E_s for at beskrive solindstrålingens påvirkning. Suits [1972] tilføjer L_o , som er radiansen i sensorens retning. Ved at anvende $E_o = \pi L_o$ kan E_o fortolkes som irradiansen fra en lambertsk flade med radians L_o . Det svarer til ligning 2.2.1. I dette tilfælde er E_o irradians og ikke emittans. Det skyldes, at

I SAIL er plantestanden beskrevet som bestående af et antal lag. Antagelserne om ét sådant lag i plantestanden i den originale udgave af SAIL er følgende:

- Et lag er horisontalt og uendelig udvidet.
- Plantestandens eneste komponenter er små, flade blade.
- Laget er homogent.

Det antages også, at bladenes orienteringer kan beskrives ved en bladvinkelfordelingsfunktion, der er beskrevet i afsnit 2.2.3.

SAIL er baseret på fire differentiaalligninger og tager højde for fordelingen af bladvinkler. Differentiaalligningerne beskriver spredning og udslukning af stråling i et enkelt lag i en plantestand.

$$\begin{aligned}
\frac{dE_s}{dx} &= kE_s \\
\frac{dE_-}{dx} &= -sE_s + aE_- - \sigma E_+ \\
\frac{dE_+}{dx} &= s'E_s + \sigma E_- - aE_+ \\
\frac{dE_o}{dx} &= wE_s + vE_- + uE_+ - KE_o
\end{aligned}
\tag{2.14}$$

Koefficienterne i ligningssystemet 2.14 er dimensionsløse. E_s, E_-, E_+ og E_o er fluxe og x en relativ optisk højde af eller tykkelse for laget. x er derfor også dimensionsløs. Koefficienterne kan fortolkes på følgende vis:

k	Udslukning af direkte sollys
s	Spredning af direkte sollys i nedadgående retning
s'	Spredning af direkte sollys i opadgående retning
a	Spredning af diffus nedadrettet stråling i nedadgående retning
σ	Spredning af diffus nedadrettet stråling i opadgående retning
w	Spredning af direkte sollys i sensorens retning (bidirektional spredning)
v	Spredning af diffus opadrettet stråling i sensorens retning
u	Spredning af diffus nedadrettet stråling i sensorens retning
K	Udslukning af stråling i sensorens retning

De fire fluxe er illustreret på figur 2.10 på side 51. Toppen af laget er $x = 0$, og bunden af laget svarer til $x = -1$. Dvs., at x -retningen er negativ for E_s og E_- og positiv for E_+ og E_o . Det svarer til, at sollyset udslukkes med udslukningskoefficienten k . E_- får et bidrag fra sollyset med spredningskoefficienten s , et bidrag fra E_+ med spredningskoefficienten σ og et tab i form af diffus stråling, der spredes i opadgående retning med spredningskoefficienten a . Omvendt får E_+ et bidrag fra solen med spredningskoefficienten s' , et bidrag fra E_- med spredningskoefficienten a og et tab til E_- med spredningskoefficienten σ . For E_o kommer der bidrag fra tre kilder: En del af sollyset spredes i sensorens retning, når det rammer toppen af plantestanden. Dette bidrag beskrives ved koefficienten w . Den anden del kommer fra E_- og E_+ beskrevet ved v og u . Der sker også en udslukning af strålingen angivet ved udslukningskoefficienten K .

Når det direkte sollys spredes, bliver det til diffus stråling. Diffus stråling kan ikke blive til direkte stråling.

Ligningssystemet kan også skrives på matrixform:

$$\frac{d}{dx} \mathbf{E} = M \mathbf{E}$$

Her er M matricen med koefficienterne:

$$M = \begin{pmatrix} k & 0 & 0 & 0 \\ -s & a & -\sigma & 0 \\ s' & \sigma & -a & 0 \\ w & v & u & -K \end{pmatrix}$$

Verhoef [1985] løser ligningssystemet for et enkelt lag ved anvende en lineær transformation Y på M for at opnå en diagonal matrix:

$$\Lambda = YMY^{-1}$$

Derefter løses følgende ligningssystem:

$$\frac{d}{dx}\mathbf{F} = \Lambda\mathbf{F}$$

Løsningen udtrykt ved \mathbf{F} kan transformeres vha. Y^{-1} . I appendiks A løses differentiaalligningssystemet. I forbindelse med løsningen af ligningssystemet indføres følgende nye koefficienter [Verhoef, 1985]:

$$\begin{aligned} h_1 &= \frac{a+m}{\sigma} & h_2 &= \frac{a-m}{\sigma} \\ C_s &= \frac{s'(k-a)-s\sigma}{k^2-m^2} & C_o &= \frac{v(K-a)-u\sigma}{K^2-m^2} \\ D_s &= \frac{-s(k+a)-s'\sigma}{k^2-m^2} & D_o &= \frac{-u(K+a)-v\sigma}{K^2-m^2} \\ H_s &= \frac{uC_s+vD_s+w}{K+k} & H_o &= \frac{sC_o+s'D_o+w}{K+k} \end{aligned} \quad (2.15)$$

Den lineære transformation kaldes Y^{-1} og består af egenvektorerne i rækkerne:

$$Y^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ D_s & h_1 & h_2 & 0 \\ C_s & 1 & 1 & 0 \\ H_s & h_1C_o - D_o & h_2C_o - D_o & 1 \end{pmatrix}$$

Løsningen udtrykt vha. \mathbf{F} er givet ved følgende:

$$\begin{aligned} F_1 &= \delta_1 \exp(kx) \\ F_2 &= \delta_2 \exp(mx) \\ F_3 &= \delta_3 \exp(-mx) \\ F_4 &= \delta_4 \exp(-Kx) \end{aligned}$$

Hvor $\delta_1, \dots, \delta_4$ er konstanter, der afhænger af randbetingelserne. Løsningen \mathbf{E} findes ved at anvende Y^{-1} på \mathbf{F} :

$$\begin{aligned}
E_s &= \delta_1 \exp(kx) \\
E_- &= D_s \delta_1 \exp(kx) + h_1 \delta_2 \exp(mx) + h_2 \delta_3 \exp(-mx) \\
E_+ &= C_s \delta_1 \exp(kx) + \delta_2 \exp(mx) + \delta_3 \exp(-mx) \\
E_o &= H_s \delta_1 \exp(kx) + (h_1 C_o - D_o) \delta_2 \exp(mx) \\
&\quad + (h_2 C_o - D_o) \delta_3 \exp(-mx) + \delta_4 \exp(-Kx)
\end{aligned} \tag{2.16}$$

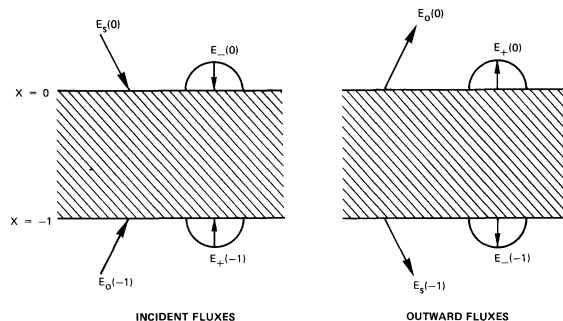
Løsningen for et enkelt lag kan udvides til flere lag. På figur 2.11 på side 55 ses en illustration af de indkommende og udgående fluxe for et lag. De indkommende fluxe er $E_s(0), E_-(0), E_+(-1)$ og $E_o(-1)$. De udgående er $E_s(-1), E_-(-1), E_+(0)$ og $E_o(0)$. Det antages, at de indkommende fluxe er kendte og derudfra bestemmes de udgående fluxe ud fra ligningerne 2.16. Der indføres igen nye koefficienter, der udtrykker reflektans ρ og transmittans τ , udtrykt vha. koefficienterne 2.15:

$$\begin{aligned}
\tau_{ss} &= \exp(-k) \\
\tau_{oo} &= \exp(-K) \\
\rho_{dd} &= \frac{\exp(m) - \exp(-m)}{h_1 \exp(m) - h_2 \exp(-m)} \\
\tau_{dd} &= \frac{h_1 - h_2}{h_1 \exp(m) - h_2 \exp(-m)} \\
\rho_{sd} &= C_s(1 - \tau_{ss}\tau_{dd}) - D_s\rho_{dd} \\
\tau_{sd} &= D_s(\tau_{ss} - \tau_{dd}) - C_s\tau_{ss}\rho_{dd} \\
\rho_{do} &= C_o(1 - \tau_{ss}\tau_{dd}) - D_o\rho_{dd} \\
\tau_{do} &= D_o(\tau_{ss} - \tau_{dd}) - C_o\tau_{oo}\rho_{dd} \\
\rho_{so} &= H_o(1 - \tau_{ss}\tau_{dd}) - C_o\tau_{sd}\tau_{oo} - D_o\rho_{sd}
\end{aligned}$$

Resultatet kan skrives på matrixform $\mathbf{E}_{\text{ud}} = \mathbf{Z}\mathbf{E}_{\text{ind}}$:

$$\begin{pmatrix} E_s(-1) \\ E_-(-1) \\ E_+(0) \\ E_o(0) \end{pmatrix} = \begin{pmatrix} \tau_{ss} & 0 & 0 & 0 \\ \tau_{sd} & \tau_{dd} & \rho_{dd} & 0 \\ \rho_{sd} & \rho_{dd} & \tau_{dd} & 0 \\ \rho_{so} & \rho_{do} & \tau_{do} & \tau_{oo} \end{pmatrix} \begin{pmatrix} E_s(0) \\ E_-(0) \\ E_+(-1) \\ E_o(-1) \end{pmatrix}$$

Resultatet kan anvendes til at beregne reflektansen fra et lag placeret over en ikke-lambertsk overflade. Dvs., at man kan tage højde for, at jordbunden ikke er lambertsk. Hvis en flade ikke er lambertsk, betyder det, at der er forskel på hvor meget stråling, der reflekteres i forskellige retninger. Derfor skal man kende til fordelingen af reflektans fra en sådan flade. Man skelner mellem fire typer reflektans: Hemisfærisk-hemisfærisk reflektans, hemisfærisk-direkte reflektans, direkte-hemisfærisk reflektans og bidirektional reflektans. "Hemisfærisk" hentyder til, at der er tale om diffus stråling, der kommer fra alle retninger. Hemisfærisk-hemisfærisk reflektans er altså den reflektans, der opstår, når diffus stråling rammer en flade og reflekteres i alle



Figur 2.11: Figuren viser de indkommende fluxe (til venstre) og de udgående fluxe (til højre) fra et enkelt lag i en plantestand. De direkte pile henviser til, at fluxen er direkte, mens halvcirklerne illustrerer, at fluxen er diffus [Verhoef, 1985].

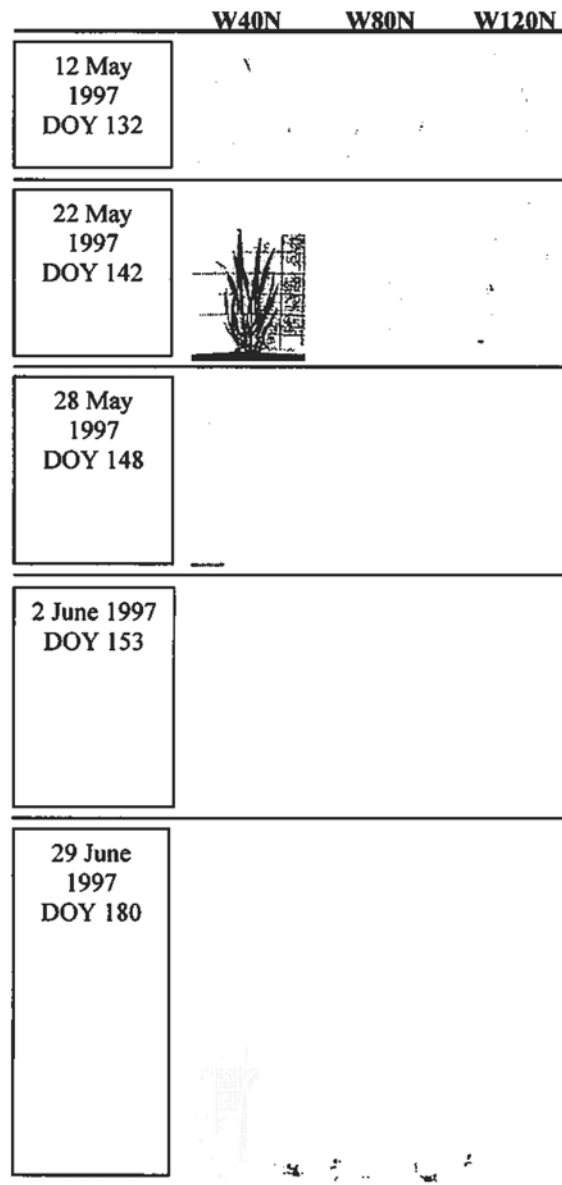
retninger. Hemisfærisk-direkte reflektans er reflektansen af diffus stråling i en bestemt retning. Direkte-hemisfærisk reflektans er reflektansen af direkte stråling i alle retninger. Birektional reflektans er reflektansen fra en direkte stråle i en bestemt retning. Overfladereflektansen fra en ikke-lambertsk flade karakteriseres altså af disse fire typer reflektans. Ud fra en overfladereflektansmatrix beregnes de fire typer af reflektans for toppen af et lag, der ligger over den ikke-lambertske flade. Modellen kan udvides til at omfatte flere lag i en plantestand ved at beregne overfladereflektansen for hvert lag rekursivt. Dvs., at man starter med overfladereflektansen for jordbunden, beregner overfladereflektansen for første lag, derefter for andet lag osv. [Verhoef, 1985].

Resultatet af SAIL er fire reflektansfaktorer, én for hver type reflektans. En reflektansfaktor er forholdet mellem reflektans fra den relevante flade og den reflektans, der under samme indstrålingsforhold, ville være kommet fra en lambertsk flade.

Bladvinkelfordeling

Bladvinkelfordelingen er en væsentlig faktor, når plantestandens reflektans skal modelleres. Det skyldes, at bladvinklerne afgør, hvor synlig jordbunden er, og hvor stor en del af bladarealet, der modtager direkte sollys. På figur 2.12 på side 56 ses billeder af hvedeplanter fra marker med varieret kvælstoftilgængelighed. Ifølge Broge [2003] er bladene generelt mere opretstående for hvedeplanter, der har mangel på kvælstof.

SAIL tager højde for fordelingen af bladenes vinkler. Bladenes orientering er relevante at tage med i betragtning, når man skal lave en model for reflektans fra vegetation. Hvis man fx betragter en plantestand fra zenit, vil horisontalt vinklede blade kunne ses i deres fulde udstrækning. Vertikalt orienterede blade vil til gengæld kun kunne ses fra siden. Dermed har bladvinkel stor indflydelse på det tilsyneladende areal af bladene, og derfor også på reflek-



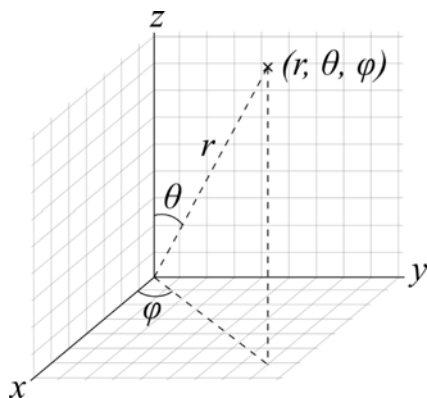
Figur 2.12: Figuren viser billeder af hvedeplanter fra tre hvedemarker ved forskellige tidspunkter i vækstsæsonen. Hvedemarkerne har forskelligt niveau af kvælstoftilgængelighed. Mod venstre er marken udsat for kvælstofmangel mens tilgængeligheden stiger mod højre. Planterne med kvælstofmangel viser en mere erektofil bladvinkelfordeling [Broge, 2003].

sion og absorption af stråling fra en given retning. Bladvinkelfordelingerne beskrives ud fra sfæriske koordinater.

I et tredimensionalt koordinatsystem kan et punkts koordinater angives med sfæriske koordinater. Punktets koordinater er givet ved afstanden til origo, r , samt to vinkler, θ og ϕ . På figur 2.13 er de sfæriske koordinater illustreret. Vinklen mellem z -aksen eller zenit kaldes θ . Den anden vinkel ϕ er azimutvinklen. Den angiver vinklen fra x -aksen til den ortogonale projektion af linjestykket mellem origo til punktet.

Man kan finde koordinaterne på formen (x, y, z) ud fra de sfæriske koordinater:

$$\begin{aligned}x &= (r \sin \theta \cos \phi) \\y &= (r \sin \theta \sin \phi) \\z &= (r \cos \theta)\end{aligned}$$



Figur 2.13: Figuren viser vinklen θ i forhold til zenit og azimutvinklen ϕ . [Wikipedia, 2014]

Bladets zenitvinkel betegnes θ_l og azimutvinklen ϕ_l . Det antages, at der er såkaldt azimutal isotropi, dvs. at azimutvinklerne er stokastisk fordelt. Ifølge Verhoef [1984] er det en rimelig antagelse eftersom kun få planter er heliotrope, dvs. vender deres blade mod solen.

Bladvinkler kan beskrives med en sandsynlighedstæthedsfunktion, den såkaldte *leaf angle distribution function* eller *leaf inclination distribution function*. Fordelingen af bladvinkler forkortes ofte LAD eller LIDF. En sandsynlighedstæthedsfunktion kendes fra definition 2.3.

I det følgende angives tæthedsfunktioner for en række idealiseringer af fordelinger af bladvinkler. For alle tæthedsfunktionerne gælder følgende:

$$\int_0^{\frac{\pi}{2}} f(\theta_l) d\theta_l = 1$$

De følgende bladvinkelfordelinger er idealiseringer og oprindeligt foreslået af de Wit [1965].

Horisontale blade: Den planofile fordeling I den planofile fordeling af bladvinkler er de fleste blade horisontalt orienterede. Den beskrives ifølge Verhoef [1997] ved følgende tæthedsfunktion:

$$f(\theta_l) = \frac{2}{\pi}(1 - \cos 2\theta_l)$$

Vertikale blade: Den erektofile fordeling I den erektofile fordeling af bladvinkler er de fleste blade vertikale orienterede. Den beskrives ifølge Verhoef [1997] ved følgende tæthedsfunktion:

$$f(\theta_l) = \frac{2}{\pi}(1 + \cos 2\theta_l)$$

Den plagiofile fordeling I den plagiofile fordeling af bladvinkler er de fleste blade orienteret med en vinkel på 45 °. Den beskrives ifølge Verhoef [1997] ved følgende tæthedsfunktion:

$$f(\theta_l) = \frac{2}{\pi}(1 - \cos 4\theta_l)$$

Den ekstremofile fordeling I den ekstremofile fordeling af bladvinkler er de fleste af bladene enten horisontalt eller vertikalt orienterede og kun få blade orienteret med en vinkel på 45 °. Den beskrives ifølge Verhoef [1997] ved følgende tæthedsfunktion:

$$f(\theta_l) = \frac{2}{\pi}(1 + \cos 4\theta_l)$$

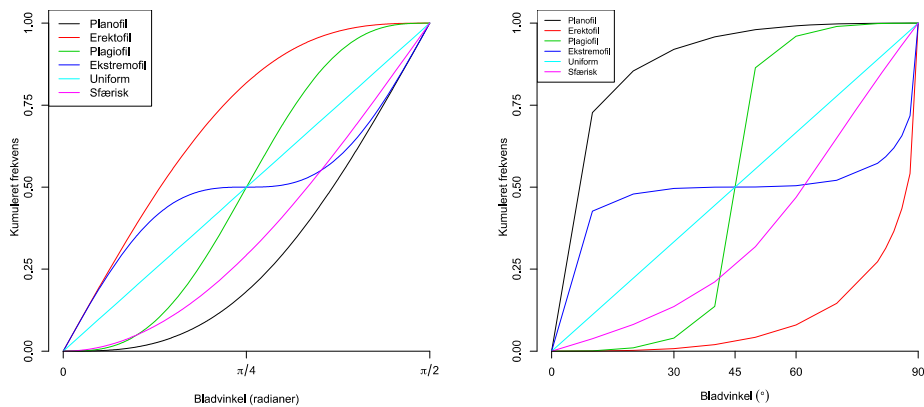
Den uniforme fordeling Den uniforme fordeling er en ligefordeling af bladenes vinkler, dvs. tæthedsfunktionen er konstant:

$$f(\theta_l) = \frac{2}{\pi}$$

Den sfæriske fordeling I den sfæriske fordeling er bladenes vinkler fordelt som på en kugleflade. Den beskrives ifølge Verhoef [1997] ved følgende tæthedsfunktion:

$$f(\theta_l) = \sin \theta_l$$

Fordelingsfunktionerne for de seks bladvinkelfordelinger er illustreret på figur 2.14a på side 59.



(a) De Wits bladvinkelfordelinger (b) Bladvinkelfordeling med Verhoefs algoritme

Figur 2.14: Figuren til venstre viser fordelingsfunktioner for De Wits inddeling af bladvinkelfordelinger, og figuren til højre viser de genererede fordelingsfunktioner for bladvinkler ud fra Verhoefs algoritme for bladvinklerne $0^\circ, 10^\circ, \dots, 80^\circ, 82^\circ, 84^\circ, \dots, 90^\circ$.

Som nævnt er bladvinkelfordelingerne idealiseringer. de Wit [1965] måler bladvinkler på forskellige afgrøder og finder, at raps og sukkerroer er har en tilnærmelsesvis plagiofil fordeling, kartofler tilnærmelsesvis planofil, mens majs og afgrøder som hvede, byg og havre har en mere eller mindre sfærisk fordeling. Græsmarker findes at have en erektofil fordeling i begyndelsen af sin vækst og en sfærisk fordeling senere i forløbet. Bladvinkelfordelingerne foreslået af de Wit [1965] anvendes dog ikke i SAIL. I stedet implementeres bladvinkelfordelinger baseret på de Wit [1965] med følgende algoritme.

Verhoefs algoritme Følgende metode kaldes Verhoefs algoritme i Wang et al. [2007] og er beskrevet i Verhoef [1997].

Fordelingsfunktionerne for bladvinklerne beskrives af en linearkombination af trigonometriske funktioner:

$$y = a \sin x + \frac{b \sin 2x}{2}$$

Der indføres restriktionen $|a| + |b| < 1$ for at sikre at fordelingsfunktionen er monotont voksende. x og y er relateret til θ_l og $F(\theta_l)$:

$$x = \frac{\pi}{2} F(\theta_l) + \theta_l$$

$$y = \frac{\pi}{2} F(\theta_l) - \theta_l$$

Algoritmen beregner $F(\theta_l)$, hvis θ_l er givet, og kan formuleres i pseudokode:

```

 $x = 2\theta_l$ 
repeat
   $y = a \sin x + \frac{b \sin 2x}{2}$ 
   $\Delta x = \frac{1}{2}(y - x + 2\theta_l)$ 
   $x = x + \Delta x$ 
until  $|\Delta x| < t$ 
 $F(\theta_l) = \frac{2}{\pi}(y + \theta_l)$ 

```

t er en tærskelværdi, der skal være lille fx 10^{-6} .

På figur 2.14b ses de genererede fordelingsfunktioner vha. Verhoefs algoritme.

Når man sammenligner figur 2.14a på side 59 og figur 2.14b på side 59 ses det, at den sfæriske fordeling er stort set sammenfaldende, mens de planofile, erektofile, plagiofile og ekstremofile fordelinger antager mere ekstreme værdier, når de genereres med Verhoefs algoritme. Det er netop formålet med Verhoefs algoritme.

Den ellipsoidale fordeling Den ellipsoidale fordeling er en generalisering af den sfæriske fordeling og antager, at bladenes vinkler er fordelt som på overfladen af en ellipsoide. Kuglen er også en ellipsoide, hvorfor den sfæriske fordeling er et særtilfælde af den ellipsoidale. Den er udviklet af Campbell [1986, 1990]. Den beskrives ved følgende tæthedsfunktion:

$$f(\theta_l) = \frac{2\chi^3 \sin \theta_l}{\Lambda(\cos^2 \theta_l + \chi^2 \sin^2 \theta_l)}$$

χ er forholdet mellem den vertikale halvakse a og den horisontale halvakse b , dvs. $\chi = \frac{b}{a}$. Λ afhænger af χ og er et normaliseret ellipsoareal. Hvis $\chi = 1$ er ellipsoiden en kugle, og tæthedsfunktionen reduceres til $\sin \theta_l$, fordi $\Lambda = 2$. Hvis $\chi < 1$, så er den horisontale halvakse kortere end den vertikale, og ellipsoiden kaldes prolat. Så er Λ givet ved følgende:

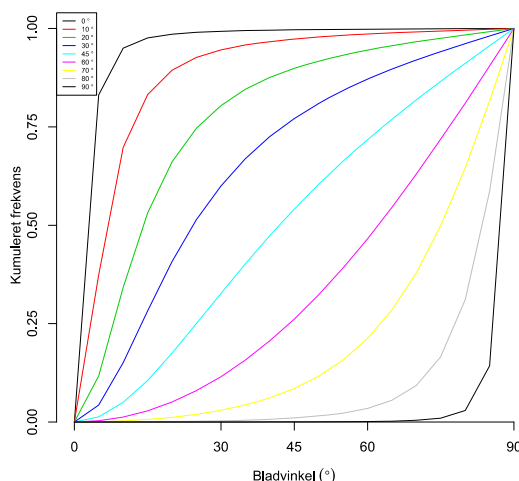
$$\Lambda = \chi + \frac{\sin^{-1} \epsilon}{\epsilon}, \text{ hvor } \epsilon = \sqrt{1 - \chi^2}$$

Hvis $\chi > 1$, så er den horisontale halvakse længere end den vertikale, og ellipsoiden kaldes oblat. Så er Λ :

$$\Lambda = \chi + \frac{\ln\left(\frac{1+\epsilon}{1-\epsilon}\right)}{2\epsilon\chi}, \text{ hvor } \epsilon = \sqrt{1 - \chi^{-2}}$$

ϵ i begge udtryk svarer til excentriciteten af ellipsoiden. Hvis $\chi = 1$, bliver excentriciteten 0.

I PROSAIL er den ellipsoidale bladvinkelfordeling implementeret således, at værdien e afhænger af middelbladvinklen ϕ_l med følgende funktion:



Figur 2.15: Figuren viser Campbells ellipsoidale bladvinkelfordelingsfunktioner for forskellige middelbladvinkler.

$$e(\theta_l) = \exp(-1.6184 \cdot 10^{-5} \cdot \phi_l^3 + 2.1145 \cdot 10^{-3} \cdot \phi_l^2 - 1.2390 \cdot 10^{-1} \cdot \phi_l + 3.2491)$$

Hvis e er lig med 1 beregnes bladvinkelfordelingen som den sfæriske fordeling. Det sker, når $\phi_l \approx 58.4^\circ$. Dvs., at e er et estimat af χ .

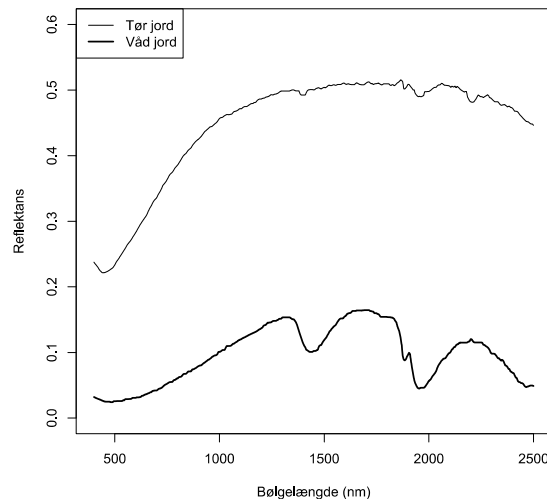
I SAIL modelleres bladvinkelfordelingen af to forskellige algoritmer. Den ene er baseret på de analytiske funktioner, der beskriver forskellige typer bladvinkelfordeling jf. figur 2.14a på side 59. Den implementeres i PROSAIL af Verhoefs algoritme jf. figur 2.14b på side 59. Når man sammenligner figur 2.14a på side 59 og figur 2.14b på side 59 ses det, at den sfæriske fordeling er stort set sammenfaldende, mens de planofile, erektofile, plagiofile og ekstremofile fordelinger antager mere ekstreme værdier, når de genereres med Verhoefs algoritme.

På figur 2.15 på side 61 ses den ellipsoidale fordeling af bladvinkler. Fordelelen ved denne fordeling er, at den kun kræver en parameter i stedet for to som i Verhoefs algoritme. Derfor er den generelt lettere at inverttere.

Bladvinkelfordelingen indgår i beregningen af koefficienterne fra lignings-systemet 2.14.

Jordbundens reflektans

Jordbundens reflektans i PROSAIL indgår som en vektor bestående af den spektrale reflektans fra 400–2500 nm med skridt på 1 nm. Jordbundsreflektansen fra Jacquemond et al. [2013] ses på figur 2.16 på side 62 for tør og våd



Figur 2.16: Figuren viser kurver for jordbundsreflektans fra datafilen, der følger med standardindstillingen af PROSAIL fra Jacquemond et al. [2013].

jord. Hvis LAI sættes til 0, beregnes reflektansen ud fra jordbundens reflektans. Hvis LAI er større end 0 indgår jordbundens reflektans i beregningen af reflektansfaktorerne.

Hotspotparameteren

Hotspottet opstår, når der er sammenfald mellem solstrålingens indfaldsvinkel og vinklen for satellittens sensor, dvs. når $\theta_s = \theta_o$. En plantestand er et tredimensionalt objekt, og inde i plantestanden forekommer der skygger. Når lyset rammer en plantestand påvirkes refleksionen af skyggerne. I hotspottet forsvinder effekten af skygger. Det forårsager, at plantestanden virker lysere i netop den retning i forhold til andre retninger. Hotspoteffekten har dog også indflydelse på refleksion af lys i andre retninger [Verhoef, 1997].

I SAIL indgår højde ikke som parameter, da der som nævnt anvendes et relativt højdemål. Hotspottet er dog relateret til plantestandens højde h gennem forholdet $\frac{\uparrow}{h}$. Forholdet kaldes hotspotparameteren. \uparrow er en horisontal korrektionslængde, der afhænger af bladets størrelse, form og orientering. Da parameteren afhænger af forholdet mellem karakteristika for blade og plantestandens højde, kan det være nærliggende at undersøge, om dette forhold er konstant over en vækstsæson eller om det forandrer sig. Planter vækst kan karakteriseres af udviklingen i LAI over en vækstsæson. Verhoef [1997] beskriver to forskellige strategier for at øge LAI i en plantestand: Enten kan planten øge sit antal af blade eller også kan planten lade sine eksisterende

Symbol	Navn	Enhed
LAI	Bladarealindeks	m^2/m^2
h_{spot}	Hotspot	–
θ_s	Solens zenitvinkel	°
θ_o	Observatørens zenitvinkel	°
ϕ	Azimutvinkel	°
LIDF	Fordelingsfunktion for bladvinkler	

Tabel 2.3: Parametre i den anvendte udgave af SAIL [Jacquemond et al., 2013].

blade vokse. Den første strategi anvendes af afgrøder som hvede, majs og kartofler, mens sukkerroer anvender den anden strategi. I praksis vil planter ofte anvende en kombination af de to strategier. Hvis planten anvender første strategi, vil højden af plantestanden øges over en vækstsæson, når antallet af blade pr. grundareal stiger. Dvs., at højden vokser proportionalt med LAI. Hvis man antager, at bladstørrelsen er konstant, så vil hotspotparameteren være proportional med det reciprokke LAI:

$$\frac{\downarrow}{h} \propto \frac{1}{LAI}$$

Verhoef og Bach [2003] estimerer proportionalitetskonstanten til 0.5 i følgende udtryk, hvor hotspotparameteren betegnes q :

$$q = \frac{0.5}{LAI}$$

Forholdet anvendes kun for $LAI > 1$ – ellers sættes q til 0.5. Ifølge Verhoef og Bach [2003] er forholdet en god beskrivelse for planter, der primært vokser i højden ved at gro nye blade frem for at bladene i sig selv vokser. Houborg et al. [2007] anvender dette forhold til at beskrive hvede, byg og løvfældende træer.

Oversigt over SAILs parametre

I tabel 2.3 ses en oversigt over inputparametrene i den anvendte version 4SAIL.

3. Data

I dette kapitel gennemgås data udvalgt fra datasættet i Boegh et al. [2013], som kan findes i den fil, der følger med den elektroniske version af denne artikel. Boegh et al. [2013] undersøger fem landskaber fordelt i Europa, heraf indgår det danske landskab i to omgange. Feltmålingerne er udført i perioden 31/5 til 21/7 2008, og det danske område indgår endnu en gang 19/4 2009. For det danske landskab er supplerende information tilgængelig vedrørende jordbundens sammensætning, reflektans og elektriske ledningsevne. Derfor er observationerne fra Danmark (2008) beskrevet grundigere og spiller en central rolle i specialet.

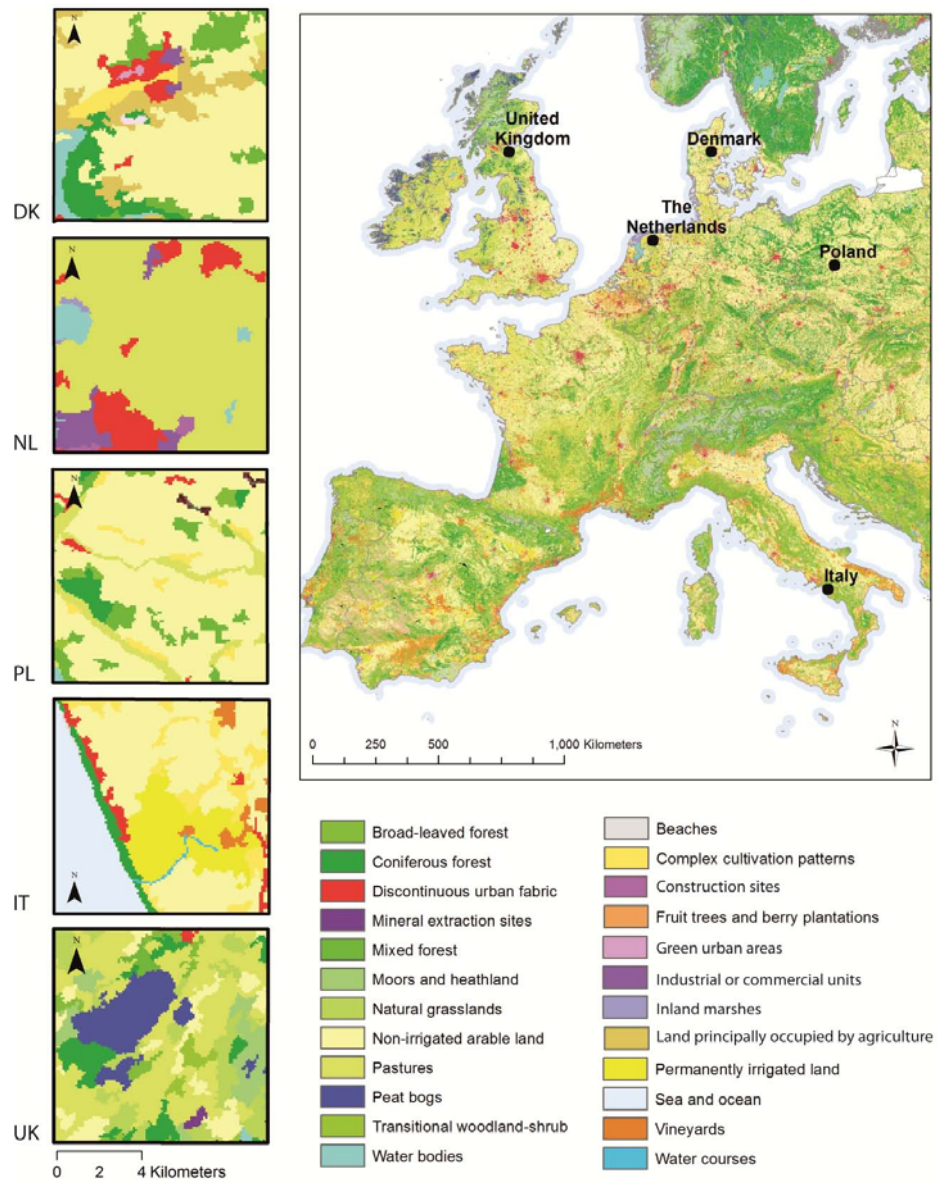
3.1 Landskaber

Der anvendes data fra fem forskellige lokaliteter i Europa. For den ene lokalitet i Danmark er der indsamlet data på to forskellige tidspunkter, nemlig midt i vækstsæsonen i 2008 og i den tidlige vækstsæson i 2009. Afgrøderne i det danske landskab er hvede og byg med enkelte raps-, græs- og majsmarker. I observationerne fra 2008 indgår både vinter- og vårafgrøder. I observationerne fra 2009 er der ingen bygmarker, men udelukkende hvedemarker i tidlig vækst.

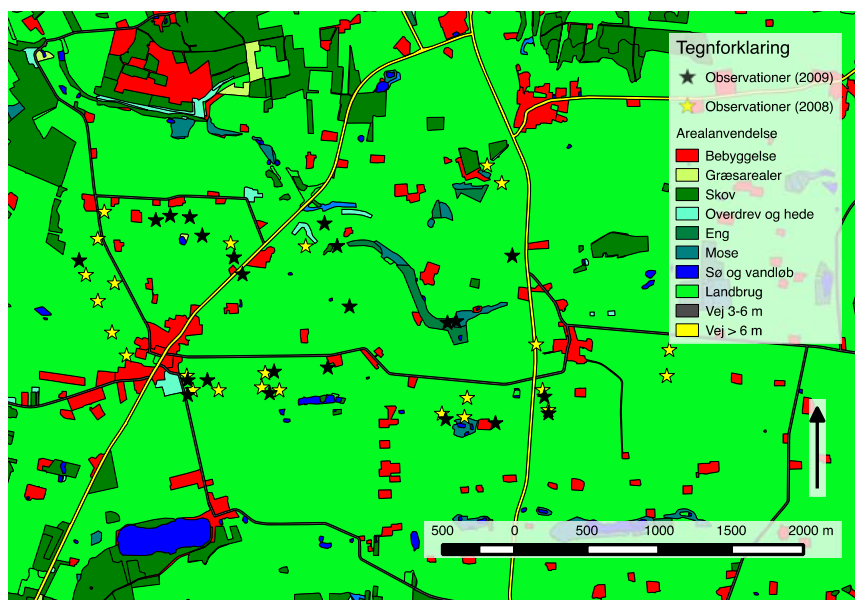
For de øvrige fire landskaber i Holland, Polen, Skotland og Italien er data indsamlet i 2008. For Skotland er der anvendt andre opmålingsmetoder end for de øvrige lokaliteter, og der er ikke foretaget målinger, der kan anvendes til at bestemme det vertikale klorofylprofil.

Det hollandske landskaber består i høj grad af græsmarker. I Polen er der en stor variation i afgrøder med marker med hvede, kartofler, raps, majs og lucerne. Det italienske landskab består af marker med majs, tomat og enkelte andre afgrøder som artiskok. Det skotske landskab består af naturlige græsser.

På figur 3.1 kan placeringen af landskaberne ses i forhold til arealdække. Det danske landskab består af landbrugsjord, skov, sø og bebyggelse. I Holland kan man se, at der er græsarealer samt bebyggelse og industri. Det polske område består af landbrugsjord, ikke-kunstvandet jord og skov. Det italienske landskab ligger ved kysten og består af både kunstvandet og ikke-kunstvandet jord samt vingårde. I Skotland er der naturlige græsser, hede, mose og skov. For det skotske område er klorofylindholdet ikke målt, hvorfor data fra Skotland ikke indgår i undersøgelsen.



Figur 3.1: Arealdækkkort over landskaber fra Boegh et al. [2013].



Figur 3.2: Arealanvendelseskort over området, hvor observationerne i Danmark i 2008 og 2009 og er foretaget. Kortet er lavet i QGIS med data fra Danmarks Miljøundersøgelser [2000].

3.1.1 Områdebeskrivelse

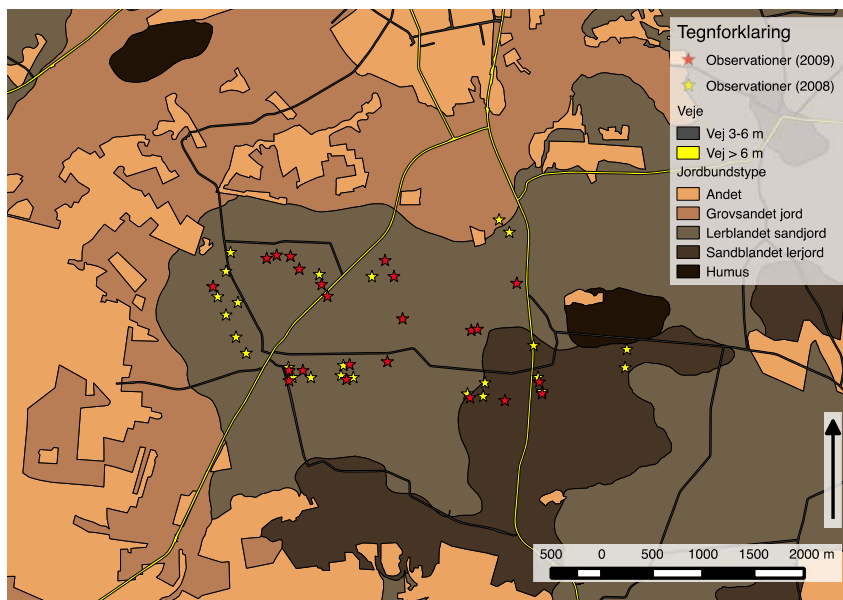
I dette afsnit præsenteres det danske landskab. På figur 3.2 på side 66 ses et kort over arealanvendelsen i det danske landskab med placeringen af observationerne fra 2008 og 2009. Som nævnt tidligere består området generelt af landbrugsjord. Der er spredt bebyggelse, vandløb og skov. Observationspunkterne i markerne er generelt placeret tæt på veje. Det skyldes, at feltmålingerne er udført i udkanten af markerne af hensyn til landmændene. På figur 3.3 på side 67 ses et kort over jordbunden i samme opløsning. De observerede marker ligger primært placeret i lerblandet sandjord og sandblandet lerjord.

Observationernes placering i forhold til jordbundstype ses på figur 3.3 på side 67. De fleste observationer er placeret i en lerblandet sandjord og et mindre antal i en sandblandet lerjord.

Vejrforhold

Vejrforholdene er undersøgt for at vurdere sandsynligheden for at jordbunden var våd eller tør.

Vha. DMIs vejrarkiv [Danmarks Meteorologiske Institut, 2014] er vejrforholdene for 31/5 2008 og 19/4 2009 undersøgt. I vejrarkivet findes oplysninger



Figur 3.3: Jordbundskort over området, hvor observationerne i Danmark i 2008 og 2009 er foretaget. Kortet er lavet i QGIS med data fra *Den danske jordklassificering* [n.d.].

om vejret på døgnbasis for de danske regioner. Da landskabet befinder sig i den nordlige del af Midtjylland, undersøges vejret for både Nordjylland og Midt- og Vestjylland.

Maj 2008 havde 2 nedbørsdage. Månedens samlede nedbørsmængde var 4 mm mod et gennemsnit på 51 mm i perioden 1961-1990. Gennemsnitstemperaturen var lidt højere end gennemsnittet i perioden 1961-1990 og antallet af solskinstimer pr. døgn væsentligt højere end gennemsnittet. Der faldt en anelse nedbør tre dage før 31/5, og der var tørt d. 31/5. Døgnets højeste temperatur var 25° C i Nordjylland og 30° C i Midt- og Vestjylland. Vinden var kraftigere i Nordjylland med ca. 8 m/s fra sydøst mod ca. 5 m/s fra øst-sydøst i Midt- og Vestjylland. Trykket var ca. 1015 hPa i Nordjylland mod ca. 1018 hPa i Midt- og Vestjylland og havde ligget stabilt over de foregående ti dage.

April 2009 var en måned med kun 2 nedbørsdage. Gennemsnitstemperaturen og antallet af solskinstimer pr. døgn var højere end gennemsnittet i perioden 1961-1990, mens nedbørsmængden var mindre end gennemsnittet. D. 19/4 2009 var døgnets højeste temperatur 15° C i Nordjylland og 15° C i Midt- og Vestjylland. Der var ligesom de foregående dage ingen nedbør og svag nordvestlig vind (under 5 m/s). Trykket var 1025 hPa og inde i en periode med stigende tryk over de foregående døgn.

På baggrund af ovenstående vurderes det overvejende sandsynligt, at jordbunden har været tør på begge tidspunkter, som er kendetegnet ved stabile højtryksperioder. Det kan dog ikke afvises, at der kan have været afvigende lokale vejrforhold.

3.2 Feltmålinger

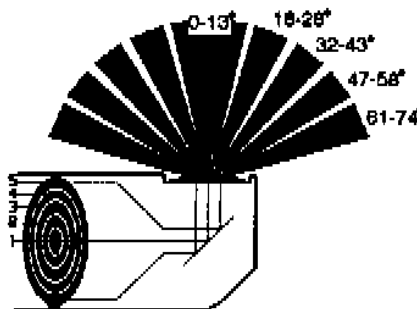
3.2.1 Bladarealindeks og middelbladvinkel

Bladarealindekset LAI er i Boegh et al. [2013] målt med LAI-2000-instrumentet. LAI-2000 ([LiCor, 1992]) måler spredningen af diffus stråling fra en plante-stand for at bestemme, hvor tætbevokset der er. Instrumentet måler desuden i fem forskellige vinkler i forhold til zenit for at bestemme løvets orienteringer. Som det ses på figur 3.4b projiceres de fem målinger ned i koncentriske cirkler, hvor den midterste cirkel repræsenterer strålingen fra lodret position. Jo mere yderlig cirklen er, desto tættere er strålingen på vandret. Se figur 3.4b. Der foretages målinger over og under plantestanden for at finde transmittansen, dvs. andelen af diffus stråling, der trænger gennem plantestanden i en bestemt vinkel. Resultatet af LAI-2000-målingen er *leaf area index* LAI, et mål for mængden af blade, og *mean tilt angle* MTA, middelbladvinklen. I praksis måler LAI-2000 ikke kun tætheden af bladene, men også af grene, stængler, frugter og andre plantelementer.

LAI-2000-målingerne er baseret på en række simple antagelser som beskrevet i LiCor [1992]. Først og fremmest antages bladene at være sorte. Hvis



(a) LAI-2000 LAI-2000 [n.d.]



(b) Vinkler LiCor [1992]

et blad er sort, vil al stråling absorberes og intet reflekteres, da emittansen fra et sort legeme er 0. På figur 2.1 ser man, at det meste lys i intervallet 400-490 nm reflekteres. Hvis man kun ser på dette spektral interval, kan man altså betragte grønne blade som tilnærmelsesvis sorte. Det optiske filter i LAI-2000 frasorterer netop stråling over 490 nm.

Dernæst antages det, at bladene er tilfældigt fordelt inden for nogle givne geometriske mønstre. Det kan fx være i parallelle rækker som rækker på en landbrugsmark. Det antages også, at bladene er relativt små. Endelig antages det ligesom i SAIL, at bladenes azimutvinkler er tilfældigt fordelt.

Teorien bag LAI-2000-estimatet af LAI er baseret på at måle mængden af huller i plantestanden i forskellige retninger.

Nilson [1971] giver en teoretisk behandling af, hvordan direkte solstråler penetrerer en plantestand. Strålen skal ramme et blad for at blive opfanget og brugt i fotosyntesen. Hvis solstrålen imidlertid rammer et hul i standen, vil den ikke blive opfanget. *Point quadrat analysis* er en metode til at undersøge fordelingen af plantelementer i en plantestand. Det er en metode, der rent praktisk udføres ved at stikke lange, tynde nåle ind i en plantestand i en fast vinkel og i systematiske placeringer. For hver gang nålen stikkes ind i et punkt registreres det, hvad nålen rammer. Det kan være, at nålen opnår kontakt med et eller flere blade. Det kan også være, at nålen rammer et hul i plantestanden, og dermed ikke opnår kontakt i det hele taget.

Der gælder følgende sammenhæng om sandsynligheden for at træffe et hul [Welles og Norman, 1991]:

$$T(\theta, \phi) = \exp(-G(\theta, \phi)\mu S(\theta, \phi))$$

$T(\theta, \phi)$ er sandsynligheden for, at en stråle ikke opfanges, når den trænger gennem plantestanden i retningen (θ, ϕ) , $G(\theta, \phi)$ er andelen af projiceret bladareal i retningen (θ, ϕ) , μ er tætheden af bladene angivet ved m^2 bladareal

pr. m^3 plantestand, og $S(\theta, \phi)$ er længden, strålen bevæger sig i m. Antages det, at azimutvinklerne er tilfældigt fordelte og den systematiske forskel kun afhænger af zenitvinklen, kan der ses bort fra ϕ og udtrykket kan omskrives:

$$G(\theta)\mu = -\frac{\log(T(\theta))}{S(\theta)} = K(\theta)$$

$K(\theta)$ er kontaktfrekvensen, dvs. det gennemsnitlige antal kontakter pr. længdeenhed, en nål i retning θ opnår. μ kan bestemmes eksakt:

$$\mu = 2 \int_0^{\frac{\pi}{2}} -\frac{\log(T(\theta))}{S(\theta)} \sin(\theta) d\theta \quad (3.1)$$

Tætheden af bladene μ afhænger af LAI (L) og plantestandens højde z på følgende vis:

$$L = \mu z$$

Det gælder, hvis plantestanden er homogen og dækkende. Der gælder også følgende sammenhæng mellem stilængden $S(\theta)$ og højden z :

$$S(\theta) = \frac{z}{\cos(\theta)}$$

Hvis man sætter de to udtryk for L og $S(\theta)$ ind i ligning 3.1, går z ud, da det multipliceres på begge sider:

$$L = 2 \int_0^{\frac{\pi}{2}} -\log(T(\theta)) \cos(\theta) \sin(\theta) d\theta \quad (3.2)$$

LAI-2000 måler i fem forskellige vinkler, og LAI bestemmes numerisk ud fra ligning 3.2:

$$L = 2 \sum_{i=1}^5 -\frac{\log(T_i)}{S_i} W_i$$

T_i er andelen af huller i retning i , S_i er stilængden i retning i , og W_i er $\sin(\theta)d\theta$ -værdien for retning i . W_i -erne summerer til 1.

LAI-2000 estimerer også middelbladvinklen MTA ud fra funktionen $G(\theta)$.

I datasættet fra Boegh et al. [2013] bestemmes LAI for en mark ved at udvælge to områder på 3 m x 3 m inden i et areal på 10 m x 10 m. I hvert område udføres fire målinger langs et transekt på 3 m. Hvis der er stor afvigelse mellem de to resultater, foretages endnu en måling. Gennemsnittet af målingerne angiver markens LAI. Estimer af middelbladvinklen indgår ikke i datasættet fra Boegh et al. [2013], men rådatafilen stillet til rådighed af Eva Bøgh for Danmark (2008) viser, at den gennemsnitlige middelbladvinkel for alle marker i landskabet er 59° med en standardafvigelse på 5° .



(a) SPAD-meter i brug Spectrum Technologies (b) SPAD-meters hoved Spectrum Technologies [n.d.a] [n.d.a]

Usikkerheden på LAI-estimerne vurderes af Boegh et al. [2013] som en kombination af måleusikkerheden for den enkelte mark og LAI-2000-instrumentets usikkerhed. For den enkelte mark bruges standardafvigelsen delt med middelværdien af de 2–4 målinger som mål for usikkerheden og er omkring 10%. Kombineret med instrumentusikkerheden vurderes den samlede usikkerhed at være 22–25%.

3.2.2 Bladklorofyl

I det følgende beskrives data for klorofylindhold fra Boegh et al. [2013].

Der er indsamlet data for klorofylindhold fra blade. Klorofylindholdet i et blad betegnes Chl_l . Målingerne er ikke-destruktive og foretaget med et håndholdt SPAD-meter.

SPAD-metret måler over et lille areal på $2 \text{ mm} \times 3 \text{ mm}$. Bladet kan maksimalt være 1.2 mm tykt. Der udsendes lys mod bladet, og en receptor på den anden side af bladet opfanger det lys, der passerer gennem bladet, dvs. det transmitterede lys. Ifølge Spectrum Technologies [n.d.b] udsendes rødt lys i bølgelængder omkring 650 nm og nærinfrarød stråling i bølgelængder omkring 940 nm. For blade vil en stor andel af det røde lys absorberes, mens den nærinfrarøde stråling i høj grad transmitteres og i mindre grad reflekteres. Der er ingen absorption af nærinfrarød stråling.

Instrumentet måler derefter transmittans for bladene. Det målte forhold mellem målt og udsendt stråling svarer til bladets transmittans er knyttet til bladenes indhold af klorofyl og kvælstof. SPAD-metrets resultat kaldes SPAD-indeks og er relateret til bladenes indhold af klorofyl og kvælstof. For at konvertere SPAD-indekset til klorofyl- og kvælstofindhold, skal der foretages en kalibrering, som udføres i laboratoriet på indsamlede blade.

Boegh et al. [2013] indsamler blade fra afgrøderne. Af hensyn til tilgængelighed er der for hver mark udvalgt et homogent udseende område tæt ved vejen. Her er der udført SPAD-målinger. For hvert blad er målingen udført to gange

og ved stor afvigelse mellem de to målinger en tredje gang. Målingerne udføres på et variabelt antal planter i det udvalgte område og i fem relative højder angivet ved en brøk ($\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$). For hver mark i hver højde er der typisk mellem 5 og 14 målinger.

Boegh et al. [2013] udfører kalibrering af SPAD-indekset i forhold til klorofyl- og kvælstofindhold ved en procedure nærmere beskrevet i artiklen. Kort sagt indsamles blade på det pågældende areal på den undersøgte mark ud fra deres farve, så farven varierer fra gullig til mørkegrøn afhængig af variationen på marken. For de indsamlede blade måles SPAD-indekset, og i laboratoriet udtrækkes klorofyl- og kvælstofindhold. Den specifikke vægt af hvert blad samt bladarealet estimeres. Man kan opgøre koncentration eller indhold af klorofyl og kvælstof i forhold til den specifikke vægt (typisk mg/g) eller arealenhed (typisk $\mu\text{g}/\text{cm}^2$). I dette tilfælde opgøres indholdet i forhold til arealenheden. Klorofylindhold i et blad målt i $\mu\text{g}/\text{cm}^2$ kaldes i denne sammenhæng bladklorofyl eller Chl_{blad} . Ud fra de indsamlede værdier af bladklorofyl findes relationen til SPAD-indekset. Jf. bl.a. Houborg og Anderson [2009] er der en eksponentiel sammenhæng mellem Chl_{blad} og SPAD-indeks, og samme kurve kan anvendes på tværs af afgrøder.

For en mark beregnes det gennemsnitlige bladklorofyl som gennemsnittet af Chl_{blad} i de relative højder, noteret $\overline{\text{Chl}}_{\text{blad}}$.

Derudfra estimeres plantestandens indhold af klorofyl ($\text{Chl}_{\text{total}}$) ved at multiplicere middelværdien af de målingerne i de fem højder med LAI:

$$\text{Chl}_{\text{total}} = \text{LAI} \cdot \overline{\text{Chl}}_{\text{blad}}$$

Boegh et al. [2013] vurderer usikkerheden på estimererne af $\overline{\text{Chl}}_{\text{blad}}$ og $\text{Chl}_{\text{total}}$. For $\overline{\text{Chl}}_{\text{blad}}$ tages en mulig vertikal struktur i betragtning og usikkerheden beregnes som en akkumuleret usikkerhed af varianser for hver relativ højde:

$$\Delta_{\text{blad}} = \frac{1}{5} \sqrt{\frac{\sigma_{\frac{1}{5}}^2 + \sigma_{\frac{2}{5}}^2 + \sigma_{\frac{3}{5}}^2 + \sigma_{\frac{4}{5}}^2 + \sigma_1^2}{\overline{\text{Chl}}_{\text{blad}}}}$$

Den gennemsnitlige usikkerhed på middelbladklorofyl for en mark er ca. 10 % for alle landskaberne, dog højere for Holland.

Usikkerheden på estimatet af $\text{Chl}_{\text{total}}$ afhænger af usikkerheden på $\overline{\text{Chl}}_{\text{blad}}$, SPAD-metrets usikkerhed, samt måleusikkerhed på LAI og LAI-2000-instrumentets usikkerhed.

$$\Delta_{\text{total}} = \sqrt{\Delta_{\text{blad}}^2 + \Delta_{\text{SPAD}}^2 + \Delta_{\text{LAI}}^2 + \Delta_{\text{LAI-2000}}^2}$$

Ifølge Boegh et al. [2013] er usikkerheden mellem 29 og 45% for alle landskaber.

Der er ni forskellige afgrøder ud over græs: Hvede, byg, majs, raps, tomat, artiskok, lucerne, kartofler og vinterrug. Profilerne er ikke målt for græs,

Hvede	30
Majs	13
Byg	10
Tomat	4
Raps	3
Lucerne	3
Kartofler	3
Artiskok	1
Vinterrug	1
I alt	68

Figur 3.4: Tabel over antal afgrødetyper for hvilke der er foretaget målinger af bladklorofyl i fem relative højder ned gennem plantestanden i datasættet fra Boegh et al. [2013].

	Grøn	Rød	NIR
μm	0.5-0.59	0.61-0.68	0.79-0.89

Tabel 3.1: Frekvensbånd for SPOT-satelitten.

hvorfor græs ikke indgår. Tabel 3.4 viser antal marker med de nævnte afgrøder for hvilke profilerne er målt. Der er markant flest hvedemarker, dernæst majs og byg, mens antallet af resterende afgrøder er fordelt på mellem én og fire marker.

3.3 Satellitdata

Der anvendes data fra SPOT (*Système pour l'Observation de la Terre*). Det drejer sig om multispektral data for radians i udvalgte frekvensbånd, som fremgår af tabel 3.1.

Der er indsamlet data fra SPOT 4 og SPOT 5 gennem instrumenterne HRVIR (High Resolution Visible and Infrared) og HRG (High Resolution Geometry). For HRVIR er den rumlige opløsning 20 m, mens HRG har en opløsning på 10 m. HRVIR er brugt til de danske marker i 2008 samt til Skotland. For de øvrige marker er opløsningen altså 10 m.

Data er behandlet af Boegh et al. [2013]. Der er udført geometrisk og atmosfærisk korrektion. Målingerne af radians er gennem modellen 6SV1 omregnet til overfladereflektans. Overfladereflektansen i de tre frekvensbånd er beregnet for 3 x 3 pixels centreret om den enkelte mark.

4. Metode

I dette kapitel opridses metoden baseret på gennemgangene i det forrige kapitel.

4.1 Karakterisering af vertikale klorofylprofiler

Det vertikale klorofylprofil undersøges ud fra højdemålinger af bladklorofyl for en enkelt mark ad gangen. For hver mark inddeles feltmålinger af klorofylindhold i fem niveauer svarende til de relative højder. For hvert niveau beregnes en middelværdi μ_i , $i = 1, 2, \dots, 5$ og en varians σ_i^2 , $i = 1, 2, \dots, 5$.

Det antages, at observationerne generelt er observerede værdier af normalfordelinger.

For at undersøge det vertikale klorofylprofil, dvs. den vertikale fordeling af klorofylindholdet ned gennem plantestanden, anvendes to statistiske modeller som beskrevet i afsnit 2.1.4 på side 26 og 2.1.4 på side 31. Modellerne adskiller sig ved fortolkningen af den forklarende variabel, højden i plantestanden. I variansanalysemodellen fortolkes højden som en kvalitativ faktor. Det betyder, at modellen ikke skelner mellem ordningen af højderne, men i stedet betragter bladklorofylmålingerne i hver enkelt højde som én gruppe bestående af normalfordelte observationer med middelværdi og varians. I regressionsmodellen fortolkes højden som en kvantitativ numerisk variabel.

Den statistiske fremgangsmåde er, at der opstilles en grundmodel med flest mulige parametre, som søges reduceret. I grundmodellen er den vertikale fordeling af bladklorofyl beskrevet ud fra fem par af parametre: Middelværdi μ_i og varians σ_i^2 for et niveau i , hvor $i = 0.2, 0.4, 0.6, 0.8, 1.0$, dvs. i alt ti parametre. Første skridt at bruge Bartlett's test til at undersøge hypotesen om ens varianser:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_5^2$$

Hvis der ikke er signifikant forskel på varianserne med signifikansniveau $\alpha = 0.05$, accepteres hypotesen, og antallet af parametre i modellen er således seks: Middelværdi μ_i for hvert niveau i og variansen σ^2 . Næste skridt er at anvende ensidet variansanalyse, hvor niveauerne i opfattes som kvalitative faktorer for at be- eller afkræfte, om der er signifikant forskel på middelværdierne. Hypotesen lyder:

$$H_1 : \mu_1 = \mu_2 = \dots = \mu_5$$

Hvis der ikke er signifikant forskel på middelværdierne med signifikansniveau $\alpha = 0.05$, accepteres hypotesen. Antallet af parametre i modellen er i

dette tilfælde reduceret til to: μ og σ^2 . I så fald omtales klorofylprofilet som uniformt. Hvis der er signifikant forskel på middelværdierne, forkastes hypotesen, og klorofylprofilet er ikke uniformt. I dette tilfælde anvendes simpel lineær regression for at undersøge om middelværdistrukturen kan beskrives som en lineær sammenhæng mellem relativ højde og bladklorofyl. Hypotesen lyder:

$$H_2 : EY_i = \alpha + \beta x_i$$

Hvis β er signifikant forskellig fra 0 med signifikansniveau $\alpha = 0.05$, accepteres hypotesen. Antallet af parametre i modellen er således reduceret til tre: α , β og σ^2 . I så fald omtales klorofylprofilet som lineært.

4.2 Variation i datasættet

Statistiske metoder anvendes til at undersøge variation i datasættets feltmålinger. Eventuelle afgrødespecifikke sammenhænge mellem højde, bladarealindeks og klorofylindhold undersøges.

4.3 Empirisk-statistiske modeller

Data for overfladereflektans undersøges med henblik på at undersøge variationen mellem landskaberne.

Den empiriske sammenhæng mellem LAI og de spektrale vegetationsindekser undersøges ud fra de forventede sammenhænge, nemlig at der vil være en tilnærmelsesvis lineær sammenhæng mellem SR og LAI, at de normerede vegetationsindekser (NDVI, GNDVI og EVI2) vil følge Baret og Guyots model og konvergere mod et maksimum for højt LAI, og at klorofylindekset CI vil være knyttet til klorofylindholdet i plantestanden.

Dvs. at følgende sammenhænge undersøges:

$$\begin{aligned} \text{SR} &= \alpha + \beta \text{ LAI} \\ \text{NDVI} &= \text{NDVI}_{\max} + (\text{NDVI}_{\min} - \text{NDVI}_{\max}) \exp(-K \cdot \text{LAI}) \\ \text{GNDVI} &= \text{GNDVI}_{\max} + (\text{GNDVI}_{\min} - \text{GNDVI}_{\max}) \exp(-K \cdot \text{LAI}) \\ \text{EVI2} &= \text{EVI2}_{\max} + (\text{EVI2}_{\min} - \text{EVI2}_{\max}) \exp(-K \cdot \text{LAI}) \\ \text{CI} &= \alpha + \beta \text{ LAI} \\ \text{CI} &= \alpha + \beta \text{ Chl}_{\text{total}} \end{aligned}$$

Der anvendes lineær og ikke-lineær regression i R. Til vurdering af resultatet af den lineære regression produceres fraktildiagrammer, residualplots og beregnes determinationskoefficienter (R^2 -værdier). Til vurdering af resultatet

af den ikke-lineære regression produceres også residualplots. Jf. afsnit 2.1.4 på side 33 kan determinationskoefficienten ikke beregnes for en ikke-lineær regressionsmodel.

Det undersøges desuden om det vertikale klorofylprofil spiller en rolle i de beskrevne sammenhænge, dvs., om korrelationen mellem vegetationsindekserne og LAI eller $\text{Chl}_{\text{total}}$ forbedres for marker med ensartet vertikalt klorofylprofil. På samme vis undersøges landskabets betydning i den empiriske sammenhæng. Inden for det enkelte landskab undersøges også, om et uniformt vertikalt klorofylprofil forbedrer den empiriske sammenhæng med LAI eller $\text{Chl}_{\text{total}}$ for vegetationsindekserne.

4.4 PROSAIL

4.4.1 Følsomhedsanalyse af PROSAIL

PROSAIL er implementeret i Matlab af Jacquemond et al. [2013]. Formålet med følsomhedsanalysen af PROSAIL er at undersøge, hvordan en ændring i inputparametrene påvirker den simulerede reflektans. PROSAIL simulerer reflektansspektret fra 400–2500 nm med skridt på 1 nm. For at sammenligne resultaterne med data fra SPOT beregnes middelværdier af den simulerede reflektans i de tre frekvensbånd fra tabel 3.1 på side 73: Grøn (500–590 nm), rød (610–690 nm) og nærinfrarød (790–890 nm). Følsomheden over for inputparametrene vurderes i forhold til den relative ændring af middelværdierne i de tre frekvensbånd.

Følsomhedsanalysen baseres på en mark fra Danmark (2008) med repræsentative egenskaber og undersøges ved forskellige værdier af LAI samt for våd og tør jordbund. De inputparametre, der er ukendte, fastsættes ud fra litteratur på området. Til brug for følsomhedsanalyse udvælges en mark med repræsentative egenskaber i forhold til bladklorofyl, vertikalt klorofylprofil, bladarealindeks og højde.

Alle inputparametre justeres én af gangen med $\pm 10\%$ for at vurdere den relative ændring i simuleret grøn, rød og nærinfrarød reflektans.

4.4.2 Vegetationsindeksernes følsomhed

Følsomhedsanalysen anvendes også til at vurdere de spektrale vegetationsindekseres følsomhed over for ændringer i PROSAIL's inputparametre. Vha. de beregnede middelværdier for reflektans af grønt og rødt lys samt nærinfrarød stråling beregnes vegetationsindekser. Dernæst beregnes de relative ændringer i vegetationsindekserne i forhold til variation af inputparametrene. Undersøgelsen foretages for forskellige niveauer af LAI. Vurderingen af vegetationsindeksernes følsomhed skal bidrage til at vurdere betydningen af det vertikale klorofylprofil i forhold til sammenhængen mellem vegetationsindekserne og LAI.

4.4.3 Simulering af reflektans

PROSAIL anvendes til at simulere reflektans for alle observationer i Danmark (2008) og Danmark (2009) med det formål at vurdere om det vertikale klorofylprofil har en betydning for sammenhængen mellem målt og simuleret reflektans.

Reflektansen simuleres for den grønne, røde og nærinfrarøde del af det elektromagnetiske spektrum. Derefter undersøges de relative og absolutte afvigelser mellem simuleret og målt reflektans. For de absolutte afvigelser udføres ensidet variansanalyse for at undersøge, om der er systematisk forskel på afvigelserne for forskellige karakteristika som vertikalt klorofylprofil, LAI, plantetype eller landskab.

5. Resultater

5.1 Landskabsvariation

I dette afsnit beskrives variationen i og i mellem landskaberne ud fra anvendte feltmålinger af højde, LAI og bladklorofyl foretaget af Boegh et al. [2013].

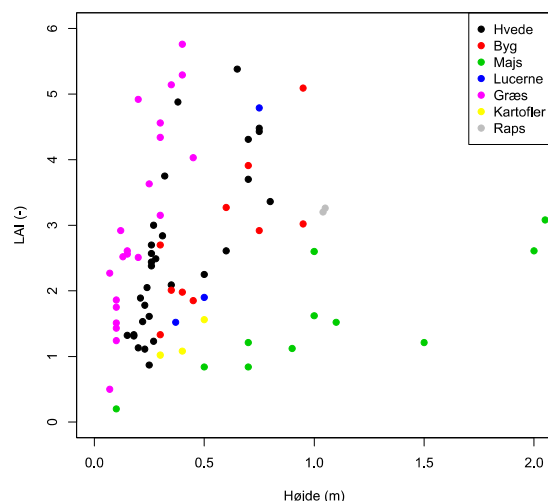
For alle marker er der målt en gennemsnitlig højde af plantestanden. Bladarealindekset er ligeledes målt for hver mark. Planternes højde afhænger af afgrødetypen. Græsmarkerne er lavest, hvede og byg er typisk under 1 m høje. Majsmarkerne har en stor spredning med nogle høje marker på omkring 2 m og enkelte lave marker.

I forhold til LAI er der en stor variation mellem afgrøderne. Bartletts test (udført med funktionen `bartlett.test` i R) for varianshomogenitet viser, at varianserne er ens med $p = 0.05$. Ensidet variansanalyse (udført med funktionen `aov` i R) viser, at der er signifikant forskel på middelværdierne for LAI mellem afgrøderne. Hypotesen om ens middelværdier forkastes med $p = 0.04$. Afgrødetyperne kan i stedet beskrives ud fra middelværdier for de enkelte afgrødetyper som vist i tabel 5.1, hvor også de gennemsnitlige højder er vist.

Således er der grund til at antage, at både højder og bladarealindeks afhænger af afgrøderne. På figur 5.1 på side 79 ses forholdet mellem højde og bladarealindeks for de hyppigst optrædende afgrøder. Simpel lineær regression viser, at der for hver enkelt afgrøde er en lineær sammenhæng mellem højde og bladarealindeks. For græs ($R^2 = 0.74$), hvede ($R^2 = 0.57$) og byg ($R^2 = 0.61$) er hældningen stejl, dvs., at der er en høj tæthed af blade ved relativt lave planter. For majs ($R^2 = 0.75$) er hældningen fladere, dvs., at majs er kendetegnet ved

Afgrøde	Højde (m)	LAI (-)	Chl _{blad} ($\mu\text{g}/\text{cm}^2$)	Chl _{total} ($\mu\text{g}/\text{cm}^2$)
Lucerne	0.54	2.74	57.29	152.12
Byg	0.58	2.81	38.37	107.32
Græs	0.21	3.07	32.26	105.46
Majs	1.24	1.70	62.35	104.27
Kartofler	0.40	1.22	37.22	45.96
Raps	1.05	3.23	71.74	216.69
Hvede	0.37	2.68	38.96	99.60

Tabel 5.1: Tabellen viser middelværdier for højde, bladarealindeks (LAI), bladklorofyl og samlet klorofylindhold for de afgrøder, der indgår i datasættet fra Boegh et al. [2013] på flere end to marker, beregnet i R med funktionen `lm`.



Figur 5.1: Figuren viser højde af plantestanden i forhold til bladarealindeks (LAI) for de hyppigst forekommende afgrøder i datasættet fra Boegh et al. [2013].

en relativt lav tæthed af blade for høje planter op til 2 m. For byg og majs er skæringen med y -aksen ikke signifikant, dvs. sammenhængen er proportional. For græs og hvede er skæringen dog signifikant men relativt lav.

For afgrødetyperne ses det i tabel 5.1 på side 78, at byg og hvede har relativt lavt klorofylindhold på omkring $40 \mu\text{g}/\text{cm}^2$, men klorofylindholdet for græs er det laveste. Hos majs ses den største spredning og generelt høje værdier med en enkelt værdi på over $120 \mu\text{g}/\text{cm}^2$. Der er få observationer af raps, og de ligger højt, mens de få observationer af kartofler har et relativt lavt klorofylindhold.

Samlet set kan det konkluderes, at højde, LAI og klorofylindhold afhænger af plantetype.

5.2 Vertikale klorofylprofiler

I dette afsnit præsenteres resultatet af undersøgelsen af de vertikale klorofylprofiler. Som beskrevet tidligere antages det, at observationerne er normalfordelte. En måde at undersøge dette på er vha. fraktildiagrammer. Nogle eksempler på fraktildiagrammer for udvalgte marker i datasættet fra Boegh et al. [2013] er vist i kapitel C.1 på side 140.

Det samlede antal marker, der testes for vertikale klorofylprofiler, er 68. Græsmarkerne fra Holland har ikke data for klorofyl i forskellige niveauer, hvorfor de ikke indgår i undersøgelsen.

	Hvede	Majs	Byg	Tomat	Raps	Lucerne	Kartofler	Artiskok	Rug	Total
Ingen varianshomogenitet	9	7				2				18
Uniform fordeling	11	4	4	2	3	1	3			28
Lineær sammenhæng	8	1	4	2				1		16
Forskellige middelværdier	2	1	2						1	6
Total	30	13	10	4	3	3	3	1	1	68

Tabel 5.2: Tabellen viser resultatet af de statistiske tests vedrørende et eventuelt vertikalt klorofylprofil i forhold til afgrøder. Der er udført ensidet variansanalyse og simpel lineær regression og hypotesen om ens varianser er testet med Bartlett's test. Tests er udført i R.

Test for varianshomogenitet udføres med funktionen `bartlett.test` i R. Der er 18 marker, som ikke har varianshomogenitet ($p < 0.05$).

For de resterende 50 marker accepteres hypotesen om varianshomogenitet, hvorfor der udføres variansanalyse, hvor de relative højder opfattes som kvalitative faktorer vha. funktionen `aov` med `as.factor` i R. For 28 marker kan hypotesen om ens middelværdier ikke forkastes, og der er ikke signifikant forskel på middelværdierne ($p > 0.05$). For disse marker betegnes det vertikale klorofylprofil som uniformt.

For de 22 resterende marker testes hypotesen om en lineær sammenhæng mellem de relative højder og bladklorofylindholdet vha. funktionen `lm` i R. Hældningskoefficienten β er signifikant for 16 marker. For disse marker betegnes det vertikale klorofylprofil som lineært. Dvs., at der er 6 marker, for hvilke middelværdierne er signifikant forskellige, men der er ikke tale om en lineær sammenhæng mellem relativ højde og bladklorofyl.

Resultatet af undersøgelsen ses fordelt på afgrøder i tabel 5.2 på side 80 og fordelt på lokaliteter i tabel 5.3 på side 81. Det ses, at markerne uden varianshomogenitet mellem bladklorofylindholdet i de relative højder fordeler sig på afgrøderne hvede, majs og lucerne. Blandt alle afgrøderne er der marker med uniformt klorofylprofil. Markerne med lineært klorofylprofil ses hos 8 hvedemarker, 4 bygmarker, en majsmark og en tomatmark.

Opdelt på lokaliteter ses det, at der er flest marker med uniformt klorofylprofil i Danmark (2009) og flest marker med lineært klorofylprofil i Danmark (2008). De italienske marker fordeler sig jævnt over kategorierne, mens de polske marker primært er med uniformt klorofylprofil eller uden varianshomogenitet. Der indgår kun to hollandske marker, eftersom de resterende hollandske græsmarker ikke kan testes for vertikale klorofylprofiler pga. manglende data.

Alt i alt ses det, at alle landskaber og alle plantetyper er repræsenteret i

	Danmark (2008)	Danmark (2009)	Italien	Holland	Polen	Total
Ingen varianshomogenitet	3	6	3	2	4	18
Uniform fordeling	6	12	4		6	28
Lineær sammenhæng	9	2	4		1	16
Forskellige middelværdier	2	1	1		2	6
Total	20	21	12	2	13	68

Tabel 5.3: Tabellen viser resultatet af de statistiske tests vedrørende et eventuelt vertikalt klorofylprofil i forhold til lokaliteter. Der er udført ensidet variansanalyse og simpel lineær regression og hypotesen om ens varianser er testet med Bartlett's test. Tests er udført i R.

	Voksende	Aftagende
Hvede	6	2
Majs	–	1
Byg	4	–
Tomat	1	1
Artiskok	1	–
Danmark (2008)	9	–
Danmark (2009)	–	2
Italien	2	2
Polen	1	–

Tabel 5.4: Tabellen viser fordelingen af henholdsvis voksende og aftagende lineære sammenhænge mellem klorofylindholdet i bladene i forhold til den relative højde fra bund til top for plantetyper såvel som for landskaber.

de forskellige kategorier vedrørende den vertikale fordeling af klorofyl ud fra datasættet. For majsplanten er der flest marker uden varianshomogenitet i klorofylindholdet i de relative højder. For landskabernes vedkommende har både Danmark (2009) og Polen flest marker med en uniform vertikal fordeling af klorofyl. Danmark (2008) har til gengæld flest marker med en lineær vertikal fordeling af klorofyl. Der er kun få marker med forskellige middelværdier gennem det vertikale klorofylprofil.

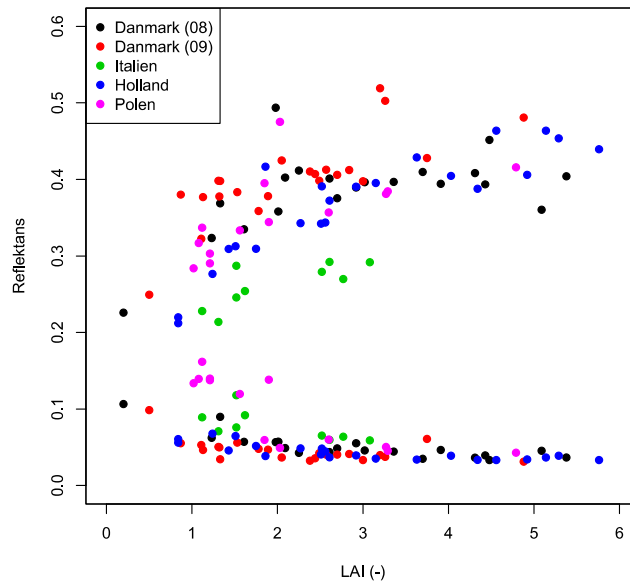
I afsnit B på side 136 ses eksempler på plots af de forskellige typer vertikalt klorofylprofil.

5.3 Empirisk sammenhængen mellem overfladereflektans og bladarealindeks

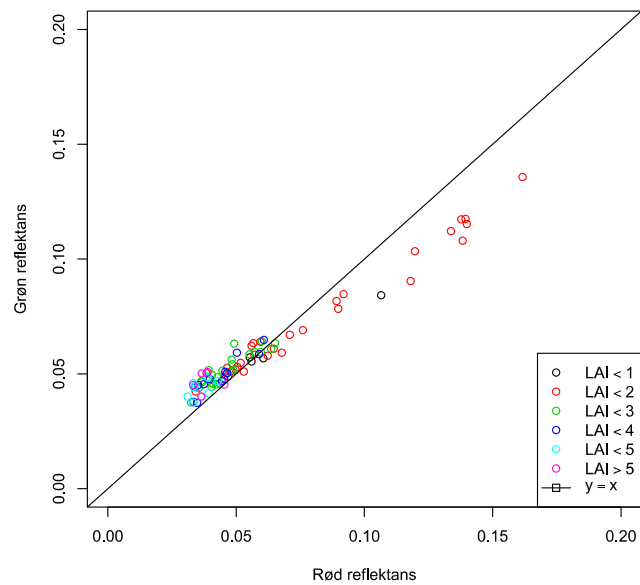
På figur 5.2 på side 83 ses rød og nærinfrarød reflektans i forhold til bladarealindeks (LAI) for landskaberne i datasættet.

Den forventede sammenhæng mellem LAI og reflektans er, at der for en given tærskelværdi af LAI vil ske en asymptotisk udvikling af kurven, hvor reflektansen ikke længere påvirkes af tætheden af bladene. For rød reflektans forventes grænseværdien at blive nået ved højt LAI, hvor reflektansen når et minimum. For nærinfrarød reflektans er der tale om et maksimum, der indtræder ved højere værdier af LAI normalt over 8. Som det ses på figur 5.2 på side 83 har den røde reflektans en tydelig asymptotisk opførsel på ca. 4–6 % i forhold til LAI over 2. For nærinfrarød reflektans er variationen større, hvilket stemmer godt overens med, at måtningen først sker ved store værdier af LAI. Især de danske landskaber samt Holland har en tydelig asymptotisk udvikling af den nærinfrarøde reflektans. For det italienske landskab er den nærinfrarøde reflektans påfaldende lavere, mens den for det polske landskabe er mere spredt. En lavere nærinfrarød reflektans ved lavt LAI kan skyldes, at jordbunden har en mørk farve.

Sammenhængen mellem rød og grøn reflektans som ses på figur 5.3 på side 84 for forskellige værdier af bladarealindekset LAI. Reflektans fra jordbunden ved lavt vegetationsdække er kendetegnet ved en lidt højere reflektans af rødt lys i forhold til grønt lys. For voksende LAI vil den røde reflektans nå et minimum, mens den grønne reflektans vil opnå et peak, der afhænger af plantespecifikke faktorer. På figuren er linjen $y = x$ tegnet ind. Hvis et punkt ligger på linjen er den grønne og røde reflektans lige stor. Et punkt, der ligger under linjen repræsenterer en højere rød end grøn reflektans, og et punkt, der ligger over linjen repræsenterer en lavere rød end grøn reflektans. Figuren viser, at reflektansen falder ved stigende LAI. Derudover ses det, at nogle af punkterne, der især repræsenterer marker med LAI under 2, har højere rød end grøn reflektans. For LAI over 2 og højere er den grønne reflektans størst. For yderligere stigende LAI når den røde reflektans et minimum, der får punkterne til at klumpe sammen.



Figur 5.2: Figuren viser rød og nærinfrarød reflektans for marker i de enkelte landskaber. For hver mark er der angivet både rød og nærinfrarød reflektans. Det laveste punkt er den røde reflektans, mens det højeste punkt er den nærinfrarøde reflektans. Data fra Boegh et al. [2013].



Figur 5.3: Figuren viser grøn reflektans i forhold til rød reflektans for alle marker i datasættet fra Boegh et al. [2013]. For alle marker er intervallet for værdien af bladarealindekset (LAI) angivet.

Landskab	Lineær regressionsmodel	R^2
Alle	SR = 2.97 + 1.92LAI	0.50
Danmark (2008)	SR = 3.77 + 1.54LAI	0.62
Danmark (2009)	SR = 5.48 + 1.76LAI	0.46
Italien	SR = 1.08 + 1.26LAI	0.77
Holland	SR = 3.44 + 1.90LAI	0.79
Polen	SR = 0.21 + 2.28LAI	0.68
Uniformt klorofylprofil	SR = 0.33 + 3.34LAI	0.65
Uniformt klorofylprofil, DK09	SR = 4.79 + 2.23LAI	0.70

Tabel 5.5: Tabellen viser lineære regressionsmodeller for sammenhængen mellem SR og LAI for de enkelte landskaber. Til sammenligning ses regressionsmodel for alle landskaber samlet samt alle marker med et uniformt vertikalt klorofylprofil.

5.4 Empirisk-statistiske modeller

I dette afsnit præsenteres resultater vedrørende sammenhængen mellem de spektrale vegetationsindekser og henholdsvis LAI og $\text{Chl}_{\text{total}}$.

5.4.1 Simple ratio

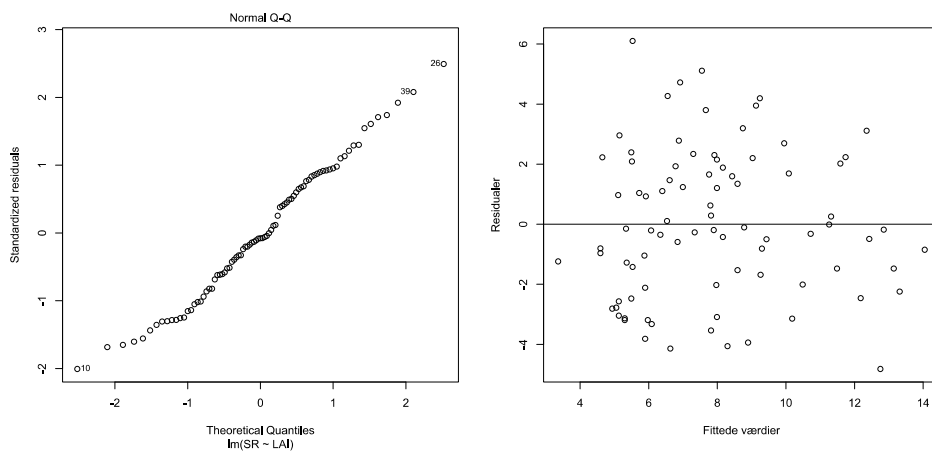
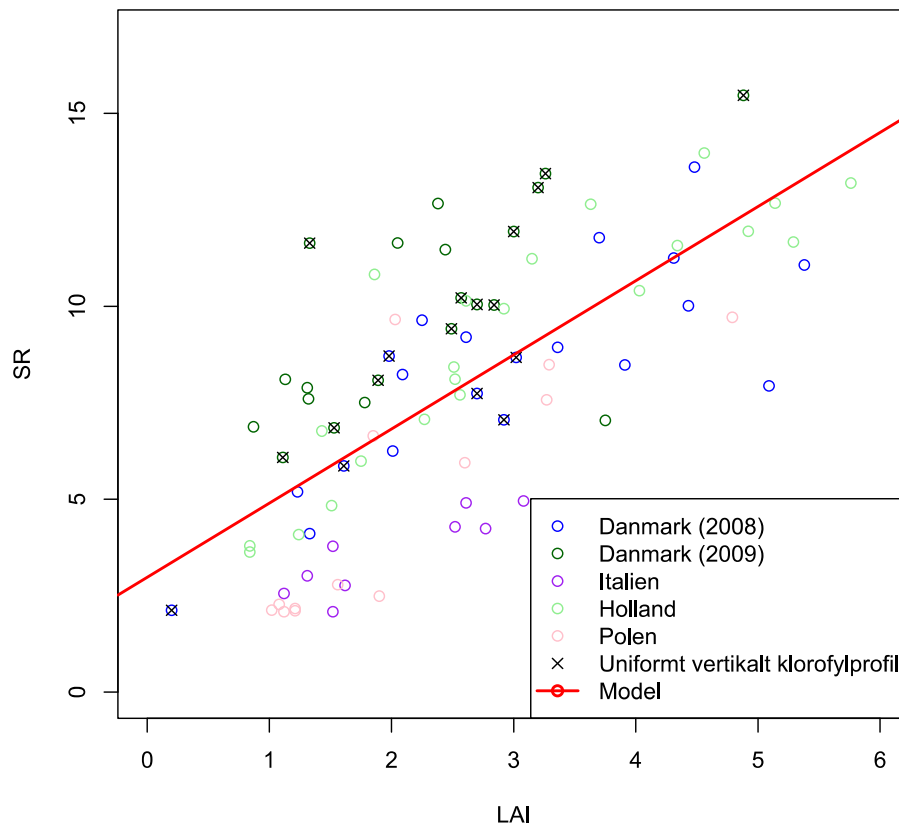
Simple ratio (SR) er beskrevet i afsnit 2.2.2:

$$\text{SR} = \frac{\rho_{\text{NIR}}}{\rho_{\text{rød}}}$$

Det forventes, at der er en lineær sammenhæng mellem SR og LAI. På figur 5.4 på side 86 ses et plot af SR i forhold til LAI for alle områder med en fælles lineær regressionslinje samt et fraktildiagram og et residualplot, der viser residualerne i forhold til de fittede y -værdier. Residualplottet viser, at der ikke er nogen systematisk afvigelse fra den lineære regressionsmodel. Fraktildiagrammet viser, at data er normalfordelt. For alle landskaberne er R^2 -værdien 0.50, og R^2 -værdierne er højere for de enkelte landskaber med undtagelse af Danmark (2009), se tabel 5.5 på side 85. For markerne med et uniformt vertikalt klorofylprofil er R^2 -værdien 0.65, dvs. højere end for alle marker samlet set, men lavere end flere landskaber. Der er også udført lineær regression for de 12 marker med uniformt vertikalt klorofylprofil i Danmark (2009), hvilket kraftigt forbedrer R^2 -værdien for landskabet. For de øvrige landskaber er der højst seks marker med uniformt vertikalt klorofylprofil, hvorfor en lineær regressionsmodel ikke anvendes.

5.4.2 NDVI, GNDVI og EVI2

NDVI, GNDVI og EVI2 er alle normerede vegetationsindekser:



Figur 5.4: Figuren viser et plot af SR i forhold til LAI for alle landskaber samt et fraktildiagram og et residualplot.

$$\begin{aligned} \text{NDVI} &= \frac{\rho_{\text{NIR}} - \rho_{\text{rød}}}{\rho_{\text{NIR}} + \rho_{\text{rød}}} \\ \text{GNDVI} &= \frac{\rho_{\text{NIR}} - \rho_{\text{grøn}}}{\rho_{\text{NIR}} + \rho_{\text{grøn}}} \\ \text{EVI2} &= \frac{2(\rho_{\text{NIR}} - \rho_{\text{rød}})}{(\rho_{\text{NIR}} + \rho_{\text{rød}} + 1)} \end{aligned}$$

Pga. normeringen vil vegetationsindekserne have en tendens til at blive mættet for store værdier af LAI. I afsnit 2.2.2 på side 44 beskrives følgende model fra Baret og Guyot [1991] for sammenhængen mellem et vegetationsindeks og LAI:

$$\text{VI} = \text{VI}_{\text{max}} + (\text{VI}_{\text{min}} - \text{VI}_{\text{max}}) \exp(-K \cdot \text{LAI})$$

I R udføres ikke-lineær regression til Baret og Guyots model for NDVI, GNDVI og EVI2. For at lave modelkontrol laves to typer af residualplots: Det første viser residualerne i forhold til x -værdierne, og det andet viser residualerne i forhold til de fittede y -værdier.

Først undersøges sammenhængen mellem NDVI og LAI for alle landskaber. Resultatet ses på figur 5.5 på side 88. Det første residualplot viser, at residualerne er store ved lavt LAI og små ved højt LAI. Det tyder på en systematisk fejl i modellen. Fejlen skyldes, at der er for mange marker med næsten ens LAI, der giver vidt forskellige NDVI-værdier. Derfor undersøges sammenhængen igen for de to danske landskaber, Danmark (2008) og Danmark (2009). På figur 5.6 på side 89 ses det, at residualplottet viser en tilfældig fordeling af residualerne. Det tyder på, at modellen passer bedre for de danske landskaber. Der er dog kun ét punkt med meget lavt LAI. NDVI lader altså til at være stærkt knyttet til landskabet. Markerne med uniformt vertikalt klorofylprofil er også markeret på figuren. Der ses ikke nogen særlig forbedring ved disse marker i forhold til modellen.

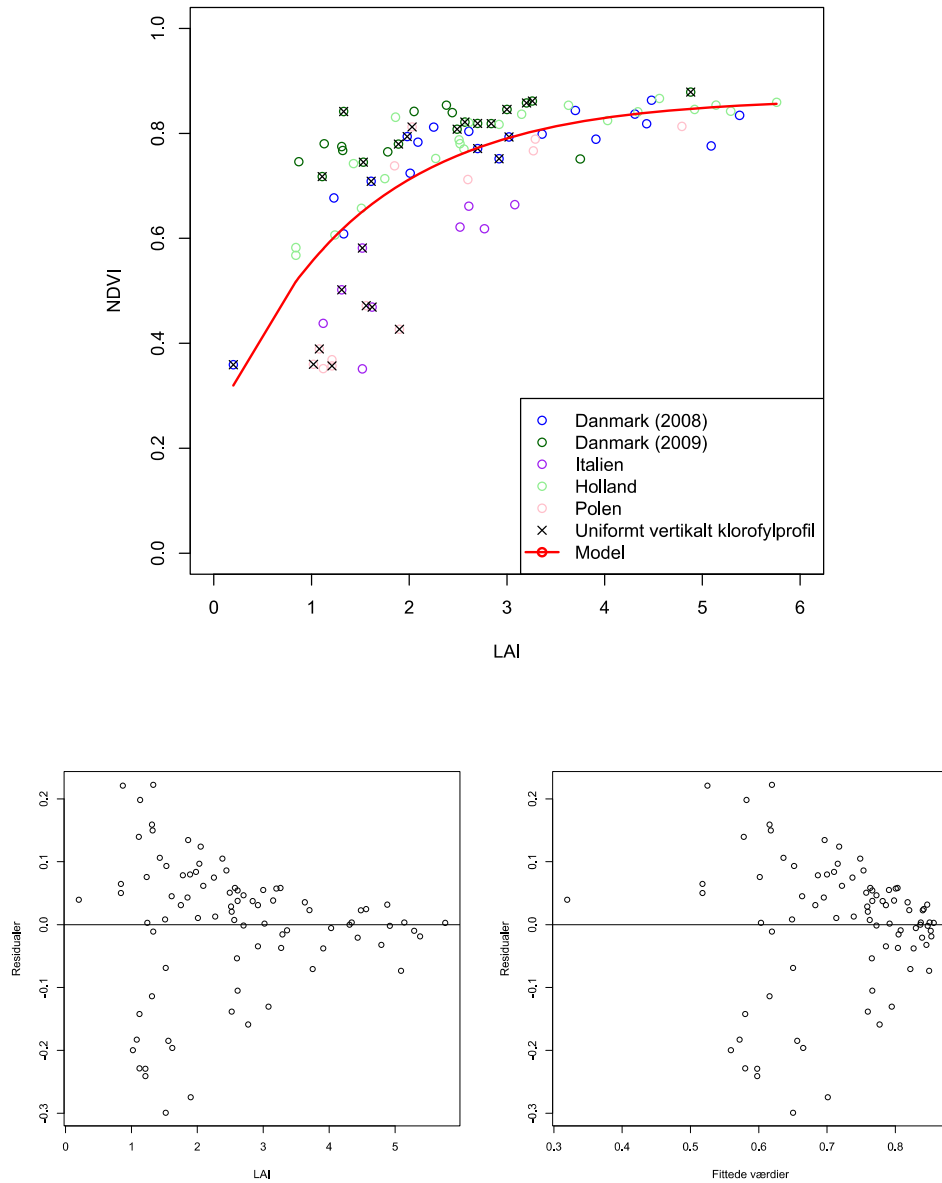
For de danske landskaber er sammenhængen mellem hhv. GNDVI og EVI2 og LAI også undersøgt og resultaterne ses på figur D.1 på side 142 og figur D.2 på side 143 i kapitel D på side 141. Generelt kan man se, at alle tre vegetationsindekser ikke udviser følsomhed over for LAI over ca. 2.

5.4.3 Klorofylindekset

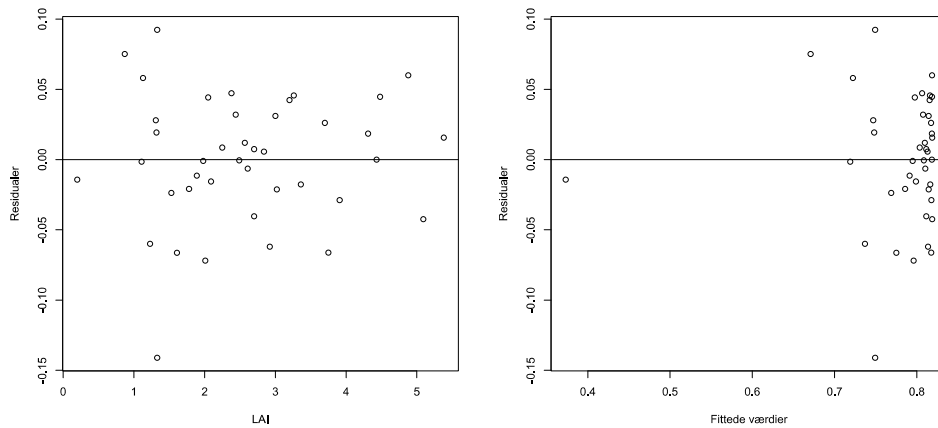
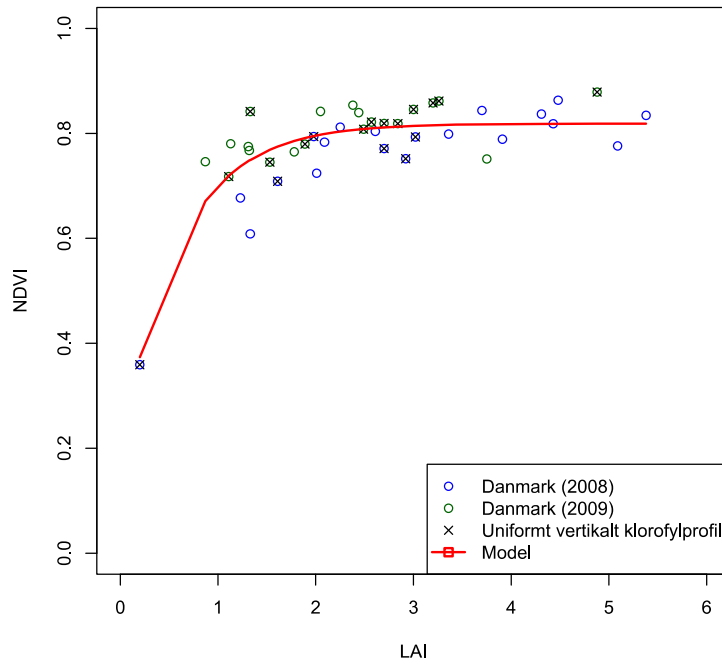
Klorofylindekset (CI) er beskrevet i afsnit 2.2.2:

$$\text{CI} = \frac{\rho_{\text{NIR}}}{\rho_{\text{grøn}}} - 1$$

CI er undersøgt i forhold til både LAI og $\text{Chl}_{\text{total}}$ pga. den forventede følsomhed over for klorofyl. På figur 5.7 på side 91 og figur 5.8 på side 92 ses data



Figur 5.5: Figuren viser sammenhængen mellem NDVI og LAI for de alle landskaber. $NDVI_{\max}$ estimeres til 0.87, $NDVI_{\min}$ til 0.24 og udslukningskoefficienten er 0.70.



Figur 5.6: Figuren viser sammenhængen mellem NDVI og LAI for de danske landskaber. $NDVI_{max}$ estimeres til 0.82, $NDVI_{min}$ til 0.20 og udslukningskoefficienten er 1.64.

Landskab	Lineær regressionsmodel	R^2
Alle	$CI = 2.56 + 1.36LAI$	0.48
Danmark (2008)	$CI = 3.03 + 1.29LAI$	0.66
Danmark (2009)	$CI = 4.93 + 1.19LAI$	0.47
Italien	$CI = 0.62 + 1.11LAI$	0.78
Holland	$CI = 3.08 + 1.11LAI$	0.78
Polen	$CI = -0.09 + 1.86LAI$	0.76
Uniformt vertikalt klorofylprofil	$CI = 0.44 + 2.50LAI$	0.65
Uniformt vertikalt klorofylprofil, DK09	$CI = 4.34 + 1.49LAI$	0.73
Alle	$CI = 3.80 + 0.02Chl_{total}$	0.25
Danmark (2008)	$CI = 3.24 + 0.02Chl_{total}$	0.62
Danmark (2009)	$CI = 5.93 + 0.03Chl_{total}$	0.48
Italien	$CI = 1.06 + 0.01Chl_{total}$	0.50
Holland	$CI = 4.01 + 0.02Chl_{total}$	0.43
Polen	$CI = 1.10 + 0.03Chl_{total}$	0.62
Uniformt vertikalt klorofylprofil	$CI = 2.73 + 0.03Chl_{total}$	0.43
Uniformt vertikalt klorofylprofil, DK09	$CI = 6.08 + 0.02Chl_{total}$	0.58

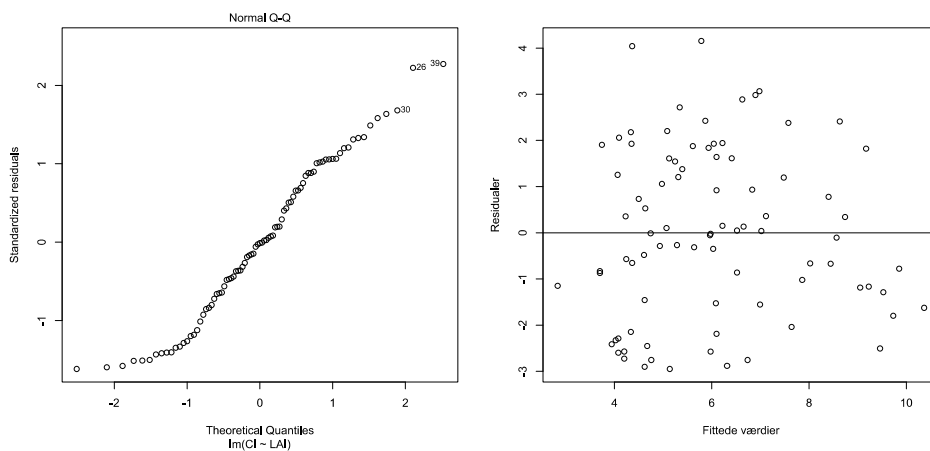
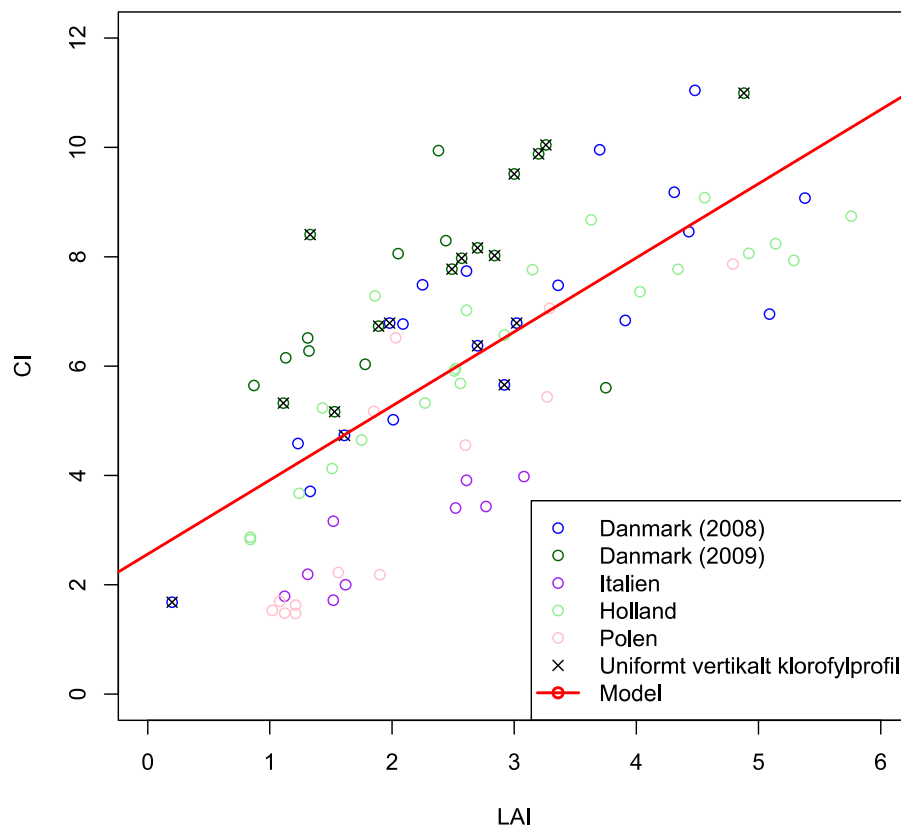
Tabel 5.6: Tabellen viser lineære regressionsmodeller for sammenhængen mellem klorofylindekset (CI) og LAI for de enkelte landskaber.

med lineære regressionsmodeller for begge sammenhænge for alle landskaber. Fraktildiagrammerne viser, om data er normalfordelt. Residualplots viser, at residualerne ligger tilfældigt, hvorfor den lineære regressionsmodel ikke har en systematisk fejl. På tabel 5.6 på side 90 ses en oversigt over de lineære regressionsmodeller samt R^2 -værdier. Generelt er billedet det samme som for SR i forhold til LAI. R^2 -værdierne er højere for de enkelte landskaber end for alle landskaber samlet set med undtagelse af Danmark (2009). Det gælder for både sammenhængen med LAI og sammenhængen med Chl_{total} . Dog er R^2 -værdierne generelt højere for sammenhængen med LAI end Chl_{total} . For markerne med uniformt vertikalt klorofylprofil er der også højere R^2 -værdi end for alle landskaber samlet, men den er lavere end for de enkelte landskaber med undtagelse af Danmark (2009). For Danmark (2009) er lineær regression dog også udført udelukkende for marker med uniformt vertikalt klorofylprofil, hvilket i begge tilfælde øger R^2 -værdien betydeligt. Dog er R^2 stadig lavere end for øvrige enkelte landskaber. Fraktildiagrammet på figur 5.8 på side 92 ser endvidere ud til at afvige fra en normalfordeling.

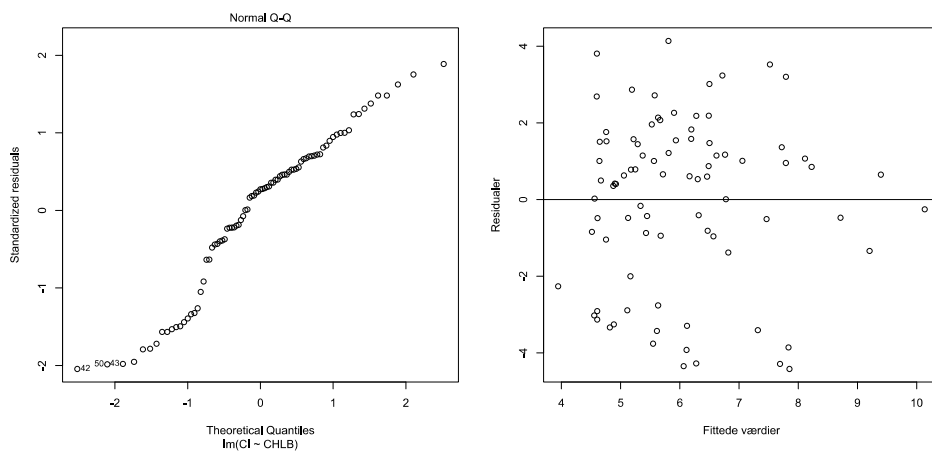
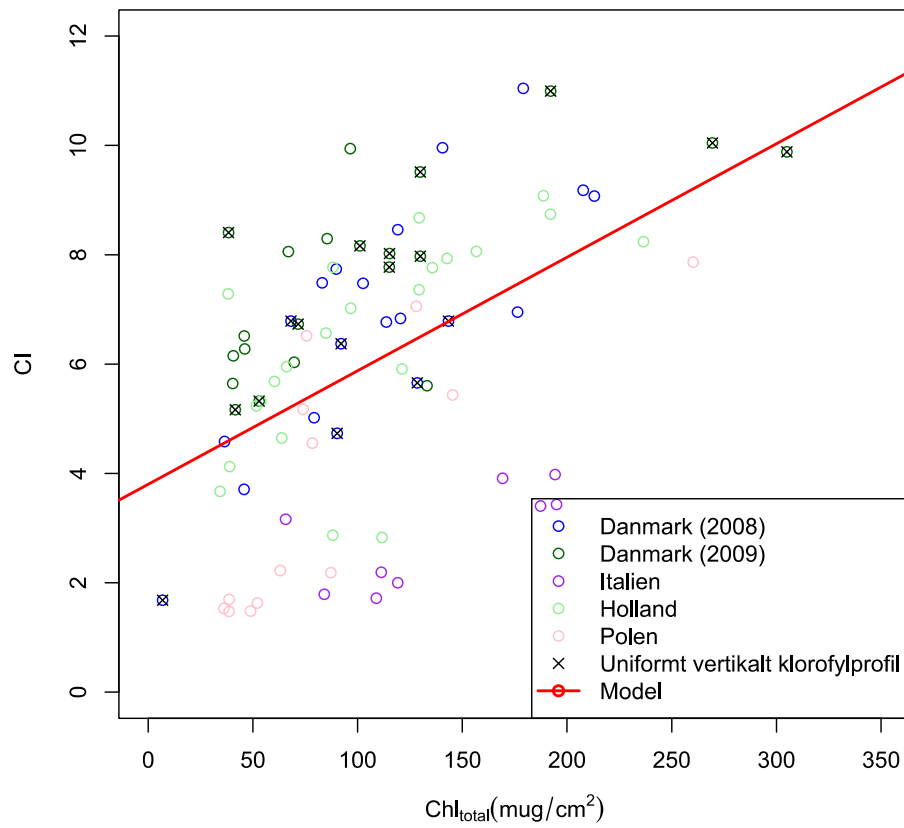
5.5 Følsomhedsanalyse af PROSAIL

I dette afsnit præsenteres resultaterne af følsomhedsanalysen af den fysisk-baserede PROSAIL-model. Først skal inputparametrene fastsættes.

For at opsætte modellen skal inputparametrene fastsættes og en oversigt



Figur 5.7: Figuren viser et plot af klorofylindekset (CI) i forhold til LAI for alle landskaber samt et fraktildiagram og et residualplot.



Figur 5.8: Figuren viser et plot af klorofyllindekset (CI) i forhold til $\text{Chl}_{\text{total}}$ for alle landskaber samt et fraktildiagram og et residualplot.

Symbol	Navn	Enhed	
Chl_{a+b}	Klorofylindhold	$\mu g/cm^2$	Målt
C_{ar}	Indhold af carotenoider	$\mu g/cm^2$	
C_{brown}	Indhold af brunt pigment	–	
C_w	Vandindhold	cm	
C_m	Tørstofindhold	g/cm^2	Målt
N	Mesofylstrukturparameter	–	
LAI	Bladarealindeks	m^2/m^2	Målt
h_{spot}	Hotspot	–	
θ_s	Solens zenitvinkel	$^\circ$	Kendt
θ_o	Observatørens zenitvinkel	$^\circ$	Kendt
ϕ	Azimutvinkel	$^\circ$	Kendt
LIDF	Fordelingsfunktion for bladvinkler		

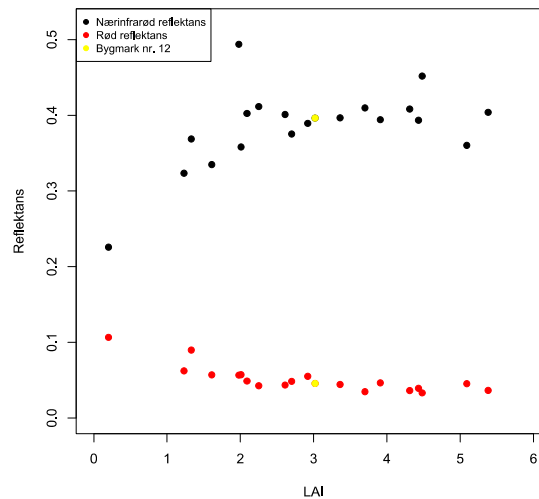
Tabel 5.7: Tabellen viser anvendte parametre i PROSAIL.

ses i 5.7.

Følsomhedsanalysen udføres for bygmark nr. 12 i Danmark (2008). Vinterbygmarken er kendetegnet ved et uniformt klorofylprofil med middelbladklorofyl på $47.5 \mu g/cm^2$, LAI på 3 og en højde på 0.95 m. Den målte overfladereflektans fra bygmarken er repræsentativ for de øvrige marker i landskab, hvilket ses på figur 5.9 på side 94. Figuren viser rød og nærinfrarød reflektans for alle marker i Danmark (2008). Det ses også på figuren, at én mark (nr. 9) har en høj nærinfrarød reflektans på omkring 50%, som er væsentlig højere end de øvrige marker. Denne pågældende mark er en bygmark, og der er ingen umiddelbar forklaring på den høje nærinfrarøde reflektans ud fra særlig højt LAI eller placering i forhold til arealdække eller jordbund.

Markerne er fordelt på to forskellige jordbundstyper som det fremgår af figur 3.3 på side 67, men det har ikke umiddelbart indflydelse på rød og nærinfrarød reflektans i forhold til LAI. Jordbundstyperne er henholdsvis sandblandet lerjord og lerblandet sandjord.

For PROSPECT er klorofylindholdet målt i Boegh et al. [2013] som beskrevet tidligere. I Boegh et al. [2013] er det gennemsnitlige tørstofindhold beregnet til $55 g/m^2$ undtagen for den skotske lokalitet, der har et lavere gennemsnitligt tørstofindhold, hvilket skyldes at der ikke er anvendt kunstgødning. Data fra Skotland anvendes som nævnt ikke her, hvorfor gennemsnittet kan anvendes. Omregnet til enheden g/cm^2 anvendes værdien $0.0055 g/cm^2$ i opsætningen af modellen. Ifølge Feret et al. [2008] er der et direkte proportionalt forhold mellem indholdet af carotenoider og klorofyl a og b i bladet. I samme artikel præsenteres datasæt for en lang række planter, der giver anledning til at bruge forhold 1:4, dvs., at indholdet af carotenoider udgør 25% af bladklorofylet. Indholdet af brunt pigment sættes til 0, dvs. det antages, at bladene er friske og ikke aldrende. Vandindholdet sættes på 0.1 cm, eftersom indflydelsen af



Figur 5.9: Figuren viser rød og nærinfrarød reflektans i forhold til bladareaindeks (LAI) fra markerne i Danmark (2008). Den udvalgte bygmark nr. 12 er markeret med gult.

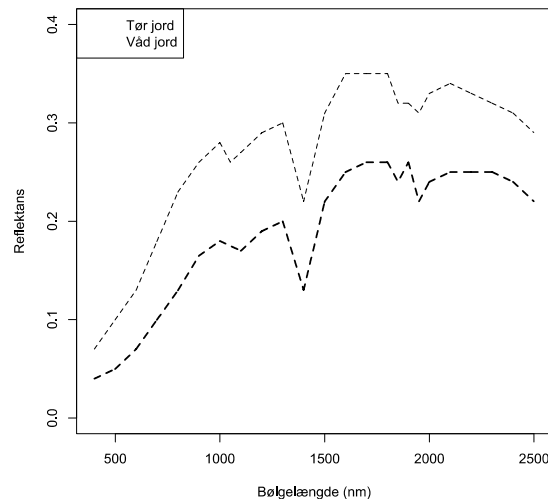
vand primært er til stede i melleminfrarøde bølgelængder.

For SAIL er LAI målt, og vinkler for sol og satellit er kendte. Hotspotparameteren fastsættes som $\frac{0.5}{LAI}$ jf. afsnit 2.2.3 på side 62, hvor forholdet er egnet til afgrøder, der har en proportional sammenhæng mellem højde og LAI. I afsnit 5.1 på side 78 ses, at hvede udviser en tilnærmelsesvis lineær sammenhæng mellem højde og LAI, mens byg og majs udviser en tilnærmelsesvis proportional sammenhæng. Det antages, at hotspotparameteren kan anvendes til alle tre afgrøder.

Bladvinkelfordelingen skal også fastlægges. Middelbladvinklen for den udvalgte bygmark er ca. 60° svarende til de øvrige byg- og hvedemarker. Det fremgår af estimererne for bladvinklen målt med LAI-2000. Derfor er den ellipsoide bladvinkelfordeling med middelbladvinkel 60° et bud på en beskrivelse af bladvinkelfordelingen. En anden mulighed er den sfæriske bladvinkelfordeling, der i Verhoefs algoritme approksimerer fordelingen med følgende tætheds- og fordelingsfunktion:

$$f(\theta_l) = \sin(\theta_l) \text{ og } F(\theta_l) = 1 - \cos(\theta_l)$$

For denne fordeling gælder det, at halvdelen af bladene vil have en bladvinkel på under 60° og halvdelen over. Ifølge implementeringen af den ellipsoide bladvinkelfordeling bør de to bladvinkelfordelinger være næsten sammenfaldende, hvis middelbladvinklen i den ellipsoide bladvinkelfordeling sættes til



Figur 5.10: Figuren viser de tilnærmede kurver for jordbundsreflektans anvendt i simuleringerne. Kurverne er approksimeret som stykkevis lineære funktioner fra punkter aflæst på figur 5.1 i Broge [2003], der viser jordbundsreflektansen for samme område målt i 1997. Middelværdien for grøn reflektans (500-590 nm) er 0.06 for våd jord og 0.11 for tør jord. For rød reflektans (610-690 nm) er middelværdien 0.08 for våd jord og 0.15 for tør jord. Den nærinfrarøde reflektans (790-890 nm) har middelværdien 0.14 for våd jord og 0.24 for tør jord.

58.4°. Teoretisk set er den ellipsoidale fordeling en generalisering af den sfæriske fordeling, men pga. forskel i implementering bliver de ikke fuldstændigt sammenfaldende.

Et andet input er jordbundens reflektans. I Broge [2003] blev reflektansen fra jordbunden for en mark i samme område undersøgt i 1997. Der blev foretaget målinger af reflektansen fra en bar plet på marken på forskellige tidspunkter i vækstsæsonen. På figur 5.10 på side 95 er denne målte reflektans approksimeret ud fra stykkevis lineære funktioner. Det ses, at den målte jordbundsreflektans fra Broge [2003] er lavere end reflektansen i standardopsætningen i Jacquemond et al. [2013], som ses på figur 2.16 på side 62. Forskellen mellem tør og våd jord er desuden mindre.

I forbindelse med følsomhedsanalysen viste det sig, at der var en fejl i Matlab-koden i Jacquemond et al. [2013]. Fejlen vedrører frekvensberegningen af de to forskellige bladvinkelfordelinger. For Verhoefs algoritme beregnes frekvenser for bladvinkelfordelingsfunktionen ved 0°, 10°, ..., 80°, 82°, 84°, ..., 88°, 90°, dvs. i alt 13 intervaller, hvor frekvenser beregnes for intervallerne midtpunkt. Årsagen til, at intervallet mellem 80° og 90° er finere inddelt skyldes, at sen-

Fordeling	Grøn reflektans	Rød reflektans	Nærinfrarød reflektans
Tør jord	0.02547	0.01789	0.3906
Våd jord	0.02327	0.01527	0.3617
Relativ forskel	9.45%	17.18%	7.99%

Tabel 5.8: Simuleret reflektans ved henholdsvis tør og våd jordbund med sfærisk bladvinkelfordeling. LAI er 3.02, bladklorofyl er $47.47 \mu\text{g}/\text{cm}^2$, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

soren typisk er placeret tæt på nadir, hvorfor bladvinkler tæt på zenit har stor indflydelse på den observerede reflektans [Verhoef, 1997]. For den ellipsoide bladvinkelfordeling beregnes frekvensen for intervaller inddelt ved $0^\circ, 5^\circ, 10^\circ, \dots, 80^\circ, 90^\circ$, dvs. i alt 18 intervaller. Her er altså ikke samme fininddeling af intervallet mellem 80° og 90° .

MATLAB-koden var skrevet så SAIL regnede videre med 13 intervalmidtpunkter. For bladvinkelfordelinger beregnet med Verhoefs algoritme betød det, at alle punkter blev regnet med, mens det for den ellipsoide bladvinkelfordeling betød, at kun de første 13 intervalmidtpunkter blev regnet med, og således blev det fininddelte interval fra 80° til 90° "optaget" af frekvenser beregnet for 40° til 65° . I Fortran-koden er den ellipsoide bladvinkelfordeling implementeret således, at der beregnes frekvenser for de samme intervaller som for Verhoefs algoritme [Jacquemon et al., 2013], og således tages der højde for intervallet mellem 80° og 90° .

I MATLAB blev koden justeret på anden vis, så SAIL for den ellipsoide bladvinkelfordeling regner videre med alle 18 intervalmidtpunkter uden at tage højde for intervallet mellem 80° og 90° . Således er der en usikkerhed tilføjet modellen.

5.5.1 Følsomhed over for tør og våd jordbund

Der er stor forskel på reflektansen fra jordbunden afhængig af, om den er tør eller våd jf. figur 5.10 på side 95. I tabel 5.8 på side 96 ses forskellen i simuleret reflektans for våd og tør jord ved en sfærisk bladvinkelfordeling. Det er især den røde reflektans, der øges ved tør jord i forhold til våd jord jf. figur 5.10 på side 95. Ændringen er på knap 20%.

5.5.2 Resultat af følsomhedsanalysen

Resultatet af en $\pm 10\%$ ændring i de øvrige inputparametre ses i tabel 5.9 på side 98 for en sfærisk bladvinkelfordeling og i tabel 5.10 på side 98 for en ellipsoide bladvinkelfordeling med middelbladvinkel på 60° . Generelt ses det, at simuleringerne af reflektans af synligt lys og nærinfrarød stråling er stabile over for ændringer i inputparametrene. De relative ændringer er næsten ens for de to forskellige bladvinkelfordelinger, hvilket stemmer godt

overens med den tætte sammenhæng mellem den sfæriske og den ellipsoide bladvinkelfordeling med middelbladvinkel tæt på 60° jf. afsnit 2.2.3 på side 55.

PROSAIL er følsom over for LAI. Et fald på 10% giver en øget grøn reflektans på 5.0% og øget rød reflektans på 10.8%. En tilsvarende stigning i LAI giver en fald i grøn og rød reflektans på henholdsvis 3.5% og 7.7%. I forhold til nærinfrarød reflektans mindskes reflektansen med faldende LAI og omvendt. Ændringen er på under 3.5%. Der er ikke en særlig grad af følsomhed over for hotspotparameteren, hvor en ændring på $\pm 10\%$ giver en ændring på mindre end $\pm 1\%$ i den simulerede reflektans.

Modellen er også følsom over for klorofylindholdet i bladene. Et 10% fald i klorofylindholdet øger den grønne reflektans med 8.1% og den røde reflektans med ca. 4.6%. En tilsvarende stigning giver et fald i grøn og rød reflektans på henholdsvis 6.6% og 3.4%. Der er ingen påvirkning af den nærinfrarøde reflektans ved ændring af klorofylindholdet. Indholdet af carotenoider påvirker kun den grønne reflektans, hvor en ændring i indholdet på $\pm 10\%$ ændrer reflektansen med under $\mp 2\%$. Indholdet af carotenoider påvirker ikke den røde og nærinfrarøde reflektans. Resultatet stemmer godt overens med figur 2.7a på side 48, der viser PROSPECTs implementering af de specifikke absorptionskoefficienter for klorofyl og carotenoider. Klorofyls absorption er høj ved både grønt og rødt lys, mens carotenoider kun absorberer grønt lys.

En $\pm 10\%$ ændring i mesofylstrukturparameteren N påvirker simuleringen af grøn og nærinfrarød reflektans med henholdsvis ca. $\pm 3.5\%$ og ca. $\pm 1.5\%$, mens påvirkningen af den røde reflektans er mindre end $\pm 1\%$. Der er ingen følsomhed over for vandindholdet i reflektans af synligt lys og en meget lille påvirkning af nærinfrarød reflektans. Ændring i tørstofindholdet giver kun anledning til meget små ændringer i reflektans af synligt lys, mens ændringen er under $\mp 2\%$ den nærinfrarøde del af spektret. Det ses også på figur 2.7a på side 48, at tørstof og vand i PROSPECT kun absorberer nærinfrarød stråling og ikke synligt lys.

Reflektansen fra jordbunden har en betydning for den simulerede reflektans. Størst indflydelse har ændring i den tørre jordbund, hvis reflektans er højere end den våde jordbund jf. figur 5.10 på side 95.

En ændring på $\pm 10\%$ påvirker den grønne reflektans med henholdsvis $\pm 1.80\%$ ved tør jord og $\pm 1.02\%$ ved våd jord, mens den røde reflektans påvirkes med henholdsvis $\pm 3.24\%$ og $\pm 2.07\%$. Den nærinfrarøde reflektans påvirkes med henholdsvis ca. $\pm 2\%$ og $\pm 1.1\%$.

I den ellipsoide bladvinkelfordeling kan middelbladvinklen også varieres. Effekten på den simulerede reflektans ses i tabel 5.10 på side 98. Ved et fald i middelbladvinkel på 10% bliver den grønne reflektans øget med 4%, den røde reflektans påvirkes ikke i høj grad og den nærinfrarøde reflektans øges med knap 8%. En stigning i middelbladvinkel på 10% giver et fald på 3% i grøn reflektans, en stigning på knap 4% i rød reflektans og et fald på

+/- 10% ændring	Grøn reflektans		Rød reflektans		NIR reflektans	
	-10%	+10%	-10%	+10%	-10%	+10%
LAI	5.02%	-3.53%	10.80%	-7.70%	-3.48%	3.27%
Hotspot	-0.70%	0.69%	-0.98%	0.97%	-0.30%	0.29%
Bladklorofyl	8.06%	-6.57%	4.62%	-3.37%	0.00%	0.00%
Carotenoider	1.72%	-1.53%	0.00%	0.00%	0.00%	0.00%
Mesofylstruktur	-2.81%	3.44%	-0.33%	0.84%	-1.56%	1.41%
Vandindhold	0.00%	0.00%	0.00%	0.00%	0.05%	-0.05%
Tørstof	0.15%	-0.14%	0.03%	-0.03%	1.64%	-1.58%
Tør jordbund	-1.80%	1.80%	-3.24%	3.24%	-1.90%	1.95%
Våd jordbund	-1.02%	1.02%	-2.07%	2.07%	-1.10%	1.12%

Tabel 5.9: Tabellen viser den procentvise ændring i middelværdier for simuleret grøn, rød og nærinfrarød reflektans med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkelfordelingen er sfærisk, jordbunden tør, LAI varieres omkring 3, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

9% i nærinfrarød reflektans. Dvs., at PROSAIL er mindre følsom over for ændringer i middelbladvinkel end over for ændringer i LAI og bladklorofyl i det synlige lysspektrum. Til gengæld er det variation i middelbladvinklen, der giver udslag i den største relative ændring af den nærinfrarøde reflektans.

+/- 10% ændring	Grøn reflektans		Rød reflektans		NIR reflektans	
	-10%	+10%	-10%	+10%	-10%	+10%
LAI	5,28%	-3,73%	11,27%	-8,07%	-3,51%	3,30%
Hotspot	-0,69%	0,67%	-0,96%	0,95%	-0,29%	0,28%
Middelbladvinkel	4,11%	-3,04%	0,33%	3,58%	7,68%	-9,42%
Bladklorofyl	8,01%	-6,53%	4,61%	-3,36%	0,00%	0,00%
Carotenoider	1,71%	-1,53%	0,00%	0,00%	0,00%	0,00%
Mesofylstruktur	-2,63%	3,27%	-0,26%	0,76%	-1,51%	1,37%
Vandindhold	0,00%	0,00%	0,00%	0,00%	0,05%	-0,05%
Tørstofindhold	0,14%	-0,14%	0,03%	-0,03%	1,64%	-1,59%

Tabel 5.10: Tabellen viser den procentvise ændring i middelværdier for simuleret grøn, rød og nærinfrarød reflektans med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkelfordelingen er ellipsoidal med middelbladvinkel på 60° , LAI varieres omkring 3, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

PROSAIL's følsomhed er også undersøgt ved højt og lavt LAI. Ved LAI på 6 ses samme følsomhed som ved LAI på 3. Resultatet vises derfor ikke. Ved LAI på 1 ændres følsomheden over for inputparametrene dog, hvilket ses i tabel 5.11 på side 99. Ved lavt LAI er følsomheden over for ændringer i LAI større i

den grønne reflektans og lavere i den røde og nærinfrarøde reflektans. Følsomheden over for bladklorofyl, carotenoider og mesofylstrukturen er mindre ved lavt LAI. Følsomheden over for middelbladvinklen ændres ved lavt LAI. Et fald på 10% i middelbladvinklen giver anledning til et fald på 3% i den grønne reflektans, et fald på 6% i den røde reflektans og en stigning på knap 6% i den nærinfrarøde reflektans. En stigning på 10% i middelbladvinklen giver derimod en stigning på 4% i den grønne reflektans, en stigning på 8% i den røde reflektans og et fald på 6% i den nærinfrarøde reflektans.

Når LAI er lavt, vil en mere horisontal bladvinkel skjule jordbunden, og synligt lys vil i højere grad absorberes af bladet i stedet for at blive reflekteret af den blottede jordbund. Omvendt til en mere vertikal bladvinkel blotte jordbunden yderligere og øge reflektansen. Ved højere LAI vil en stejlere bladvinkel blotte flere blade og derfor øge absorptionen af synligt lys.

+/- 10% ændring	Grøn reflektans		Rød reflektans		NIR reflektans	
	-10%	+10%	-10%	+10%	-10%	+10%
LAI	6,71%	-6,04%	9,05%	-8,11%	-1,74%	1,76%
Hotspot	-0,35%	0,35%	-0,43%	0,43%	-0,24%	0,23%
Middelbladvinkel	-3,06%	4,26%	-6,30%	8,25%	5,70%	-6,16%
Bladklorofyl	2,96%	-2,47%	1,38%	-1,04%	0,00%	0,00%
Carotenoider	0,65%	-0,59%	0,00%	0,00%	0,00%	0,00%
Mesofylstruktur	-0,45%	0,71%	0,17%	-0,03%	-1,11%	1,01%
Vandindhold	0,00%	0,00%	0,00%	0,00%	0,02%	-0,02%
Tørstofindhold	0,05%	-0,05%	0,01%	-0,01%	0,65%	-0,64%

Tabel 5.11: Tabellen viser den procentvise ændring i middelværdier for simuleret grøn, rød og nærinfrarød reflektans med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkelfordelingen er ellipsoidal med middelbladvinkel på 60° , LAI varieres omkring 1, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

5.6 Vegetationsindeksernes følsomhed

De spektrale vegetationsindekseres følsomhed over for ændringer i LAI, middelbladvinkel og bladenes indhold af klorofyl og andre pigmenter er undersøgt med PROSAIL.

Følsomheden er undersøgt for LAI på 3 samt for lavt og højt LAI. I tabel 5.12 på side 102 ses de relative ændringer i vegetationsindekserne ved variation på $\pm 10\%$ i parametrene for LAI på 3. I tabel 5.13 på side 103 og tabel 5.14 på side 103 er de relative ændringer beregnet for LAI på henholdsvis 1 og 6. Generelt er vegetationsindekserne især følsomme over for ændringer i LAI, middelbladvinkel og til en vis grad bladklorofyl. Følsomheden over for LAI er særlig stor ved SR og CI, hvor variation med $\pm 10\%$ i LAI giver ændringer på

$\pm 10\%$ eller højere for LAI under 3. Følsomheden aftager ved voksende LAI, og ændringer på $\pm 10\%$ i LAI giver således en forskel på omkring 2–4 % hos SR og CI, når LAI er 6. NDVI ændres med ca. 4.5%, når LAI på 1 ændres med 10%. For højere LAI er ændringen i NDVI markant mindre. Ved en ændring på 10% i LAI på 6 ændres NDVI med blot 0.1–0.2%. Der sker altså en markant udslukning af følsomheden over for voksende LAI. Det samme billede ses hos GDNVI, der er lidt mindre følsom over for LAI end NDVI. EVI2 har omtrentligt samme følsomhed som NDVI ved lavt LAI. Der sker dog ikke samme mætning, selv om EVI2's følsomhed over for LAI mindskes ved stigende LAI.

Alle vegetationsindekserne er mere følsomme over for ændringer i middelbladvinklen end over for LAI. Det skyldes, at middelbladvinklen påvirker den nærinfrarøde reflektans betydelig jf. tabel 5.10 på side 98. Der er størst følsomhed over for middelbladvinklen hos SR, CI og EVI2. NDVI og GDNVI er primært følsomme over for ændringer i middelbladvinklen ved lavt LAI.

Nogle af vegetationsindekserne er følsomme over for bladklorofyl. Det drejer sig især om CI, hvor følsomheden over for bladklorofyl vokser ved stigende LAI. Ved LAI over 3 giver en ændring på $\pm 10\%$ en ændring på $\pm 8-9\%$ i CI. SR udviser også følsomhed for bladklorofyl, dog i lavere grad. NDVI og EVI2 er meget lidt følsomme over for bladklorofyl, da en ændring på $\pm 10\%$ ændrer mindre end $\pm 1\%$ på indekserne, og følsomheden aftager yderligere ved voksende LAI. GNDVI ændres med ca. $\pm 1\%$, når bladklorofylet varieres med $\pm 10\%$. Indholdet af carotenoider påvirker kun den grønne reflektans, som indgår i GNDVI og CI. Følsomheden over for carotenoider er dog lav hos GNDVI. CI ændres med $\pm 2\%$, når indholdet af carotenoider ændres med $\pm 10\%$ ved LAI over 3.

Ændring i jordbundens reflektans påvirker vegetationsindekserne på forskellig vis. Ved LAI på 1 øges alle vegetationsindekserne på nær EVI2 ved lavere reflektans fra jordbunden. For SR og NDVI skyldes stigningen ved et fald i jordbundsreflektansen, at ændringen i den røde reflektans påvirker vegetationsindekset mere end ændringen i den nærinfrarøde reflektans. For GNDVI og CI gælder det samme blot med den grønne reflektans i stedet for den røde. EVI2 er derimod tilsyneladende mere følsom over for den ændrede nærinfrarøde reflektans, hvorfor EVI2 falder ved et fald i jordbundsreflektansen. Ved LAI på 3 øges SR og NDVI også ved et fald i jordbundsreflektansen. Det skyldes, at SR og NDVI er følsomme over for den røde reflektans. GDNVI, EVI2 og CI falder til gengæld ved lavere jordbundsreflektans. For GNDVI og CIs vedkommende skyldes det, at den grønne reflektans ikke falder lige så meget som den røde reflektans jf. figur 5.10 på side 95. For EVI2s vedkommende skyldes det igen, at følsomheden over for ændringer i nærinfrarød reflektans er større end følsomheden over for ændringer i rød reflektans. Ved LAI på 6 falder alle vegetationsindekser ved lavere jordbundsreflektans, men generelt påvirker jordbundsreflektansen kun indekserne meget lidt ved LAI på 6.

Variation af de øvrige inputparametre har en forholdsvis lille indflydelse

på vegetationsindekserne.

+/- 10% ændring	SR		NDVI		GNDVI		EVI2		CI	
	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%
LAI	-13,28%	12,36%	-1,41%	1,03%	-1,19%	0,90%	-3,43%	3,02%	-8,93%	7,82%
Hotspot	0,68%	-0,66%	0,06%	-0,06%	0,05%	-0,05%	-0,17%	0,16%	0,42%	-0,41%
Middelbladvinkel	7,33%	-12,55%	0,64%	-1,33%	0,44%	-0,92%	5,80%	-7,70%	3,67%	-7,04%
Bladklorofyl	-4,41%	3,48%	-0,43%	0,31%	-1,05%	0,86%	-0,28%	0,21%	-7,94%	7,48%
Carotenoider	0,00%	0,00%	0,00%	0,00%	-0,22%	0,20%	0,00%	0,00%	-1,80%	1,66%
Mesofylstruktur	0,60%	0,60%	0,06%	0,06%	-0,25%	-0,25%	1,01%	1,01%	-1,96%	-1,96%
Vandindhold	0,05%	-0,05%	0,00%	0,00%	0,01%	-0,01%	0,04%	-0,04%	0,05%	-0,05%
Tørstofindhold	1,61%	-1,55%	0,15%	-0,15%	0,19%	-0,19%	1,26%	-1,23%	1,60%	-1,55%
Jordbundsreflektans	1,53%	-1,38%	0,14%	-0,13%	-0,01%	0,01%	-1,32%	1,35%	-0,05%	0,10%

Tabel 5.12: Tabellen viser den procentvise ændring i de spektrale vegetationsindekser simuleret med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkel fordelingen er ellipsoid med middelbladvinkel på 60° , LAI varierer omkring 3, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

+/- 10% ændring	SR		NDVI		GNDVI		EVI2		CI	
	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%
LAI	-9,89%	10,74%	-4,68%	4,29%	-3,26%	2,98%	-4,60%	4,34%	-9,72%	10,18%
Middelbladvinkel	12,81%	-13,31%	5,04%	-6,50%	3,23%	-4,20%	7,85%	-9,09%	11,10%	-12,27%
Bladklorofyl	-1,36%	1,05%	-0,60%	0,45%	-1,13%	0,95%	-0,42%	0,32%	-3,53%	3,11%
Carotenoider	0,00%	0,00%	0,00%	0,00%	-0,25%	0,23%	0,00%	0,00%	-0,80%	0,73%
Jordbundsreflektans	2,97%	-2,41%	1,26%	-1,07%	0,62%	-0,52%	-3,69%	3,66%	2,00%	-1,63%

Tabel 5.13: Tabellen viser den procentvise ændring i de spektrale vegetationsindekser simuleret med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkelfordelingen er ellipsoid med middellbladvinkel på 60° , LAI varierer omkring 1, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

+/- 10% ændring	SR		NDVI		GNDVI		EVI2		CI	
	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%	-10%	+10%
LAI	-3,24%	2,08%	-0,19%	0,11%	-0,24%	0,16%	-1,88%	1,50%	-2,60%	1,82%
Middelbladvinkel	-6,23%	7,41%	-0,37%	0,39%	-0,35%	0,34%	3,84%	-5,43%	-3,69%	3,83%
Bladklorofyl	-5,80%	4,67%	-0,34%	0,25%	-0,89%	0,73%	-0,23%	0,17%	-8,98%	8,64%
Carotenoider	0,00%	0,00%	0,00%	0,00%	-0,19%	0,17%	0,00%	0,00%	-2,04%	1,89%
Jordbundsreflektans	-0,23%	0,25%	-0,01%	0,01%	-0,03%	0,03%	-0,30%	0,31%	-0,36%	0,38%

Tabel 5.14: Tabellen viser den procentvise ændring i de spektrale vegetationsindekser simuleret med PROSAIL, når parameteren justeres med $\pm 10\%$. Bladvinkelfordelingen er ellipsoid med middellbladvinkel på 60° , LAI varierer omkring 6, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$.

	Målt	Simuleret	Afvigelse
Grøn reflektans	0.0509	0.02626	-48.41%
Rød reflektans	0.0457	0.01793	-60.76%
NIR reflektans	0.3964	0.41335	4.28%

Tabel 5.15: Tabellen viser afvigelsen mellem målt og simuleret reflektans for bygmark nr. 12 i Danmark (2008) med LAI = 3.02 og klorofylindhold på 47.5 $\mu\text{g}/\text{cm}^2$. Middelbladvinklen er 57°.

5.7 Simulering af reflektans med PROSAIL

I dette afsnit præsenteres resultaterne af simulering af reflektans med PROSAIL på en række marker i Danmark (2008). Bladvinkelfordelingen er ellipsoidal med middelbladvinkel på 57°, jordbunden tør, $\theta_s = 34.90^\circ$, $\theta_o = 10.04^\circ$ og $\psi = 148.21^\circ$. Reflektansen fra markerne i Danmark (2009) simuleres også med $\theta_s = 46.67^\circ$, $\theta_o = 12.85^\circ$ og $\psi = 148.14^\circ$.

Først simuleres reflektansen for bygmark nr. 12, der blev anvendt i følsomhedsanalysen.

Endelig sammenlignes med afvigelsen mellem målt og simuleret reflektans for marker uden uniformt klorofylprofil.

5.7.1 Målt og simuleret reflektans i forhold til profiler

For alle markerne fra Danmark (2008) er reflektansspektret simuleret med PROSAIL, og middelværdierne for grøn, rød og nærinfrarød reflektans er beregnet og sammenlignet med de målte reflektanser fra SPOT. Resultatet ses i tabel E.1 på side 145 i appendiks E fordelt efter det vertikale klorofylprofil på den enkelte mark. Afgrøder, højde og LAI fremgår også af tabellen.

Modellen er god til at simulere reflektansen ved meget lavt LAI. Majs-marken med LAI på 0.2 simuleres med en afvigelse på mindre end 20 % på reflektanserne og mindre end 1 % afvigelse for fire ud af fem vegetationsindekser. For alle marker er der en god overensstemmelse mellem målt og simuleret nærinfrarød reflektans. Den eneste mark med en relativ afvigelse over 30 % er mark nr. 9, som har en usædvanlig høj nærinfrarød reflektans jf. figur 5.9 på side 94. Den røde reflektans simuleres til gengæld for lavt for alle marker med LAI over 1.3. Estimerne er mere end 50 % for lave ved LAI over 2.7. Den grønne reflektans simuleres også for lavt, men der er ikke et klart mønster i forhold til LAI. Til gengæld simuleres den grønne reflektans bedst for marker med lavest bladklorofyl (bladklorofylindholdet er ikke vist i tabellen).

SR påvirkes i høj grad af de for lave estimer for rød og estimeres alt for højt (mere end 100 %) ved LAI over 2.7. NDVI påvirkes ikke lige så meget pga. robustheden over for den røde reflektans og bliver derfor aldrig mere end lidt over 20 % for høj. GNDVI estimeres maksimalt 20 % for højt og i de fleste

tilfælde mindre end 10 % fra det målte. EVI2 er også robust over for fejlen på det røde reflektansestimater. I to tilfælde er der mere end 20 % forskel på målt og simuleret EVI2. Der er ikke et klart mønster i de relative afvigelser for CI.

Der er ikke en tydelig indflydelse fra det vertikale klorofylprofil. I det følgende afsnit undersøges de absolutte afvigelser.

5.7.2 Absolutte afvigelser

I dette afsnit undersøges de absolutte afvigelser mellem målt og simuleret reflektans.

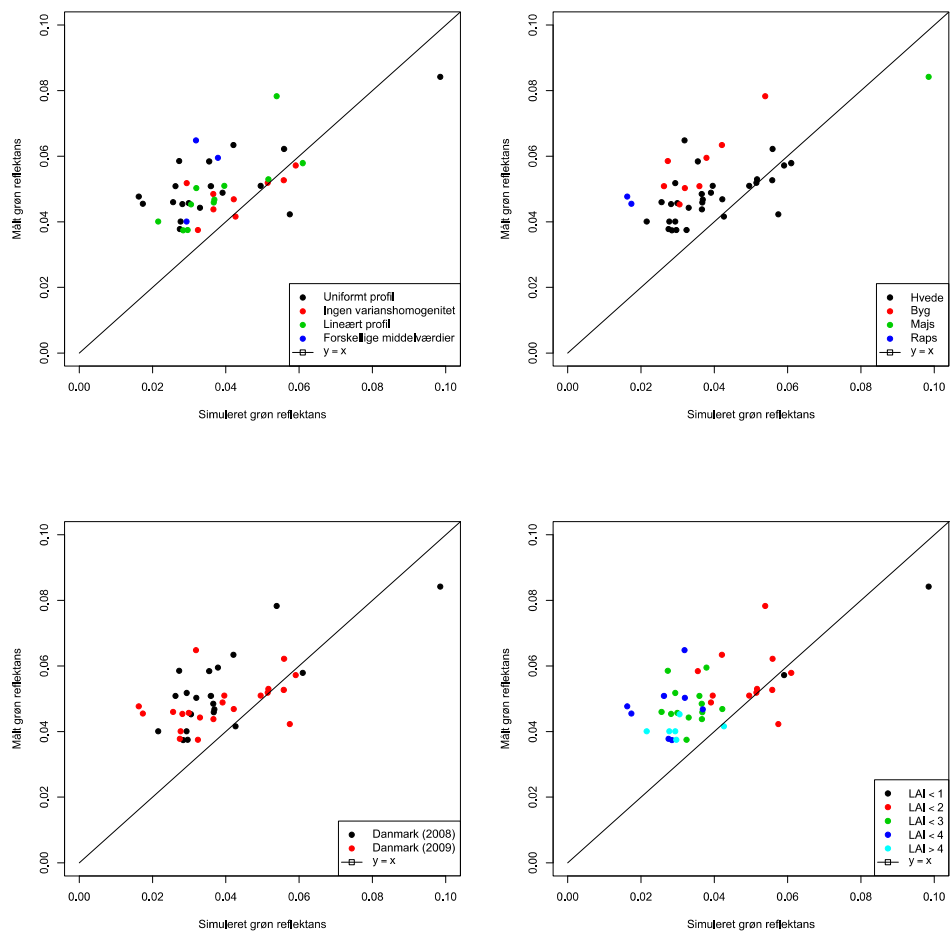
Ensidet variansanalyse anvendes til at undersøge, om der er systematisk forskelle i de absolutte afvigelser mellem de forskellige vertikale klorofylprofiler for markerne. I den forbindelse undersøges også landskab (Danmark (2008) eller Danmark (2009)), afgrøde (hvede eller byg) samt LAI (under 2, mellem 2 og 3, mellem 3 og 4, over 4).

Grøn reflektans Der er ikke forskel i de absolutte afvigelser mellem de forskellige vertikale klorofylprofiler, og der er heller ikke forskel på de absolutte afvigelser mellem de to landskaber. Der er forskel på afgrøderne. Således er de absolutte afvigelser større for byg (middelværdi på 0.021) end for hvede (middelværdi 0.009). Der er også forskel på størrelsen af LAI. Den absolutte afvigelse er størst ved LAI mellem 3 og 4 (middelværdi 0.0206). Den laveste absolutte afvigelse er for LAI under 2 (middelværdi 0.004).

På figur 5.11 på side 106 ses to figurer, der begge viser sammenhængen mellem simuleret og målt grøn reflektans. Til venstre er det vertikale klorofylprofil markeret, og der ses som ventet ud fra variansanalysen af afvigelserne ikke et særligt mønster i forhold til, at reflektansen simuleres bedre for nogle typer af vertikalt profil end for andre. Til højre er plantetype markeret, og her ses det, at alle bygmarkerne har en simuleret grøn reflektans, der er lavere end den målte. Hvedemarkerne er bedre fordelt, men der er flest med for lavt simuleret grøn reflektans.

Rød reflektans For den røde reflektans er der forskel på de absolutte afvigelser mellem marker med forskelligt vertikalt klorofylprofil. Den største afvigelse er hos marker med forskellige middelværdier ned gennem profilet (middelværdi 0.031). Derefter følger marker med lineært klorofylprofil (middelværdi 0.020) og uniformt klorofylprofil (middelværdi 0.017). Den laveste absolutte afvigelse findes hos markerne uden varianshomogenitet (middelværdi 0.008).

Variansanalysen viser, at forskellen mellem de to landskabers absolutte afvigelse er tæt på grænsen for signifikansniveauet ($p = 0.049$). Det viser sig, at den gennemsnitlige afvigelse for Danmark (2008) er højere end den gennemsnitlige afvigelse for Danmark (2009). For den røde reflektans er der ikke



Figur 5.11: Figuren viser sammenhængen mellem simuleret og målt rød refleksans for alle danske marker. Øverst til venstre er det vertikale klorofylprofil fremhævet, mens figuren øverst til højre viser sammenhængen i forhold til plantetype. Nederst til venstre er landskabet fremhævet og nederst til højre LAI.

varianshomogenitet for de absolutte afvigelser mellem forskellige afgrødetyper eller mellem forskellige niveauer af LAI, hvorfor indflydelsen herfra ikke kan testes ved ensidet variansanalyse.

På figur 5.12 på side 108 ses sammenhængen mellem målt og simuleret rød reflektans i forhold til det vertikale klorofylprofil, plantetype, landskab og LAI. På figureerne ser det ud til, at den røde reflektans simuleres bedst for hvedemarker, i landskabet Danmark (2009) og ved lavt LAI. Variansanalysen bekræfter, at afvigelsen er lavest for Danmark (2009). I forhold til de vertikale klorofylprofiler ses den laveste afvigelse ved marker uden varianshomogenitet i overensstemmelse med variansanalysen.

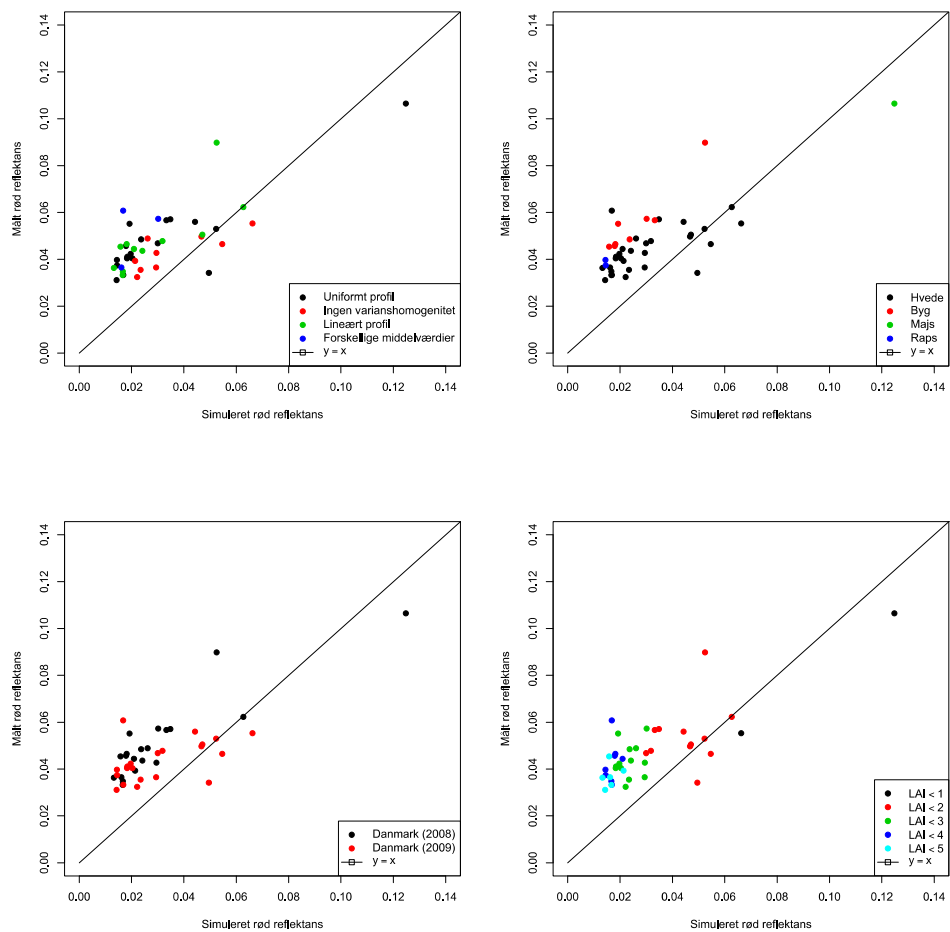
Nærinfrarød reflektans For den nærinfrarøde reflektans er der ikke forskel mellem de absolutte afvigelser for forskellige vertikale klorofylprofiler. Der er heller ikke forskel mellem hvede eller byg. Til gengæld er der forskel ved forskellige niveauer af LAI. Den største gennemsnitlige afvigelse findes ved LAI under 2 (0.051). Afvigelsen er mindre for LAI mellem 2 og 3 (0.034) og mellem 3 og 4 (0.017). For LAI større end 4 er den gennemsnitlige afvigelse -0.041. Dvs. her er den simulerede reflektans højere end den målte. Der er ikke varianshomogenitet mellem de to landskaber, hvorfor variansanalyse ikke kan anvendes til at vurdere forskelle mellem målt og simuleret reflektans.

På figur 5.13 på side 109 vises sammenhængen mellem målt og simuleret nærinfrarød reflektans i forhold til det vertikale klorofylprofil, plantetype, landskab og LAI. Generelt kan man se, at den nærinfrarøde reflektans simuleres både for højt og for lavt af PROSAIL, og at de simulerede nærinfrarøde reflektanser er mere spredte end de målte. Ligesom variansanalysen viste, ses der ikke systematisk forskel baseret på de vertikale klorofylprofiler eller plantetyper. Der ses heller ikke noget særligt mønster ved forskellige landskaber.

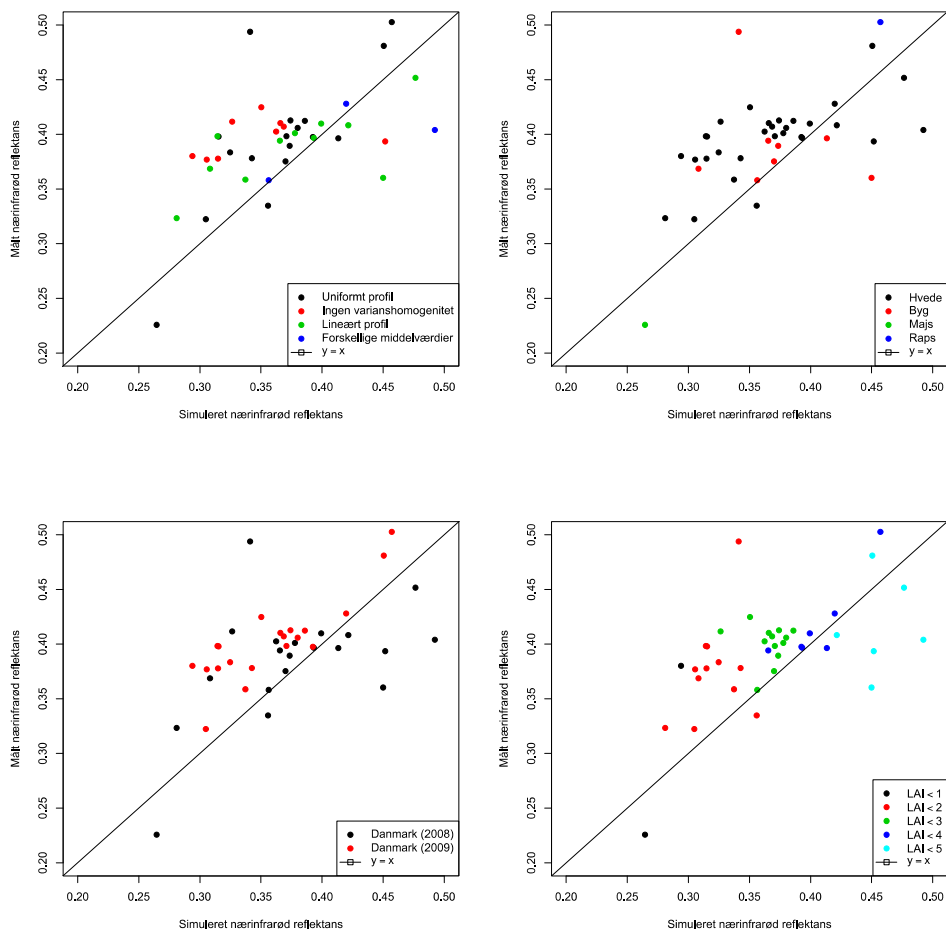
SR Der er forskel i de absolutte afvigelser for SR ved forskellige vertikale klorofylprofiler, og mønstret følger den røde reflektans, hvor den største gennemsnitlige afvigelse ses hos markerne med forskellige middelværdier i det vertikale klorofylprofil (-14.3), efterfulgt af markerne med lineært profil (-9.0) og uniformt profil (-7.9). Den mindste gennemsnitlige afvigelse ses hos marker uden varianshomogenitet (-2.3). Alle de gennemsnitlige afvigelser er negative, da SR generelt estimeres for højt af PROSAIL.

Der er også forskel på de absolutte afvigelser mellem forskellige niveauer af LAI. Den mindste gennemsnitlige afvigelse ses ved LAI <2 (-0.10). Ved LAI mellem 2 og 3 er den gennemsnitlige afvigelse -6.69, mellem 3 og 4 -14.24 og ved LAI >4 -17.13.

NDVI For NDVI viser variansanalysen, at der er forskel på de absolutte afvigelser mellem de forskellige vertikale klorofylprofiler, forskellige niveauer



Figur 5.12: Figuren viser sammenhængen mellem simuleret og målt rød reflektans for alle danske marker. Øverst til venstre er det vertikale klorofylprofil fremhævet, mens figuren øverst til højre viser sammenhængen i forhold til plantetype. Nederst til venstre er landskabet fremhævet og nederst til højre LAI.



Figur 5.13: Figuren viser sammenhængen mellem simuleret og målt nærinfrarød reflektans for alle danske marker. Øverst til venstre er det vertikale klorofylprofil fremhævet, mens figuren øverst til højre viser sammenhængen i forhold til plantetype. Nederst til venstre er landskabet fremhævet og nederst til højre LAI.

af LAI, forskellige landskaber samt forskellige afgrøder. De gennemsnitlige absolutte afvigelser ved de forskellige typer vertikalt klorofylprofil følger samme mønster som ved den røde reflektans og SR. Afvigelsen er størst ved marker med forskellige middelværdier i klorofylindholdet (-0.131), derefter marker med lineært vertikalt klorofylprofil (-0.071) og uniformt profil (-0.062), og den laveste gennemsnitlige afvigelse er for marker uden varianshomogenitet (-0.005).

For de forskellige niveauer af LAI ses det, at den gennemsnitlige afvigelse er lavest ved LAI <2 (0.003). Derefter følger LAI mellem 2 og 3 (-0.076). For LAI mellem 3 og 4 er den gennemsnitlige afvigelse størst (-0.102), mens den næststørste afvigelse ses ved LAI >4 (-0.097). Den gennemsnitlige afvigelse er større for byg end for hvede og for Danmark (2008) i forhold til Danmark (2009).

GNDVI GNDVI viser samme mønster som NDVI med hensyn til LAI, afgrøder og landskab. En undtagelse er for de vertikale klorofylprofiler, hvor den største afvigelse ses hos marker med forskellige middelværdier i klorofylprofilet efterfulgt af marker med uniformt klorofylprofil og lineært profil. Det er også stadig markerne uden varianshomogenitet, som har lavest gennemsnitlig afvigelse.

EVI2 For EVI2 er der forskel i de absolutte afvigelser for både vertikale klorofylprofiler, LAI og landskab. Der er ikke forskel i de absolutte afvigelser mellem plantetyper.

For de vertikale klorofylprofiler er den gennemsnitlige afvigelse størst hos marker med forskellige middelværdier i profilet (-0.078). Derefter følger marker uden varianshomogenitet i klorofylindholdet (0.042). Marker med lineært vertikalt klorofylprofil har en gennemsnitlig afvigelse på -0.018, mens marker med uniformt profil har en gennemsnitlig afvigelse på 0.003.

For LAI er den gennemsnitlige afvigelse mindst for marker med LAI mellem 2 og 3 (0.001). Afvigelsen for LAI mellem 3 og 4 er -0.027. For LAI <2 er afvigelsen 0.046, mens den største afvigelse er for LAI over 4 (-0.077).

For Danmark (2008) er afvigelsen -0.029, mens afvigelsen for Danmark (2009) er 0.027. Dvs., at EVI2 generelt estimeres for lavt for Danmark (2008) og for højt for Danmark (2009).

CI CI estimeres meget upræcist for de to rapsmarker fra Danmark (2009). Hvis man ser bort fra de absolutte afvigelser for de to rapsmarker, er der varianshomogenitet for de fire kategorier, der undersøges i denne variansanalyse. Variansanalysen viser, at der ikke er forskel på de absolutte afvigelser mellem byg og hvede. Til gengæld er der forskel mellem landskab, vertikalt klorofylprofil og forskellige niveauer af LAI. For Danmark (2008) er afvigelsen større end for Danmark (2009). Der er mindst absolut forskel på målt og simuleret

CI ved lavt LAI, og afvigelsen bliver større ved stigende LAI. For de vertikale klorofylprofiler er den absolutte afvigelse mindst ved marker, hvor der ikke er varianshomogenitet. Afvigelsen er størst ved marker, hvor middelværdierne af klorofylindholdet er forskelligt gennem det vertikale profil.

5.7.3 Opsamling

PROSAILS simuleringer af reflektans mellem 400 og 2500 nm er omregnet til gennemsnit for grøn, rød og nærinfrarød reflektans og ud fra dette er de simulerede vegetationsindekser beregnet. De simulerede reflektanser og vegetationsindekser er sammenlignet med de målte for at undersøge overensstemmelsen. Alt i alt er der forskel på de målte og simulerede reflektanser og vegetationsindekser. Der er ikke et iøjnefaldende mønster i afvigelserne mellem målt og simuleret reflektans. Derfor er forskelle i de absolutte afvigelser undersøgt for fire udvalgte kategorier: De vertikale klorofylprofiler, plantetyper, landskab og niveau af LAI.

Grøn og rød reflektans simuleret bedst ved lavt LAI, hvor de laveste absolutte afvigelser ses, mens nærinfrarød reflektans simuleres bedre ved lidt højere LAI. Alle vegetationsindekserne påvirkes også af forskelle i LAI.

De vertikale klorofylprofiler påvirker sammenhængen mellem målt og simuleret rød reflektans, men der er ingen forskel i de absolutte afvigelser mellem målt og simuleret grøn eller nærinfrarød reflektans. Plantetyper spiller en rolle i forhold til sammenhængen for grøn reflektans og ser også ud til at betyde noget i forhold til den røde reflektans. Dette kan dog ikke bekræftes gennem en variansanalyse af de absolutte afvigelser. Plantetype spiller ikke en rolle i forhold til nærinfrarød reflektans. Landskabet har indflydelse på sammenhængen mellem målt og simuleret grøn og rød reflektans, men lader ikke til at have betydning for nærinfrarød reflektans.

De spektrale vegetationsindekser påvirkes af alle fire faktorer, hvilket ikke er overraskende eftersom de er konstrueret ud fra de simulerede reflektanser. Plantetype har dog kun en betydning for NDVI og GNDVI, mens SR heller ikke påvirkes af landskabet.

6. Diskussion

I dette kapitel diskuteres de opnåede resultater. Kapitlet er inddelt i tre afsnit. Det første afsnit omhandler undersøgelsen af det vertikale klorofylprofil. Det andet afsnit beskæftiger sig med den empirisk-statistiske model og det sidste afsnit den fysisk-baserede model.

6.1 Det vertikale klorofylprofil

Det vertikale klorofylprofil betegner i denne rapport den vertikale fordeling af klorofylindholdet i en plantestand. For 68 marker med forskellige afgrøder og i forskellige landskaber er det vertikale klorofylprofil undersøgt ud fra SPAD-målinger i fem relative højder i Boegh et al. [2013]. Den statistiske metode, der anvendes i nærværende rapport, er baseret på normalfordelingsmodeller. Derfor indgår der to centrale antagelser om data: For det første antages det, at data er observerede værdier af en normalfordelt stokastisk variabel, og for det andet antages det, at der er varianshomogenitet. Den første antagelse kan testes ved at undersøge fraktildiagrammer, hvor de empiriske fraktiler sammenlignes med de teoretiske fraktiler. I afsnit C på side 139 ses fire eksempler, hvor punkter ikke afviger systematisk fra at ligge på en ret linje. Dvs. at antagelsen om normalfordeling af data kan accepteres.

Antagelsen om varianshomogenitet testes med Bartlett's test for varianshomogenitet. Der er 18 marker fordelt i landskaberne, som ikke har varianshomogenitet. Ifølge Tjur [2002b] vil man i nogle tilfælde antage varianshomogenitet på trods af, at Bartlett's test ikke viser ens varianser. Det vil typisk være, hvis datamaterialet er stort og der ikke er store relative afvigelser mellem gruppernes varianser. I dette tilfælde er antallet af observationer i de enkelte relative højder mellem 5 og 14, og det vurderes ikke, at datamaterialets størrelse tillader at acceptere hypotesen om ens varianser på trods af Bartlett's test. Derfor holdes de 18 marker adskilt i den videre analyse, og der opstilles ikke en ny hypotese om sammenhængen mellem klorofylindhold og plantens højde for disse marker.

For de øvrige 50 marker opstilles hypotesen om ens middelværdier. En accept af den hypotese betyder, at variationen i klorofylindholdet i bladene er større inden for det enkelte lag i plantestanden end mellem de relative højder. Det er ensbetydende med, at der ikke er systematisk forskel på middelværdierne mellem de fem relative højder. Eftersom variansanalysemodellen opfatter de relative højder som kvalitative faktorer og ikke ordnede kvantitative faktorer, kan man ikke afvise, at der i forhold til den relative højde er en systematisk sammenhæng med klorofylindholdet. Styrken ved variansanalysemodellen er

dog, at variationer mellem de relative højder vurderes i forhold til variationen i en relativ højde. Hvis hypotesen om ens middelværdier kan accepteres, vurderes forskellen mellem de relative højder således til at være for lille til at retfærdiggøre en hypotese om en vertikal gradient i klorofylindholdet.

6.1.1 Sammenligning med Boegh et al. [2013]

Boegh et al. [2013] anvender en simpel lineær regressionsmodel på middelværdierne for hver relativ højde uden at tage højde for eventuel manglende varianshomogenitet og bruger hældningens signifikans med et signifikansniveau på 0.1 samt $R^2 > 0.68$ til at bekræfte eller afvise en lineær sammenhæng. I denne tilgang er grundmodellen, at der er en lineær sammenhæng mellem relativ højde og klorofylindholdet. Hypotesen, der testes, er, at hældningen ikke er signifikant. En mark med uniformt vertikalt klorofylprofil er således en mark for hvilken hypotesen om en lineær sammenhæng mellem relativ højde og klorofylindholdet må afvises. Det fører til uoverensstemmelse mellem inddelingen af marker med uniformt eller lineært vertikalt klorofylprofil i Boegh et al. [2013] og i afsnit 5.2 på 79.

Resultatet fra Boegh et al. [2013] er, at 74% af markerne har en uniform fordeling af klorofylindholdet vertikalt gennem plantestanden, mens 26% har en vertikal gradient. Den vertikale gradient er voksende fra bund til top hos 72% af markerne og aftagende hos 28%. I denne undersøgelse ses det, at 18 marker ikke har varianshomogenitet. Sammenlignes fordelingen af markerne uden varianshomogenitet med resultaterne i Boegh et al. [2013] ses det, at 72% af markerne uden varianshomogenitet klassificeres som havende uniformt klorofylprofil i Boegh et al. [2013] mens 28% er klassificeret med vertikal gradient. Dvs. markerne uden varianshomogenitet findes i lige høj grad i begge kategorier fra Boegh et al. [2013].

Af de resterende 50 marker i undersøgelsen ses det, at 56% klassificeres som havende uniformt vertikalt klorofylprofil, 32% som havende lineært vertikalt klorofylprofil mens 12% hverken har uniformt eller lineært vertikalt klorofylprofil, fordi middelværdierne i de relative højder er forskellige men ikke lineært fordelt. Denne kategori kan dække over en klokkeformet fordeling eller andre typer af fordelinger. De 6 marker i kategorien klassificeres alle af Boegh et al. [2013] som marker med uniformt klorofylprofil. Det skyldes, at de ikke har en signifikant hældning i en lineær regressionsmodel. Denne undersøgelse viser dog, at der er signifikant forskel på klorofylindholdet i de fem relative højder, hvorfor der ikke er tale om en ligefordeling.

I denne undersøgelse ses det, at i alt 28 marker karakteriseres som havende uniformt klorofylprofil. Sammenlignes markerne med inddelingen i Boegh et al. [2013], ses det at 23 også af Boegh et al. [2013] havde et uniformt klorofylprofil, mens 5 havde en vertikal gradient i klorofylindholdet. Årsagen til, at Boegh et al. [2013] klassificerer nogle af disse marker med vertikal gradient, kan skyldes, at grundmodellen i Boegh et al. [2013] er en lineær sammenhæng

baseret på middelværdierne uden at tage højde for variansen i den relative højde. I variansanalysemodellen indgår variationen mellem de relative højder i forhold til variationen inden for et horisontalt lag i plantestanden. Dvs. at selv om middelværdierne i hver relativ højde kan placeres på en ret linje med signifikant hældning, så er variationen inden for den enkelte relative højde større end variationen mellem lagene. Derved karakteriseres marken anderledes i variansanalysemodellen i forhold til fremgangsmåden i Boegh et al. [2013].

Det ses endvidere, at 16 marker klassificeres med lineært vertikalt klorofylprofil, dvs. at der er en vertikal gradient i klorofylindholdet. Af dem er 7 klassificeret af Boegh et al. [2013] som havende en uniform fordeling af klorofyl, mens 9 er klassificeret med en vertikal gradient. At nogle af disse marker ændrer klassifikation med den her anvendte metode skyldes som før, at variationen i de relative højder indgår i den statistiske model. I Boegh et al. [2013] forkastes hypotesen om vertikal gradient, hvis hældningen for middelværdier ikke er signifikant med signifikansniveau på 0.1 eller hvis R^2 -værdien er under 0.67. I denne undersøgelse viser det sig altså, at hældningen i nogle tilfælde er signifikant med et signifikansniveau på 0.05, hvis alle observationer i en relativ højde indgår i stedet for blot middelværdien.

I datasættet fra Boegh et al. [2013] anvendes fem relative højder for hver mark uanset plantetype eller andre forhold som højde eller LAI. Dvs. at metoden ikke er tilpasset til en bestemt vegetationstype med særlige karakteristika eftersom det ikke er formålet med datasættet. Når datasættet imidlertid anvendes til at vurdere den vertikale fordeling af klorofyl, er der nogle begrænsninger forbundet med dette valg. Nogle af majsplanterne er tæt på 2 m høje, mens flere hvedemarker er under 20 cm høje. Nogle planters blade har et langt højere klorofylindhold end andre. Der er således stor variation mellem plantekarakteristika.

6.1.2 Andre fremgangsmåder til karakteristik af klorofylprofil

Ifølge Huang et al. [2011] har vinterhvede ligesom andre afgrøder en aftagende vertikal gradient i bladernes klorofylkoncentration fra top til bund. I nærværende rapport omtales dette aftagende forhold som en voksende lineær sammenhæng mellem klorofylindhold og relativ højde, idet klorofylindholdet stiger i forhold til højden. I tabel 5.4 på side 81 ses det, at alle hvede- og bygmarker fra Danmark (2008) samt en enkelt hvedemark fra Polen med lineære vertikale klorofylprofiler har et sådant aftagende klorofylindhold fra top til bund. Omvendt ses det ved to hvedemarker fra Danmark (2009), at klorofylindholdet er størst i bunden og aftager mod toppen. En enkelt italiensk majsmark har også samme aftagende klorofylindhold fra bund til top. For to italienske tomatmarker med lineært vertikalt klorofylprofil forholder det sig sådan, at den ene mark har en aftagende lineær sammenhæng mellem højde

og klorofylindhold mens den anden har en voksende. I forhold til hvede og byg stemmer resultatet om aftagende klorofylindhold fra top til bund overens med Huang et al. [2011] for de marker, hvor klorofylindholdet er bestemt midt i vækstsæsonen. For de danske marker fra den tidlige vækstsæson i 2009 forholder det sig derimod omvendt med et voksende klorofylindhold fra top til bund omend de fleste af disse hvedemarker har et uniformt vertikalt klorofylprofil.

I det følgende beskrives to andre fremgangsmåder i forbindelse med bestemmelse af den vertikale fordeling af klorofyl i afgrøder. Ciganda et al. [2012] undersøger, hvordan sammenhængen mellem et spektralt vegetationsindeks og klorofylindholdet i bladene på majsplanter afhænger af den vertikale fordeling af klorofyl i bladene. Det anvendte spektrale vegetationsindeks er det røde klorofylindeks fra Gitelson et al. [2005] jf. afsnit 2.2.2 på side 43:

$$CI_{\text{red edge}} = \frac{\rho_{\text{NIR}}}{\rho_{\text{red edge}}} - 1$$

Den spektrale reflektans måles med to radiometre, der peger henholdsvis op og ned. For tre majsmarker udvælges der på 13–15 dage fordelt over en vækstsæson tre planter, hvis vækst er repræsentativ for markens vækststadiet. For hver plante nummereres bladene, hvor blad 1 er det øverste blad, og analyseres i laboratoriet. Reflektansen fra det enkelte blad bestemmes på en sådan måde, at der tages højde for farveforskelle i bladene. Ud fra bladets reflektans estimeres klorofylindholdet.

Resultaterne fra Ciganda et al. [2012] viser, at majsplanter på alle tidspunkter i vækstsæsonen har en klokkeformet vertikal fordeling af klorofylindhold i bladene. I løbet af vækstsæsonen stiger klorofylindholdet for derefter at aftage, når planterne aldres. Dvs., at toppunktet på klokkeformen antager et maksimum på et tidspunkt i vækstsæsonen. En statistisk analyse viser, at den empiriske sammenhæng mellem det røde klorofylindeks og klorofylindholdet i majsmarken afhænger af de øverste 7–9 blade i plantestanden. Der er altså intet bidrag fra blade placeret mere end 1–1.25 m fra toppen. Klorofylindholdet i de øverste 7–9 blade udgør kun ca. 55–65% af det samlede klorofylindhold i plantestanden. På trods af dette resultat finder Ciganda et al. [2012] dog en god empirisk sammenhæng mellem det røde klorofylindeks og plantestandens samlede klorofylindhold.

Winterhalter et al. [2012] undersøger ligeledes den vertikale fordeling af klorofylindholdet i majs. Med et SPAD-meter foretages målinger for blade fordelt på tre majsmarker. SPAD-værdien, der repræsenterer klorofylindholdet, omregnes ikke til et konkret indhold af klorofyl, men anvendes som udtryk for det relative klorofylindhold. Resultatet viser samme klokkeformede fordeling af klorofylindholdet i majsplanten som Ciganda et al. [2012]. Den konvekse fordeling af klorofylindholdet har også her et maksimum ved de midterste blade. Winterhalter et al. [2012] finder endvidere, at en sensor kan opfange

information om bladets biomasse fra de nederste blade i plantestanden, dog med en tendens til at underestimere.

Fremgangsmåden i Ciganda et al. [2012] og Winterhalter et al. [2012] adskiller sig på flere måder fra fremgangsmåden i dataindsamlingen foretaget i Boegh et al. [2013]. Først og fremmest er de kun baseret på majsplanten. Dernæst estimeres planternes klorofylindhold blad for blad eller lag for lag. Boegh et al. [2013] estimerer klorofylindholdet i fem relative højder. Dvs. at der kan være en plantespecifik variation mellem bladenes indhold af klorofyl, der ikke opfanges, fordi nogle blade springes over. En anden forskel mellem fremgangsmåderne er, at det spektrale signal i Boegh et al. [2013] er satellitmålt og derfor påvirkes af atmosfæren.

Marker med majsplanter i undersøgelsen har oftest ikke varianshomogenitet. Det kan fx skyldes, at der er foretaget målinger på forskellige majsplanter. Ciganda et al. [2012] finder som nævnt, at majsplanten har en klokkeformet fordeling af klorofyl i bladene fra top til bund. Klokkeformen varierer over vækstsæsonen. Der er ikke beskrevet, om man kan forvente, at bladene vil sidde i samme højde på tværs af planter. Hvis forskellige planter har en variation i bladenes placering således, at de sidder tættere på hinanden hos nogle og længere fra hinanden hos andre, og der i øvrigt er en variation i klorofylindholdet, kan det bidrage til manglende varianshomogenitet, når data indsamles på tværs af planter i samme højde. Klokkeformen, som Ciganda et al. [2012] beskriver for majs, er et eksempel på en vertikal fordeling af klorofyl, som ikke har ens middelværdier ned gennem profilet og heller ikke har en lineær fordeling af klorofyl. Denne type fordeling ses hos meget få marker i denne undersøgelse i det hele taget og hos kun én majsmark.

Alt i alt viser resultaterne vedrørende det vertikale klorofylprofil, at de fleste plantetyper og alle landskaber er repræsenteret ved alle typer vertikalt profil. Dvs. at der ikke er et plante- eller landskabsspecifikt mønster i tilstedeværelsen af en bestemt type vertikalt klorofylprofil. Hos hvedemarker i den tidlige vækstsæson i Danmark (2009) er der dog en overvægt af marker med uniform vertikal klorofylfordeling. Der er dog blandt alle landskaber en del marker, der frasorteres pga. manglende varianshomogenitet.

6.2 Empirisk–statistiske modeller

De empirisk–statistiske modeller er baseret på sammenhængen mellem feltmålinger af biofysiske variable og spektrale vegetationsindekser. I afsnit 5.3 på side 82 er den målte overfladereflektans undersøgt. Det forventes, at den grønne reflektans er større end den røde reflektans for planter, hvilket bekræftes for marker med LAI over 2 svarende til, at jordbunden er dækket med vegetation. Den røde reflektans forventes at konvergere mod et minimum for voksende LAI, mens den nærinfrarøde reflektans omvendt konvergerer mod et maksimum. Begge dele bekræftes også for alle landskaber. Dog er den

nærinfrarøde reflektans afhængig af landskabet og på figur 5.2 på side 83 ses det fx, at den nærinfrarøde reflektans fra de italienske marker er lavere end reflektansen fra de andre landskaber. Hvis man sammenholder den målte reflektans fra de danske marker med jordbundskortet på figur 3.3 på side 67 ses det, at den grønne reflektans er lavere fra markerne med en lerblandet sandjord i forhold til markerne med en sandblandet lerjord. Ensidet variansanalyse viser dog, at forskellen ikke er signifikant ($p = 0.21$). Der er ikke forskel på reflektansen af rød og nærinfrarød stråling.

Den empiriske sammenhæng er undersøgt for forskellige landskaber og ikke for plantetyper. Det skyldes dels, at der forventes at være en stærk afhængighed af konteksten i form af jordbund og atmosfæriske forhold jf. afsnit 2.2.2 på 38, og dels at der er en skæv fordeling af plantetyper i undersøgelsen med en stor andel hvedemarker. Til gengæld er det undersøgt, om tilstedeværelsen af en vertikal gradient i klorofylindholdet har en betydning for sammenhængen på tværs af landskaber.

For at undersøge den empiriske sammenhæng mellem de spektrale vegetationsindekser og LAI eller $\text{Chl}_{\text{total}}$ anvendes simpel lineær regression og ikke-lineær regression. Simpel lineær regression anvendes ved SR og CI, mens ikke-lineær regression anvendes ved de normerede vegetationsindekser baseret på NDVI pga. det asymptotiske forhold mellem disse indekser og LAI. I bl.a. Viña et al. [2011] skelnes på samme vis mellem lineær og ikke-lineær regression i forhold til forskellige typer af vegetationsindekser. SR og CI findes i Viña et al. [2011] at have et lineært forhold til LAI, mens sammenhængen for NDVI og EVI er en ikke-lineær funktion på formen $y = c_1 + c_2(1 - \exp(-Kx))$, hvor konstanterne c_1 , c_2 og K bestemmes ved ikke-lineær regression.

6.2.1 Lineær regression

I forhold til anvendelse af den lineære regressionsmodel må man som tidligere nævnt foretage en række nødvendige antagelser bl.a. at x -værdierne er ikke-stokastiske. I dette tilfælde er x -værdierne LAI og $\text{Chl}_{\text{total}}$ ikke målt uden måleusikkerhed. Som nævnt i afsnit 3.2.1 på side 68 og afsnit 3.2.2 på side 71 er usikkerheden på LAI mellem 22–25% og på $\text{Chl}_{\text{total}}$ 29–45%. Der er altså i høj grad tale om måleusikkerhed, hvorfor antagelsen om ikke-stokastiske x -værdier ikke er opfyldt. Det bidrager til en indbygget usikkerhed i modellen.

Generelt viser resultaterne, at der er en sammenhæng mellem de spektrale vegetationsindekser og LAI. For SR og CI kan sammenhængen beskrives ved en lineær model med en statistisk signifikant hældning og hvis R^2 -værdier forbedres, når sammenhængen undersøges for det enkelte landskab. Det gælder dog ikke for observationerne fra Danmark (2009). Som nævnt i afsnit 2.1.4 på side 33 er der svagheder forbundet med at anvende R^2 til modelkontrol. Det gøres dog for sammenhængen mellem spektrale vegetationsindekser og LAI i bl.a. Boegh et al. [2013].

Undersøges sammenhængen mellem SR og CI for markerne med et uni-

formt vertikalt klorofylprofil ses også en forbedring af R^2 -værdien i forhold til alle landskaber samlet. R^2 -værdien bliver dog ikke lige så høj som for de enkelte landskaber. Det kan forklares ved vegetationsindeksernes kontekstafhængighed. For at undersøge betydningen af det vertikale klorofylprofil kunne man derfor også have undersøgt, om der var en stærkere sammenhæng mellem LAI og vegetationsindekset for markerne med uniformt vertikalt klorofylprofil inden for et enkelt landskab. Landskabet med flest marker med uniformt vertikalt klorofylprofil er Danmark (2009) har samtidig en svag sammenhæng mellem LAI og vegetationsindekser. Her ses det dog, at R^2 -værdi forbedres betydeligt for markerne med uniformt vertikalt klorofylprofil. Der er dog stadig en højere R^2 -værdi for landskaberne Italien og Holland. De øvrige landskaber har maksimalt seks marker med uniformt vertikalt klorofylprofil, hvilket anses som for få observationer at foretage lineær regression.

Selv om CI teoretisk set burde være knyttet til klorofylindholdet i plantestanden, er sammenhængen med LAI stærkere end sammenhængen med $\text{Chl}_{\text{total}}$ i resultaterne i denne rapport. Til sammenligning finder Gitelson et al. [2005] en nær lineær sammenhæng mellem $\text{Chl}_{\text{total}}$ og klorofylindekset med $R^2 = 0.92$ for majs og $R^2 = 0.93$ sojabønner. Som nævnt i afsnit 2.2.2 på side 43 findes klorofylindekset også i en udgave, der bruger den røde reflektans tæt på grænsen til nærinfrarøde bølgelængder. Dette røde klorofylindeks har i Gitelson et al. [2005] en bedre sammenhæng med $\text{Chl}_{\text{total}}$ på tværs af plantetyper. En lineær model på tværs af de to plantetyper majs og sojabønne vises at have $R^2 = 0.70$ for det grønne klorofylindeks og $R^2 = 0.89$ for det røde klorofylindeks. Resultaterne i nærværende rapport har langt lavere R^2 -værdier som det fremgår af tabel 5.6 på side 90. På tværs af landskaber og plantetyper er $R^2 = 0.25$. Det enkelte landskab på tværs af plantetyper har maksimalt $R^2 = 0.62$, mens hvedemarker med uniformt vertikalt klorofylprofil fra Danmark (2009) har $R^2 = 0.58$. Der er dog en ret høj usikkerhed på estimatet af $\text{Chl}_{\text{total}}$ på 29–45%.

6.2.2 Ikke-lineær regression

Sammenhængen mellem NDVI, GNDVI og EVI2 og LAI beskrives vha. Baret og Guyots "semi-empiriske" ikke-lineære model:

$$VI = VI_{\text{max}} + (VI_{\text{min}} - VI_{\text{max}})\exp(-K \cdot \text{LAI})$$

Sammenhængen er ikke-lineær pga. udslukningskoefficienten K , mens forholdet mellem VI og $\exp(-K \cdot \text{LAI})$ er lineært. Det betyder, at sammenhængen ikke kan transformeres til en lineær sammenhæng. Havde det været tilfældet, kunne lineær regression anvendes på transformerede x og y -værdier. VI_{max} og VI_{min} er ukendte værdier. VI_{max} er den maksimale værdi af vegetationsindekset for det givne område ved det tættest mulige vegetationsdække, mens VI_{min} er værdien for bar jord.

I R foretages ikke-lineær regression ved, at der vælges startværdier for de

tre ukendte parametre i modellen og derefter itereres der mod den sammensætning af parametrene, der minimerer de kvadratiske afvigelser mellem data og model. Dvs. at der er følsomhed over for valg af startværdier. Ikke-lineær regression er derved en anderledes metode end lineær regression, og der er ikke samme mulighed for at foretage modelkontrol. Det skyldes, at der ikke er en grundmodel at teste imod. Derfor inddrages residualplots, der kan afsløre systematiske fejl i modellen.

Når modellen anvendes i nærværende rapport for observationer fra alle landskaber, er der en systematisk afvigelse ved lavt LAI. Denne systematiske fejl ses ikke, når modellen anvendes på de danske landskaber. Det skyldes formentlig, at sammenhængen mellem vegetationsindekser og LAI er kontekstafhængig. Endvidere kan man se på figur 5.6 på side 89, at NDVI er højt for markerne i Danmark (2009) uanset LAI. På figur 5.6 på side 89 er marker med et uniformt vertikalt klorofylprofil markeret. Ud fra punkternes placering er der ikke noget, der tyder på, at den empiriske sammenhæng mellem LAI og vegetationsindekserne forbedres væsentligt for marker med uniformt vertikalt klorofylprofil.

6.2.3 Følsomhedsanalyse af vegetationsindekserne

Vegetationsindeksernes følsomhed over for forskellige biofysiske størrelser er undersøgt vha. simuleret reflektans fra PROSAIL. Følsomheden er testet med udgangspunkt i landskabet Danmark (2008) for vinkler samt de øvrige plante-standsspecifikke parametre. Det betyder, at bladvinklen er varieret omkring 60° . Det er en relativt stejl bladvinkel, og sandsynligvis ville følsomheden over for bladvinklen aftage ved lavere hældning, da jordbundens indflydelse mindskes. Følsomheden er dog testes for forskellige niveauer af LAI henholdsvis 1, 3 og 6 for at vurdere betydningen af LAI. Det er valgt, da datasættet indeholder LAI, der varierer mellem lige over 0 og tæt på 6. Klorofylindholdet er $47.5 \mu \text{g/cm}^2$, og der er ikke undersøgt følsomhed ved andre niveauer af klorofylindhold, da variationen i bladklorofyl er lav for de danske landskaber.

Resultatet af følsomhedsanalysen viser, at vegetationsindekser er mest følsomme over for LAI, middelbladvinkel og klorofylindhold i bladene. Bladvinklens påvirkning er størst for den nærinfrarøde reflektans og vegetationsindekserne SR, EVI2 og CI. Det grønne klorofylindeks CI har klart den største følsomhed over for bladklorofyl. Det grønne NDVI, der burde være klorofylfølsomt, har dog en lav følsomhed over for bladklorofyl, der overstiges af både SR og CI.

I forhold til LAI ses det, at den største følsomhed findes hos SR og CI. Følsomheden over for LAI for alle vegetationsindekserne aftager ved store værdier af LAI. Hos NDVI bliver følsomheden dog meget lille (0.1-0.2 % ved ændring på 10 % i LAI). NDVI har altså en tydelig tendens til at blive mættet ved højt LAI. EVI2 har ikke en ligeså høj grad af mætning, hvilket svarer til resultater for EVI fra bl.a. Boegh et al. [2002].

Huete et al. [1997] finder, at NDVI er tæt knyttet til den røde reflektans, mens EVI er mere knyttet til den nærinfrarøde reflektans. Den samme forskel mellem NDVI og EVI2 bekræftes i høj grad af følsomhedsanalysen i nærværende rapport. Følsomheden over for jordbundens reflektans viser, at EVI2 falder, når den nærinfrarøde reflektans falder, mens NDVI stiger ved faldende nærinfrarød reflektans for lavt LAI og aftager ved faldende nærinfrarød reflektans for højt LAI som det ses på tabel 5.12 på side 102, tabel 5.13 på side 103 og tabel 5.14 på side 103.

Baret og Guyot [1991] undersøger følsomheden i relationen mellem NDVI og LAI over for jordbund og middelbladvinkel og finder, at NDVI forstyrres af jorden. Endvidere er NDVI mest følsom over for bladvinklen for LAI <0.5 og LAI >3, især hvis middelbladvinklen er lav. Forstyrrelsen falder ved voksende LAI op til LAI=3, da jordbundens indflydelse mindskes. SAVI er til gengæld i stand til at reducere forstyrrelsen for lavt LAI. Desuden finder Baret og Guyot [1991], at det generelt er svært at estimere LAI ud fra vegetationsindekser, når vegetationsindeksets værdi er tæt på VI_{\max} .

Når vegetationsindeksernes følsomhed sammenlignes med resultaterne for den empirisk-statistiske model i denne rapport bekræftes det, at der er et tilnærmelsesvist lineært forhold mellem SR og CI på den ene side og LAI på den anden side. Af de øvrige vegetationsindekser findes EVI2 til at være mere følsomt over for LAI end NDVI og GNDVI. Det kan dog ikke bekræftes af resultaterne i undersøgelsen, hvor EVI2 lader til at konvergere hurtigere end NDVI og GNDVI.

I forhold til det vertikale klorofylprofil viser følsomhedsanalysen af vegetationsindekserne, at der er en ret lille følsomhed over for klorofylindholdet i bladene hos NDVI, GDNVI og EVI2. Det lader altså til, at variation i klorofylindhold under de omstændigheder, følsomhedsanalysen er lavet ud fra, ikke vil påvirke den empiriske sammenhæng mellem de normerede vegetationsindekser og LAI. SR og især CI er til gengæld mere følsomme over for variation i klorofylindholdet, og resultaterne viser en forbedring af den empiriske sammenhæng for marker med uniformt vertikalt klorofylprofil. Forbedringen er dog endnu bedre blandt marker inden for samme landskab, hvilket bekræfter vegetationsindeksernes kontekstafhængighed.

6.3 Fysisk-baseret modellering

I dette afsnit diskuteres resultaterne vedrørende PROSAIL. Først diskuteres følsomhedsanalysen og til sidst sammenhængen mellem målt og simuleret reflektans for de udvalgte marker.

6.3.1 Følsomhedsanalyse

PROSAILs følsomhed er vurderet ved at justere på de fleste af inputparametrene og beregne den relative ændring i reflektansen i SPOTs frekvensbånd for grøn, rød og nærinfrarød stråling gennem middelværdien i bølglængderne for de tre intervaller. For at undersøge betydningen af LAI er samme procedure foretaget for LAI på 1, 3 og 6, hvor LAI på 3 og 6 viste samme resultat. I opsætningen er data for en mark fra landskabet Danmark (2008) anvendt og ukendte parametre er fastsat efter litteratur. I forbindelse med følsomhedsanalysen er Matlab-implementeringen af PROSAIL justeret for at rette op på en fejl i koden vedrørende den ellipsoide bladvinkelfordeling. Justeringen i sig selv er dog også en fejlkilde, da den ikke fininddeler intervallet mellem 80° og 90° i simuleringen af reflektansen.

Det vurderes dog, at denne fejlkilde er lav, eftersom følsomheden ved anvendelse af den sfæriske bladvinkelfordeling med oprindelig Matlab-implementering, der tager højde for intervallets betydning, er sammenlignelig med følsomheden ved anvendelse af den ellipsoide bladvinkelfordeling med middelbladvinkel på 60° jf. tabel 5.9 på side 98 og tabel 5.10 på side 98.

Følsomhedsanalysen viser, at den grønne reflektans er mest følsom over for ændringer i klorofylindholdet og dernæst LAI og middelbladvinkel. Hvis LAI er 1, er den grønne reflektans dog mere følsom over for LAI og middelbladvinkel end klorofyl. Den røde reflektans påvirkes mest af LAI og middelbladvinkel og dernæst klorofylindholdet uanset om LAI er 1 eller højere. Den nærinfrarøde reflektans er følsom over for ændringer i middelbladvinkel og LAI. Klorofyl har som forventet ingen indflydelse på nærinfrarød reflektans.

I følsomhedsanalysen er de indgåede parametre varieret, mens de øvrige holdes konstante. Dette scenarium er dog urealistisk eftersom de biofysiske variable er tæt korreleret [Jacquemoud et al., 2009]. Gennem en vækstsæson varierer bladenes sammensætning, bladvinkel og LAI samtidig.

Jacquemoud et al. [2009] præsenterer resultaterne af en anden følsomhedsanalyse af PROSAIL, der vha. en anden metode undersøger de relative bidrag fra de enkelte variable og kombinerede effekter fra samvirkende variable. Når variablene betragtes hver for sig, konkluderes det, at klorofylindhold i bladene styrer ca. 60% af variationen i reflektansen af synligt lys. Variationen i den nærinfrarøde reflektans styres ligeligt af LAI og middelbladvinklen, der hver især forklarer ca. 40% af variationen. I PROSAIL interagerer variablene, hvis effekten fra én variabel på den spektrale reflektans afhænger af niveauet på en anden variabel. Dvs. en kombineret ændring i to variable kan producere en større effekt på resultatet end summen af effekter fra hver variabel. I det synlige lysspektrum ses det, at variabelparrene klorofylindhold og LAI samt klorofylindhold og middelbladvinkel begge forklarer ca. 8% af variationen i den simulerede reflektans. I den nærinfrarøde del af spektret forklarer LAI og middelbladvinkel 5-6% af variationen.

I følsomhedsanalysen i denne rapport ses det også, at variation i middel-

bladvinkel og LAI har størst betydning for variation i PROSAILs simulering af nærinfrarød reflektans. I forhold til synligt lys, er grøn reflektans i denne undersøgelse mest følsom over for klorofyl mens rød reflektans er mest følsom over for LAI. Det stemmer ikke overens med følsomhedsanalysen, der præsenteres i Jacquemoud et al. [2009], hvor klorofyl som nævnt vises at være styrende for reflektansen af synligt lys. Fremgangsmåderne i de to følsomhedsanalyser er dog ikke umiddelbart sammenlignelige. Følsomhedsanalysen i denne rapport tager ikke højde for kombinerede effekter af flere variable.

I følsomhedsanalysen i denne rapport er simuleringerne lavet for forskellige niveauer af LAI på henholdsvis 1, 3 og 6, da det repræsenterer variationen i datasættet. Middelbladvinklen er kun varieret omkring 60° . En følsomhedsanalyse omkring en lavere hældning på bladene kan som førnævnt formentlig ændre på resultatet vedrørende middelbladvinklens betydning for reflektansen.

Simuleringerne er lavet med en række inputparametre, der er kendte eller fastsat ud fra litteraturen. LAI samt indhold af klorofyl og tørstof er målt i Boegh et al. [2013]. Estimatere af LAI og klorofylindhold er dog behæftet med en grad af usikkerhed som tidligere nævnt. Som det ses i følsomhedsanalysen, har LAI og klorofylindhold en stor betydning for reflektansen, hvorfor estimaternes usikkerhed skal tages i betragtning. Middelbladvinklen er for Danmark (2008) estimeret sammen med LAI. Data mangler for Danmark (2009) og usikkerheden på estimatere fra 2008 er ikke vurderet.

I undersøgelsen vælges den ellipsoidale bladvinkelfordeling, hvor middelbladvinklen sættes til den estimerede med LAI-2000. Pga. den ukendte usikkerhed på estimatet af middelbladvinklen og det relativt lave antal målinger pr. mark er det usikkert, om den valgte bladvinkelfordeling er repræsentativ for de marker, der indgår i undersøgelsen. Plantetyperne er hvede, byg, raps og majs. Hvedemarkerne er fra to forskellige tidspunkter i vækstsæsonen i to forskellige år (2008 og 2009), og i 2008 er der både vår- og vinterhvede. Bygmarkerne fra 2008 er også en blanding af vår- og vinterbyg. Der er kun to rapsmarker og en enkelt majsmark med LAI tæt på nul. LAI-2000-estimatere fra 2008 af hvede og byg angiver bladvinkler tæt på 60° .

Ifølge Wang et al. [2009] vil bladvinkelfordeling for en afgrøde variere over forskellige vækststadier og mellem forskellige plantetyper. Andre faktorer som mangel på næringsstoffer og miljøforhold kan også give ændringer i bladens orientering. Huang et al. [2011]; Wang et al. [2009] undersøger vinterhvede og inddeler hvedesorter i erektofile og planofile typer ud fra en størrelse kaldet *leaf orientation value* (LOV), der er baseret på bladets hældning i forhold til bladets længde. Huang et al. [2011] introducerer også kategorien horisontale typer for særligt lave værdier af LOV.

Jordbundens reflektans er fastsat ved aflæsning af grafer baseret på målinger af jordbundens reflektans i samme område i 1997 foretaget af Broge [2003]. Det er ikke klart, hvilken jordtype der er undersøgt i Broge [2003]. Ud fra

vejrforholdene vurderet med oplysninger fra DMI bør jordbunden have været tør på begge tidspunkter for målinger, hvorfor data for tør jordbund er anvendt. Majsmark nr. 5 i Danmark (2008) har et estimeret LAI tæt på 0, hvorfor den SPOT-målte reflektans kan give en indikation af, om der er overensstemmelse mellem det anvendte estimat og den faktiske jordbundsreflektans. Den grønne reflektans fra majsmarken er ca. 0.08, mens den anvendte jordbundsreflektans er ca. 0.11 i den grønne del af spektret. For den røde del af spektret har majsmarken en reflektans på 0.11 mens jordbundens estimerede reflektans er ca. 0.15. I den nærinfrarøde del af spektret er majsmarkens reflektans 0.23 mens den estimerede jordbundsreflektans er ca. 0.24. Majsmarken er dog placeret tæt på en vej jf. figur 3.3 på side 67. Det lader dog til, at der kan være en uoverensstemmelse mellem den faktiske og den anvendte estimerede jordbundsreflektans.

Hotspotparameteren er estimeret som funktion af LAI jf. afsnit 2.2.3 på side 62. Følsomhedsanalysen viser, at hotspotparameteren kun har en lille betydning for den simulerede reflektans. De øvrige parametre vedrører bladpigmenter og omfatter indhold af carotenoider, vand og brunt pigment. Indholdet af carotenoider er fastsat som funktion af klorofylindholdet jf. afsnit 2.2.3 på side 46 og har udelukkende betydning i den grønne reflektans. De anvendte frekvensbånd er meget lidt følsomme over for vandindholdet. Indholdet af brunt pigment ses der helt bort fra i denne undersøgelse, hvorfor det dermed antages, at bladene ikke er aldret men udelukkende friske, grønne blade. Der er således også et bidrag til simuleringernes usikkerhed fra disse parametre.

6.3.2 Simuleret og målt reflektans

Sammenligning af PROSAILs simulering af reflektansen med den målte reflektans viser, at der generelt er afvigelser mellem simuleret og målt reflektans. Afvigelsen er mindst for lavt LAI, hvor reflektans fra jordbunden spiller en større rolle. Simuleringerne af grøn og rød reflektans er har de mindste afvigelser ved lavt LAI. For den røde reflektans spiller det vertikale klorofylprofil tilsyneladende en rolle. De absolutte afvigelser mellem målt og simuleret reflektans er mindst for marker uden varianshomogenitet og dernæst for marker med uniformt vertikalt klorofylprofil. Afvigelsen mellem målt og simuleret rød reflektans påvirkes tilsyneladende også af plantetype og landskab. For den grønne reflektans er der ikke et mønster i afvigelserne i forhold til de vertikale klorofylprofiler. Til gengæld er der en indflydelse fra plantetype. Den nærinfrarøde reflektans simuleres bedst ved lidt højere LAI, og der er tilsyneladende ingen indflydelse fra plantetype, landskab eller vertikalt klorofylprofil.

Vegetationsindekserne påvirkes af afvigelserne på estimerede af grøn, rød og nærinfrarød reflektans og både vertikalt klorofylprofil, LAI, plantetype og landskab lader til at have indflydelse på afvigelserne mellem målt og simuleret værdi af vegetationsindekset.

Når man sammenligner markernes placering på arealanvendelseskortet

i figur 3.2 på side 66, ses der intet klart mønster i afvigelserne. En enkelt hvedemark (nr. 15) fra 2009 ligger tæt på bebyggelse og har en høj reflektans af synligt lys på over 0.06, hvilket kan skyldes en forstyrrelse af signalet. Der ses ingen øvrig sammenhæng mellem afvigelser i simuleringen af reflektansen og placering nær vej, skov eller mose på arealanvendelseskortet. I forhold til jordbunden på figur 3.3 på side 67 simuleres nærinfrarød reflektans tilsyneladende bedst for marker med lerblandet sandjord. Det kan dog ikke bekræftes statistisk. Der ses ingen forskel på afvigelserne for rød og grøn reflektans i forhold til jordbunden.

7. Konklusion

Formålet med specialet er at undersøge, hvorvidt den vertikale fordeling af klorofylindholdet i en plantestand har en betydning for overfladereflektansen fra en landbrugsmark. Problemformuleringen er besvaret på baggrund af et datasæt fra Boegh et al. [2013].

For 68 marker fordelt i fem landskaber findes data for klorofylindholdet målt på et antal blade i fem relative højder gennem plantestanden. Der er foretaget en statistisk behandling af datasættet med henblik på at kategorisere markernes vertikale fordeling af klorofyl. I den forbindelse er der i nærværende rapport opstillet tre kategorier til at karakterisere det vertikale klorofylprofil: Uniformt, lineært og uden ens middelværdier. Et uniformt vertikalt klorofylprofil er kendetegnet ved, at der ikke er systematisk forskel på klorofylindholdet i de enkelte lag i plantestanden. Et lineært vertikalt klorofylprofil er kendetegnet ved, at der er en lineær sammenhæng mellem relativ højde og klorofylindhold i plantestanden. Denne sammenhæng kan være voksende eller aftagende fra bund til top. Den sidste kategori indeholder marker, hvor begge hypoteser må forkastes: Her er der systematisk forskel på klorofylindholdet i de enkelte lag, men der er ikke en lineær sammenhæng mellem relativ højde og klorofylindhold. Der er anvendt ensidet variansanalyse og simpel lineær regression, som begge er lineære normalfordelingsmodeller. En forudsætning for at anvende metoderne er, at data kan betragtes som observationer af en normalfordelt stokastisk variabel, og at der er ens varianser. Det viser sig, at der for et antal marker ikke er ens varianser i mellem de relative højder, hvor klorofylindholdet er målt, hvorfor der ses bort fra disse marker.

Datasættet er i forvejen behandlet i Boegh et al. [2013] med henblik på at identificere en eventuel vertikale gradient i klorofylindholdet, men specialets fremgangsmåde giver anledning til en anden inddeling af markerne end Boegh et al. [2013]. Det skyldes, at der i specialet først undersøges, om klorofylindholdet er uniformt fordelt, og dernæst om der er en lineær sammenhæng mellem klorofylindhold og højde. Fremgangsmåden i Boegh et al. [2013] er at bestemme den lineære sammenhæng for dernæst at teste hældningens signifikans.

Alt i alt ses det, at det vertikale klorofylprofil hverken er plante- eller landskabsspecifikt. Dog er der en overvægt af hvedemarker med uniformt vertikalt klorofylprofil i den tidlige vækstsæson. Der er også en stor andel af majsmarkerne, som mangler varianshomogenitet.

Den empiriske sammenhæng mellem vegetationsindekserne og LAI og $\text{Chl}_{\text{total}}$ er generelt bedst for enkelte landskaber. For marker med uniformt vertikalt klorofylprofil forbedres sammenhængen vurderet ved R^2 -værdien på

tværs af landskaber, men den bliver ikke bedre end sammenhængen inden for enkelte landskaber. Inden for landskabet Danmark (2009) forbedres sammenhængen dog betragteligt, når man ser bort fra marker uden uniformt vertikalt klorofylprofil.

Der er udvalgt fem spektrale vegetationsindekser med forskellige egenskaber. Simple ratio (SR) er tilnærmelsest lineært knyttet til LAI og er et ofte anvendt vegetationsindeks. NDVI er en normaliseret udgave af SR, der har en tendens til at opføre sig asymptotisk ved høje værdier af LAI. Knyttet til NDVI er GNDVI og EVI2. GNDVI er en teoretisk set klorofylfølsom udgave af NDVI, mens EVI2 har en mindre følsomhed over for jordbundens reflektans. Det sidste indeks er klorofylindekset CI, der som navnet antyder er udviklet til at estimere klorofylindhold.

De spektrale vegetationsindeksernes følsomhed er undersøgt ved at anvende simuleret reflektans fra PROSAIL og variere på inputparametrene. Reflektansen er simuleret for værdier af LAI fra 1 til 6, svarende til lav, mellem og høj vegetationstæthed. Det viser sig, at alle vegetationsindekserne er følsomme over for bladvinkel og LAI. Bladvinklen er en planteegenskab, der afhænger af faktorer som plantetype, adgang til næringsstoffer og tidspunkt i vækstsæsonen. Vegetationsindekserne vises at være i høj grad følsomme over for bladvinklen. Det kan til dels forklare vegetationsindeksernes kontekstafhængighed. SR og CI har den største følsomhed over for LAI for alle niveauer af vegetationstæthed, mens især NDVI og GNDVI som forventet mættes ved LAI over 3. EVI2 mættes i lidt mindre grad. Følsomhedsundersøgelsen bekræfter, at EVI2 er tæt knyttet til den nærinfrarøde reflektans, mens NDVI er tættere relateret til den røde reflektans. Især klorofylindekset CI har i øvrigt en høj følsomhed over for klorofylindholdet i bladene. GNDVI vises til gengæld at være meget lidt følsom over for klorofyl.

PROSAILs følsomhed over for inputparametrene er vurderet. Det ses, at den grønne reflektans påvirkes mest af ændringer i klorofylindholdet og dernæst af LAI samt middelbladvinkel. For lavt LAI er den grønne reflektans dog mere følsom over for LAI og middelbladvinkel end klorofylindhold. Den røde reflektans er primært følsom over for LAI og middelbladvinkel og dernæst klorofylindhold, mens den nærinfrarøde reflektans slet ikke er klorofylfølsom. Her spiller middelbladvinklen til gengæld en stor rolle sammen med LAI. De øvrige inputparametre har en mindre betydning for den simulerede reflektans. Af de mest betydningsfulde er mesofylstrukturparameteren i forhold til den grønne reflektans og jordbundens reflektans for den røde reflektans.

Reflektansen er simuleret for de danske marker fra 2008 og 2009. Der er generelt afvigelser mellem den målte og den simulerede reflektans. Der er flere forhold, der påvirker overensstemmelsen mellem målt og simuleret reflektans. En væsentlig faktor er LAI. Den grønne og røde reflektans simuleres bedst ved lavt LAI, mens den nærinfrarøde reflektans simuleres bedst ved lidt højere LAI. I forhold til simuleringen af den røde reflektans er der forskel på

overensstemmelsen mellem de forskellige typer af vertikalt klorofylprofil. Den mindste absolutte afvigelse ses hos markerne uden varianshomogenitet og dernæst hos markerne med uniformt vertikalt klorofylprofil.

Appendiks

A. Ligningssystemet fra SAIL

Dette kapitel indeholder mellemregninger til afsnit 2.2.3.

En diagonal matrix er en matrix på formen:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Λ er en diagonal matrix med egenverdierne til M i diagonalen.

Bestemmelse af egenverdierne Egenverdierne bestemmes ved at finde nulpunkterne til det karakteristiske polynomium $\det(M - xI)$, hvor I er identitetsmatricen. I dette tilfælde er det følgende determinant, der skal bestemmes:

$$\begin{aligned} \begin{vmatrix} k-x & 0 & 0 & 0 \\ -s & a-x & -\sigma & 0 \\ s' & \sigma & -a-x & 0 \\ w & v & u & -K-x \end{vmatrix} &= (-K-x) \begin{vmatrix} k-x & 0 & 0 \\ -s & a-x & -\sigma \\ s' & \sigma & -a-x \end{vmatrix} \\ &= (-K-x)(k-x) \begin{vmatrix} a-x & -\sigma \\ \sigma & -a-x \end{vmatrix} \\ &= (-K-x)(k-x)(x^2 - a^2 + \sigma^2) \end{aligned}$$

Hvis man indfører $m = \sqrt{a^2 - \sigma^2}$, så kan udtrykket omskrives til $(-K-x)(k-x)(x^2 - m^2)$. Dvs. egenverdierne er $\lambda_1 = k$, $\lambda_2 = m$, $\lambda_3 = -m$ og $\lambda_4 = -K$. Dermed er matricen Λ :

$$\Lambda = \begin{pmatrix} k & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & -m & 0 \\ 0 & 0 & 0 & -K \end{pmatrix}$$

Den lineære transformation består af en matrix med egenvektorerne som søjler. Da der er ni forskellige koefficienter i ligningssystemet, består egenvektorerne af kombinationer heraf og det viser sig at være behændigt at indføre nye koefficienter jf. Verhoef [1985]:

$$\begin{aligned}
h_1 &= \frac{a+m}{\sigma} & h_2 &= \frac{a-m}{\sigma} \\
C_s &= \frac{s'(k-a) - s\sigma}{k^2 - m^2} & C_o &= \frac{v(K-a) - u\sigma}{K^2 - m^2} \\
D_s &= \frac{-s(k+a) - s'\sigma}{k^2 - m^2} & D_o &= \frac{-u(K+a) - v\sigma}{K^2 - m^2} \\
H_s &= \frac{uC_s + vD_s + w}{K+k} & H_o &= \frac{sC_o + s'D_o + w}{K+k}
\end{aligned}$$

Det gælder, at $h_2 = \frac{1}{h_1}$:

$$\frac{1}{h_1} = \frac{\sigma}{a+m} = \frac{\sigma(a-m)}{a^2 - m^2} = \frac{\sigma(a-m)}{a^2 - a^2 + \sigma^2} = \frac{a-m}{\sigma} = h_2$$

Bestemmelse af egenvektorerne For hver egenværdi $\lambda_1, \dots, \lambda_4$ bestemmes den tilhørende egenvektor ved at løse ligningssystemet: $(M - \lambda_i \cdot I)\mathbf{v} = \mathbf{0}$.

For $\lambda_1 = k$ ser ligningssystemet således ud:

$$\begin{aligned}
&\left(\begin{array}{cccc|c} k-k & 0 & 0 & 0 & 0 \\ -s & a-k & -\sigma & 0 & 0 \\ s' & \sigma & -a-k & 0 & 0 \\ w & v & u & -K-k & 0 \end{array} \right) \\
&\sim \left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ -s & -(k-a) & -\sigma & 0 & 0 \\ s' & \sigma & -(k+a) & 0 & 0 \\ w & v & u & -(K+k) & 0 \end{array} \right)
\end{aligned}$$

Der kan opskrives tre ligninger:

$$-sx_1 - (k-a)x_2 - \sigma x_3 = 0 \quad (\text{A.1})$$

$$s'x_1 + \sigma x_2 - (k+a)x_3 = 0 \quad (\text{A.2})$$

$$wx_1 + vx_2 + ux_3 - (K+k)x_4 = 0 \quad (\text{A.3})$$

Da der er en nulrække i matricen, kan x_1 vælges frit. x_1 sættes lig t , og dette indsættes i ligning A.1 og x_3 isoleres:

$$-st - (k-a)x_2 - \sigma x_3 = 0 \Leftrightarrow x_3 = -\frac{k-a}{\sigma}x_2 - \frac{s}{\sigma}t$$

Udtrykket for x_3 indsættes i ligning A.2:

$$\begin{aligned}
s't + \sigma x_2 - (k+a)\left(-\frac{k-a}{\sigma}x_2 - \frac{s}{\sigma}t\right) &= 0 & (\text{A.4}) \\
\Leftrightarrow \frac{s(k+a) + s'\sigma}{\sigma}t + \frac{\sigma^2 - a^2 + k^2}{\sigma}x_2 &= 0 \\
\Leftrightarrow \frac{s(k+a) + s'\sigma}{\sigma}t &= -\frac{\sigma^2 - a^2 + k^2}{\sigma}x_2 \\
\Leftrightarrow (s(k+a) + s'\sigma)t &= -(k^2 - m^2)x_2 \\
\Leftrightarrow x_2 &= \frac{-s(k+a) - s'\sigma}{k^2 - m^2}t = D_s t
\end{aligned}$$

Udtrykket for x_2 i ligning A.4 sættes ind i ligning A.1 og x_3 isoleres:

$$\begin{aligned}
-st - (k-a)D_s t - \sigma x_3 &= 0 \\
\sigma x_3 &= (-s - (k-a)\frac{-s(k+a) - s'\sigma}{k^2 - m^2})t \\
x_3 &= \left(-\frac{s}{\sigma} - \frac{-s(k^2 - a^2) - s'\sigma(k-a)}{\sigma(k^2 - m^2)}\right)t \\
&= \frac{-s(k^2 - m^2) + s(k^2 - a^2) + s'\sigma(k-a)}{\sigma(k^2 - m^2)}t \\
&= \frac{-sk^2 + sk^2 - sm^2 + sa^2 + s'\sigma(k-a)}{\sigma(k^2 - m^2)}t \\
&= \frac{-s(a^2 - \sigma^2) + sa^2 + s'\sigma(k-a)}{\sigma(k^2 - m^2)}t \\
&= \frac{s'\sigma(k-a) - s\sigma^2}{\sigma(k^2 - m^2)}t \\
&= \frac{s'(k-a) - s\sigma}{k^2 - m^2}t = C_s
\end{aligned}$$

Dermed er både x_1 , x_2 og x_3 udtrykt vha. t , og x_4 kan således også udtrykkes vha. t ved at indsætte de øvrige udtryk i ligning A.3:

$$\begin{aligned}
wt + vD_s t + uC_s t - (K+k)x_4 &= 0 \\
(K+k)x_4 &= (w + vD_s + uC_s)t \\
x_4 &= \frac{w + vD_s + uC_s}{K+k}t = H_s t
\end{aligned}$$

Dermed er egenvektoren x_1 knyttet til egenværdien λ_1 givet ved $x_1 = (1, D_s, C_s, H_s)t$.

De øvrige egenvektorer bestemmes vha. samme metode. Se udregning i A. Egenvektoren v_2 knyttet til egenværdien $\lambda_2 = m$ er givet ved $v_2 = (0, h_1, 1, h_1 C_o - D_o)t$ og for $\lambda_3 = -m$ er egenvektoren $v_3 = (0, h_2, 1, h_2 C_o - D_o)t$. Den fjerde egenvektor er $v_4 = (0, 0, 0, 1)t$.

Den lineære transformation Den lineære transformation kaldes Y^{-1} og består af egenvektorerne i rækkerne:

$$Y^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ D_s & h_1 & h_2 & 0 \\ C_s & 1 & 1 & 0 \\ H_s & h_1 C_o - D_o & h_2 C_o - D_o & 1 \end{pmatrix}$$

Y^{-1} kan inverteres ved at opskrive ligningen $Y^{-1}Y = I$:

$$\begin{aligned} & \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ D_s & h_1 & h_2 & 0 & 0 & 1 & 0 & 0 \\ C_s & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ H_s & h_1 C_o - D_o & h_2 C_o - D_o & 1 & 0 & 0 & 0 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & h_1 & h_2 & 0 & -D_s & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -C_s & 0 & 1 & 0 \\ 0 & h_1 C_o - D_o & h_2 C_o - D_o & 1 & -H_s & 0 & 0 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & h_1 - h_2 & 0 & 0 & h_2 C_s - D_s & 1 & -h_2 & 0 \\ 0 & 1 & 1 & 0 & -C_s & 0 & 1 & 0 \\ 0 & h_1 C_o - D_o & h_2 C_o - D_o & 1 & -H_s & 0 & 0 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{h_2 C_s - D_s}{h_1 - h_2} & \frac{1}{h_1 - h_2} & \frac{-h_2}{h_1 - h_2} & 0 \\ 0 & 1 & 1 & 0 & -C_s & 0 & 1 & 0 \\ 0 & h_1 C_o - D_o & h_2 C_o - D_o & 1 & -H_s & 0 & 0 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{h_2 C_s - D_s}{h_1 - h_2} & \frac{1}{h_1 - h_2} & \frac{-h_2}{h_1 - h_2} & 0 \\ 0 & 0 & 1 & 0 & \frac{-h_1 C_s + D_s}{h_1 - h_2} & \frac{-1}{h_1 - h_2} & \frac{h_1}{h_1 - h_2} & 0 \\ 0 & h_1 C_o - D_o & h_2 C_o - D_o & 1 & -H_s & 0 & 0 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{h_2 C_s - D_s}{h_1 - h_2} & \frac{1}{h_1 - h_2} & \frac{-h_2}{h_1 - h_2} & 0 \\ 0 & 0 & 1 & 0 & \frac{-h_1 C_s + D_s}{h_1 - h_2} & \frac{-1}{h_1 - h_2} & \frac{h_1}{h_1 - h_2} & 0 \\ 0 & 0 & 0 & 1 & -H_0 & -C_0 & D_0 & 1 \end{array} \right) \end{aligned}$$

Dvs. at Y er givet ved:

$$Y = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{h_2 C_s - D_s}{h_1 - h_2} & \frac{1}{h_1 - h_2} & \frac{-h_2}{h_1 - h_2} & 0 \\ \frac{-h_1 C_s + D_s}{h_1 - h_2} & \frac{-1}{h_1 - h_2} & \frac{h_1}{h_1 - h_2} & 0 \\ -H_0 & -C_0 & D_0 & 1 \end{pmatrix}$$

Løsning af differentiallygningsystemet Løsningen udtrykt vha. \mathbf{F} er givet ved følgende:

$$F_1 = \delta_1 \exp(kx)$$

$$F_2 = \delta_2 \exp(mx)$$

$$F_3 = \delta_3 \exp(-mx)$$

$$F_4 = \delta_4 \exp(-Kx)$$

Hvor $\delta_1, \dots, \delta_4$ er konstanter, der afhænger af randbetingelserne.

Løsningen \mathbf{E} findes ved at anvende Y^{-1} på \mathbf{F} :

$$E_s = \delta_1 \exp(kx)$$

$$E_- = D_s \delta_1 \exp(kx) + h_1 \delta_2 \exp(mx) + h_2 \delta_3 \exp(-mx)$$

$$E_+ = C_s \delta_1 \exp(kx) + \delta_2 \exp(mx) + \delta_3 \exp(-mx)$$

$$E_o = H_s \delta_1 \exp(kx) + (h_1 C_o - D_o) \delta_2 \exp(mx) + (h_2 C_o - D_o) \delta_3 \exp(-mx) + \delta_4 \exp(-Kx)$$

Bestemmelse af de øvrige egenvektorer

I dette afsnit bestemmes egenvektorerne hørende til de tre egenverdier λ_2 , λ_3 og λ_4 .

For $\lambda_2 = m$ ser ligningssystemet således ud:

$$\begin{aligned} & \left(\begin{array}{cccc|c} k-m & 0 & 0 & 0 & 0 \\ -s & a-m & -\sigma & 0 & 0 \\ s' & \sigma & -a-m & 0 & 0 \\ w & v & u & -K-m & 0 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & a-m & -\sigma & 0 & 0 \\ 0 & \sigma & -(a+m) & 0 & 0 \\ 0 & v & u & -(K+m) & 0 \end{array} \right) \\ & \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & \sigma(a-m) & -\sigma^2 & 0 & 0 \\ 0 & \sigma(a-m) & -(a^2-m^2) & 0 & 0 \\ 0 & v & u & -(K+m) & 0 \end{array} \right) \end{aligned}$$

Det udnyttes, at $a^2 - m^2 = \sigma^2$, og række 3 kan dermed reduceres til en nulrække:

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & a-m & -\sigma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & v & u & -(K+m) & 0 \end{array} \right)$$

Ligningerne kan skrives op:

$$\begin{aligned} x_1 &= 0 \\ (a - m)x_2 - \sigma x_3 &= 0 \end{aligned} \tag{A.5}$$

$$vx_2 + ux_3 - (K + m)x_4 = 0 \tag{A.6}$$

$$\tag{A.7}$$

x_3 kan vælges frit og sættes til t og x_2 isoleres i ligning A.5:

$$(a - m)x_2 - \sigma t = 0 \Leftrightarrow x_2 = \frac{\sigma}{a - m}t = h_1 t$$

De to udtryk sættes ind i ligning A.6 og x_4 isoleres:

$$vh_1 t + ut - (K + m)x_4 = 0$$

$$(K + m)x_4 = (vh_1 + u)t$$

$$\begin{aligned} x_4 &= \frac{vh_1 + u}{K + m}t \\ &= \left(h_1 \frac{v(K - m)}{K^2 - m^2} + \frac{u(K - m)}{K^2 - m^2} \right)t \\ &= \left(h_1 \frac{vK - va + u\sigma - vm + va - u\sigma}{K^2 - m^2} + \frac{uK - um}{K^2 - m^2} \right)t \\ &= \left(h_1 \frac{v(K - a) + u\sigma}{K^2 - m^2} + h_1 \frac{-vm + va - u\sigma}{K - m^2} + \frac{uK - um}{K^2 - m^2} \right)t \\ &= \left(h_1 C_o + \frac{a + m}{\sigma} \frac{-vm + va - u\sigma}{K^2 - m^2} + \frac{\sigma(uK - um)}{\sigma(K^2 - m^2)} \right)t \\ &= \left(h_1 C_o + \frac{-vam + va^2 + au\sigma - vm^2 + vam + um\sigma + uK\sigma - um\sigma}{\sigma(K^2 - m^2)} \right)t \\ &= \left(h_1 C_o + \frac{u\sigma(K + a) + v(a^2 - m^2)}{\sigma(K^2 - m^2)} \right)t \\ &= \left(h_1 C_o + \frac{u\sigma(K + a) + v\sigma^2}{\sigma(K^2 - m^2)} \right)t \\ &= \left(h_1 C_o + \frac{u(K + a) + v\sigma}{K^2 - m^2} \right)t = (h_1 C_o - D_o)t \end{aligned}$$

Egenvektoren v_2 knyttet til egenværdien $\lambda_2 = m$ er givet ved $v_2 = (0, h_1, 1, h_1 C_o - D_o)t$.

For $\lambda_3 = -m$ er udledningen helt parallel med den netop udførte for $\lambda_2 = m$ og $v_3 = (0, h_2, 1, h_2 C_o - D_o)t$.

For $\lambda_4 = -K$ ser ligningssystemet således ud:

$$\left(\begin{array}{cccc|c} K + k & 0 & 0 & 0 & 0 \\ -s & K + a & -\sigma & 0 & 0 \\ s' & \sigma & K - a & 0 & 0 \\ w & v & u & 0 & 0 \end{array} \right)$$

Række 1 fratrækkes rækkerne 2, 3 og 4.

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & K+a & -\sigma & 0 & 0 \\ 0 & \sigma & K-a & 0 & 0 \\ 0 & v & u & 0 & 0 \end{array} \right)$$

Række 2 multipliceres med σ og fratrækkes række 3 multipliceret med $(K+a)$:

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & K+a & -\sigma & 0 & 0 \\ 0 & 0 & K^2 - a^2 - \sigma^2 & 0 & 0 \\ 0 & v & u & 0 & 0 \end{array} \right)$$

Den reducerede række 3 multipliceret med hhv. σ og u kan nu fratrækkes rækkerne 2 og 4:

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & K+a & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & v & 0 & 0 & 0 \end{array} \right)$$

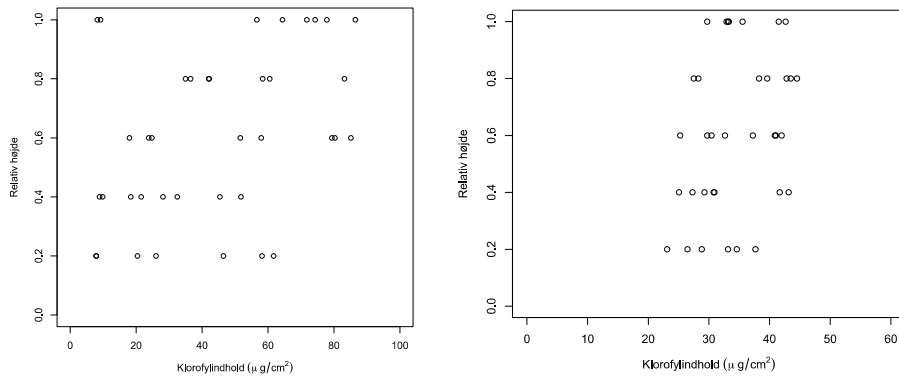
Endelig indses det, at række 4 reduceres til en nulrække:

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

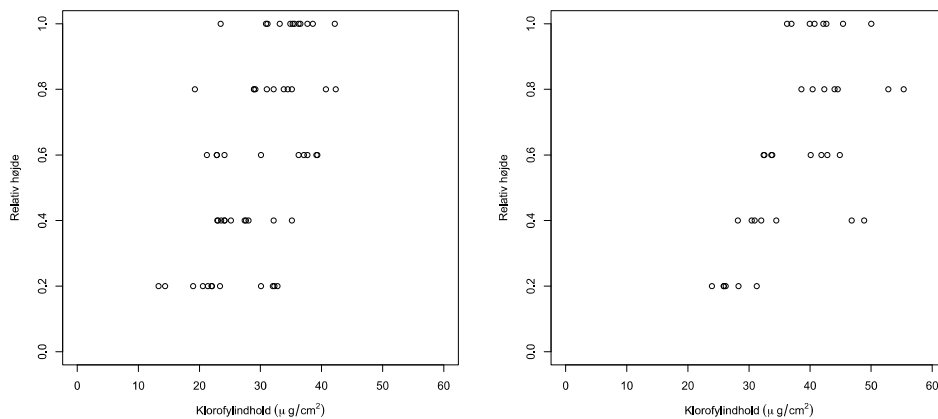
Den tilhørende egenvektor er dermed på formen $v_4 = (0, 0, 0, 1)t$.

B. Vertikale klorofylprofiler

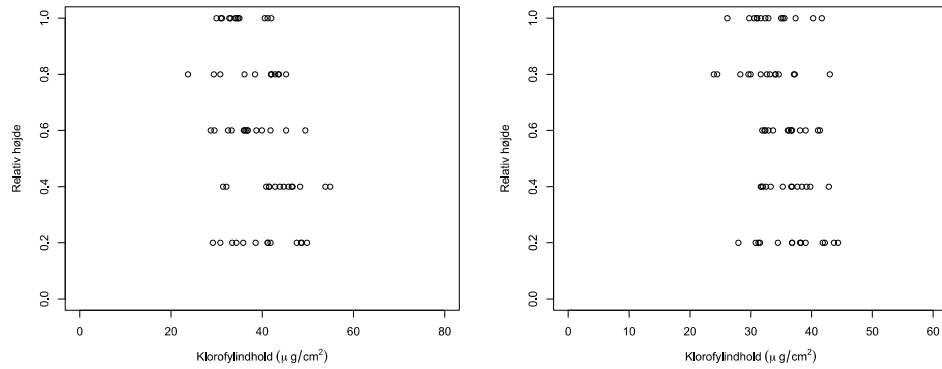
Formålet med dette afsnit er at præsentere plots af data for forskellige typer af vertikale klorofylprofiler.



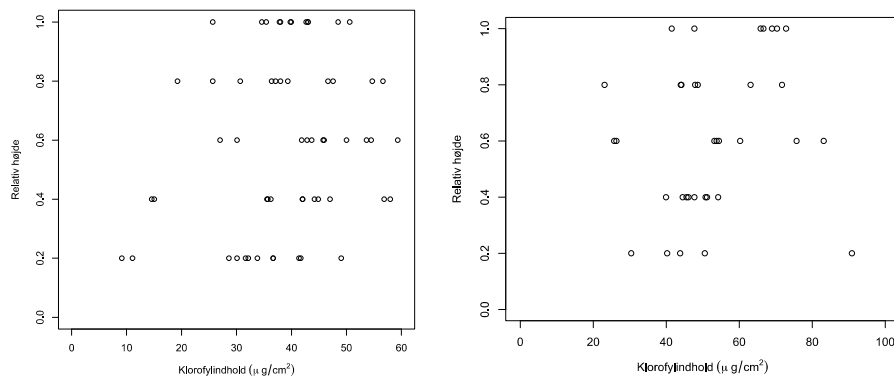
Figur B.1: Figuren viser to plots af marker med et uniformt vertikalt klorofylprofil hhv. nr. 7 og 9 fra Danmark (2008).



Figur B.2: Figuren viser to plots af marker med et voksende lineært vertikalt klorofylprofil hhv. nr. 1 og 2 fra Danmark (2008).



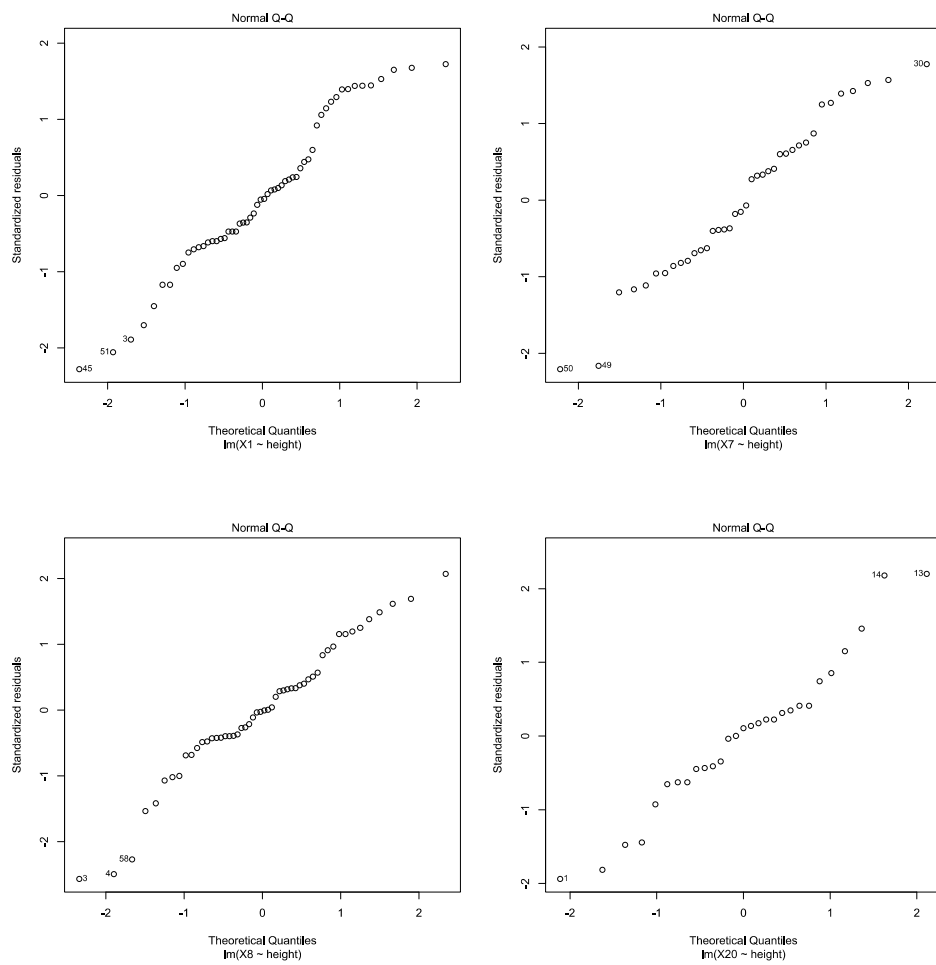
Figur B.3: Figuren viser to plots af marker med et aftagende lineært vertikalt klorofylprofil hhv. nr. 8 og 16 fra Danmark (2009).



Figur B.4: Figuren til venstre viser et plot af en mark med hverken uniformt eller lineært vertikalt klorofylprofil, mens figuren til højre viser et plot af en mark, hvor data for klorofylindholdet ikke har varianshomogenitet, hhv. nr. 19 og 3 fra Danmark (2008).

C. Fraktildiagrammer

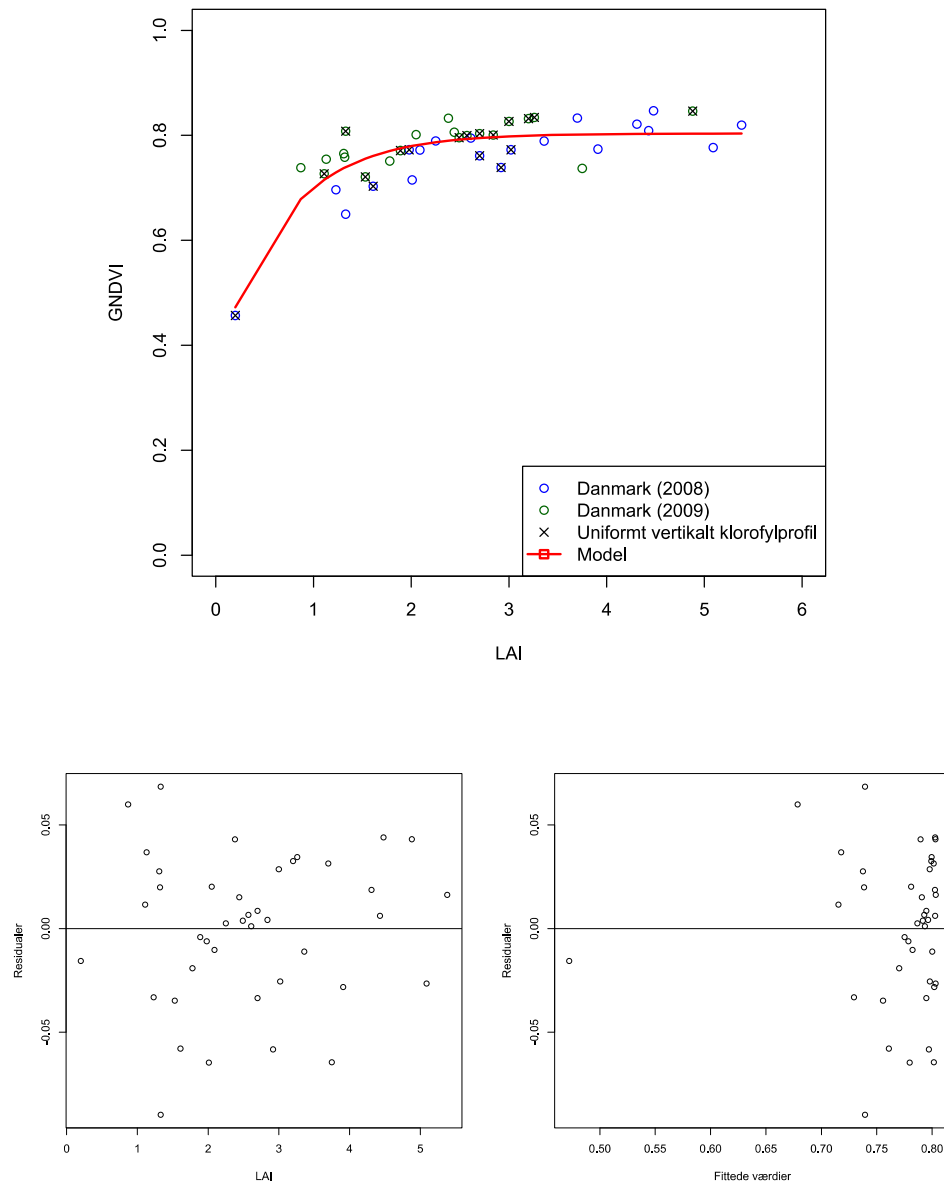
Her vises eksempler på fraktildiagrammer fra et udvalg af marker hvis vertikale klorofylprofil er testet i afsnit 5.2 på side 79.



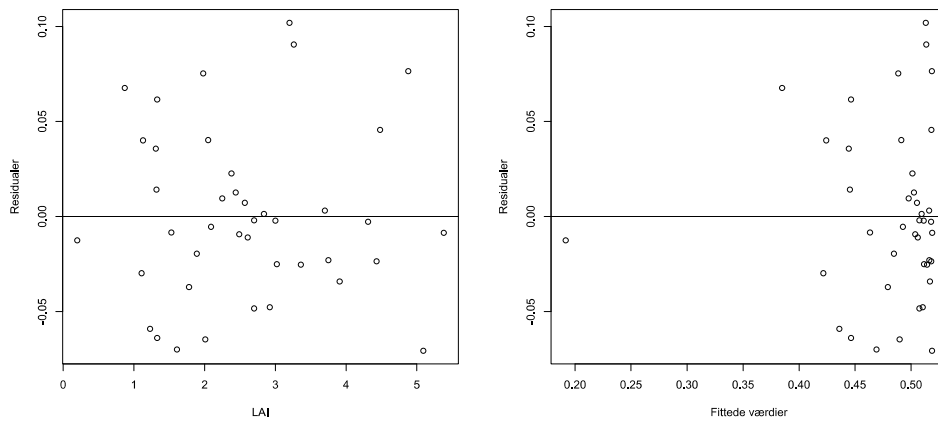
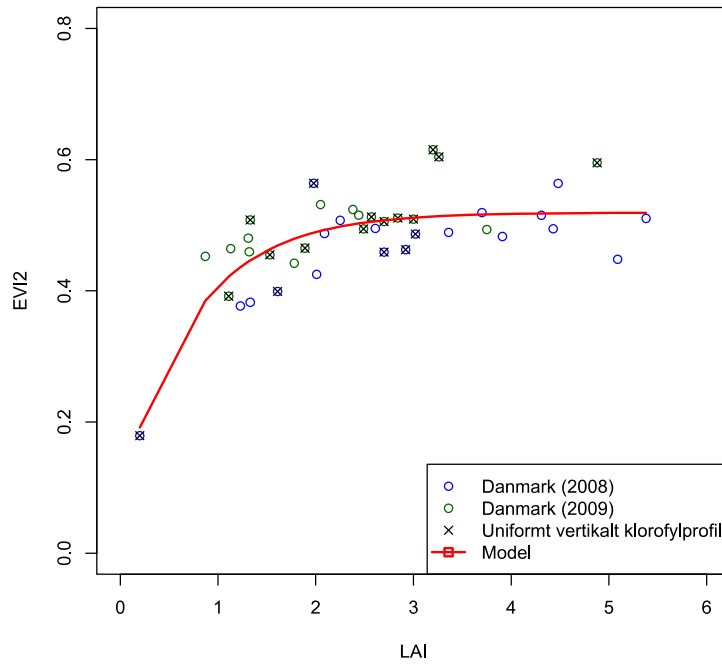
Figur C.1: Figuren viser fraktildiagrammer fra fire marker fra Danmark (2008).

D. GNDVI og EVI2

I dette kapitel præsenteres den empiriske sammenhæng mellem LAI og henholdsvis GNDVI og EVI2 for de danske landskaber.



Figur D.1: Figuren viser sammenhængen mellem GNDVI og LAI for de danske landskaber. $GNDVI_{\max}$ estimeres til 0.83, $GNDVI_{\min}$ til 0.35 og udslningskoefficienten er 0.67.



Figur D.2: Figuren viser sammenhængen mellem EVI2 og LAI for de danske landskaber. $EVI2_{max}$ estimeres til 0.55, $EVI2_{min}$ til 0.09 (ikke signifikant) og udslningskoefficienten er 0.64.

E. Resultat af simuleringer

Her præsenteres to tabeller, der viser de relative afvigelser mellem simuleret og målt reflektans for henholdsvis Danmark (2008) og Danmark (2009).

	Højde (m)	LAI	$\rho_{\text{grøn}}$	$\rho_{\text{rød}}$	ρ_{NIR}	SR	NDVI	GNDVI	EV12	CI
Marker med uniformt vertikalt klorofylprofil										
Majs (nr. 5)	0.10	0.20	16,97%	17,15%	17,18%	0,02%	0,03%	0,15%	12,39%	0,28%
Byg (nr. 7)	0.75	2.92	-53,35%	-65,15%	-4,09%	175,19%	20,00%	16,92%	9,95%	124,28%
Byg (nr. 9)	0.40	1.98	-33,62%	-41,30%	-30,92%	17,70%	3,55%	1,03%	-20,55%	4,67%
Byg (nr. 12)	0.95	3.02	-48,41%	-60,76%	4,28%	165,73%	15,58%	14,00%	13,60%	117,16%
Byg (nr. 14)	0.30	2.70	-29,41%	-51,25%	-1,39%	102,29%	14,11%	8,13%	8,30%	45,94%
Hvede (nr. 17)	0.25	1.61	-39,24%	-38,97%	6,30%	74,18%	15,95%	16,46%	15,70%	90,78%
Marker med lineært vertikalt klorofylprofil										
Hvede (nr. 1)	0.27	1.23	5,34%	0,64%	-13,12%	-13,67%	-6,18%	-7,61%	-13,79%	-21,35%
Hvede (nr. 2)	0.70	3.70	-24,08%	-52,10%	-2,55%	103,46%	9,06%	4,14%	4,11%	31,20%
Byg (nr. 6)	0.30	1.33	-31,17%	-41,59%	-16,39%	43,13%	16,58%	8,11%	-1,69%	27,26%
Byg (nr. 10)	0.95	5.09	-32,62%	-65,20%	24,90%	258,88%	20,09%	12,41%	32,23%	97,65%
Hvede (nr. 13)	0.80	3.36	-21,24%	-52,86%	-0,86%	110,33%	12,55%	5,03%	7,70%	29,35%
Hvede (nr. 15)	0.60	2.61	-20,03%	-44,66%	-5,82%	70,18%	9,45%	3,55%	1,94%	20,06%
Hvede (nr. 16)	0.75	4.48	-21,04%	-49,86%	5,49%	110,38%	8,04%	4,29%	9,28%	36,65%
Hvede (nr. 18)	0.70	4.31	-46,24%	-63,47%	3,23%	182,59%	12,23%	9,93%	10,49%	102,05%
Byg (nr. 20)	0.70	3.91	-36,45%	-60,92%	-7,30%	137,23%	14,74%	8,46%	3,99%	52,59%
Marker med forskellige middelværdier i det vertikale klorofylprofil										
Hvede (nr. 8)	0.65	5.38	-26,89%	-55,76%	21,87%	175,44%	12,25%	8,32%	23,74%	74,05%
Byg (nr. 19)	0.35	2.01	-36,38%	-47,35%	-0,49%	89,00%	16,54%	12,99%	10,70%	67,67%
Marker uden varianshomogenitet i klorofylindholdet mellem relative højder										
Hvede (nr. 3)	0.35	2.09	-43,37%	-46,59%	-9,96%	68,59%	10,50%	10,14%	-0,59%	67,70%
Hvede (nr. 4)	0.50	2.25	-24,61%	-30,98%	-20,69%	14,90%	2,75%	1,18%	-13,66%	5,88%
Hvede (nr. 11)	0.75	4.43	2,52%	-45,74%	14,80%	111,56%	11,17%	2,31%	18,20%	13,40%

Tabel E.1: Tabellen viser de relative afvigelser mellem de simulerede reflektanser og vegetationsindekser og de målte for markerne i Danmark (2008).

	Højde (m)	LAI	$\rho_{\text{grøn}}$	$\rho_{\text{rød}}$	ρ_{NIR}	SR	NDVI	GNDVI	EVI2	CI
Marker med uniformt vertikalt klorofylprofil										
Hvede (nr. 2)	0.26	2.57	-44,34%	-54,67%	-9,34%	100,00%	10,35%	9,06%	-0,26%	70,76%
Hvede (nr. 7)	0.18	1.33	35,82%	44,80%	-20,76%	-45,28%	-13,45%	-14,37%	-23,32%	-46,61%
Hvede (nr. 9)	0.22	1.53	-10,14%	-20,96%	-15,29%	7,18%	2,02%	-2,01%	-9,91%	-6,84%
Raps (nr. 10)	1.04	3.20	-65,97%	-63,71%	-12,38%	141,45%	9,40%	11,95%	-2,53%	173,39%
Raps (nr. 11)	1.05	3.26	-61,84%	-61,26%	-9,05%	134,81%	8,95%	11,13%	-0,43%	152,14%
Hvede (nr. 13)	0.38	4.88	-30,96%	-54,05%	-6,32%	103,88%	6,83%	4,51%	0,10%	38,93%
Hvede (nr. 14)	0.23	1.11	-2,91%	-1,35%	-5,45%	-4,16%	-1,46%	-0,87%	-4,99%	-3,10%
Hvede (nr. 17)	0.28	2.49	-37,96%	-53,41%	-6,90%	99,84%	11,27%	7,98%	2,18%	56,50%
Hvede (nr. 18)	0.21	1.89	-20,05%	-35,94%	-9,38%	41,46%	7,61%	3,14%	-2,03%	15,34%
Hvede (nr. 19)	0.27	3.00	-27,33%	-49,33%	-1,25%	94,90%	8,54%	5,19%	4,71%	39,65%
Hvede (nr. 21)	0.26	2.70	-25,55%	-49,69%	-6,37%	86,12%	9,71%	4,62%	1,65%	28,92%
Hvede (nr. 22)	0.31	2.84	-34,56%	-55,15%	-6,38%	108,74%	11,01%	6,96%	2,47%	48,43%
Marker med lineært vertikalt klorofylprofil										
Hvede (nr. 8)	0.23	1.78	-22,41%	-33,48%	-5,97%	41,35%	8,22%	5,19%	0,94%	24,70%
Hvede (nr. 16)	0.18	1.31	-2,56%	-6,79%	-21,08%	-15,33%	-4,57%	-6,19%	-18,22%	-21,92%
Marker med forskellige middelværdier i det vertikale klorofylprofil										
Hvede (nr. 15)	0.32	3.75	-50,81%	-72,34%	-1,92%	254,57%	22,86%	16,53%	13,73%	117,11%
Marker uden varianshomogenitet i klorofylindholdet mellem relative højder										
Hvede (nr. 1)	0.24	2.05	-10,06%	-19,48%	-17,52%	2,43%	0,41%	-2,00%	-12,45%	-9,33%
Hvede (nr. 3)	0.25	0.87	3,24%	19,74%	-22,66%	-35,41%	-15,24%	-9,88%	-26,01%	-29,54%
Hvede (nr. 4)	0.26	2.44	-16,41%	-34,02%	-9,46%	37,22%	4,87%	1,68%	-3,74%	9,31%
Hvede (nr. 5)	0.20	1.13	5,90%	17,49%	-18,93%	-31,00%	-10,73%	-8,42%	-20,53%	-27,26%
Hvede (nr. 12)	0.15	1.32	-0,84%	-6,00%	-16,65%	-11,33%	-3,37%	-5,19%	-14,31%	-18,49%
Hvede (nr. 20)	0.26	2.38	-13,57%	-31,74%	-10,84%	30,62%	3,79%	0,57%	-5,46%	3,48%

Tabel E.2: Tabellen viser de relative afvigelser mellem de simulerede reflektanser og vegetationsindekser og de målte for markerne i Danmark (2009).

Litteratur

- Allen, W. A., Gayle, T. V. og Richardson, A. J. [1970], 'Plant-canopy irradiance specified by the Duntley equations', *Journal of the Optical Society of America* **60**(3), 372–376.
- Baret, F. og Buis, S. [2008], Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems, i S. Liang, red., 'Advances in Land Remote Sensing: System, modeling, inversion and application', Springer, kapitel 7, side 173–201.
- Baret, F. og Fourty, T. [1997], Radiometric estimates of nitrogen status of leaves and canopies, i G. Lemaire, red., 'Diagnosis of the nitrogen status in crops', Springer, kapitel 12, side 201–227.
- Baret, F. og Guyot, G. [1991], 'Potentials and limits of vegetation indices for LAI and APAR assessment', *Remote Sensing of Environment* **35**, 161–173.
- Bartlett, M. S. [1937], 'Properties of sufficiency and statistical tests', *Proceedings of the Royal Society of London* .
- Beadle, C. L. [1993], Growth analysis, i D. O. Hall, J. M. O. Shurlock, H. R. Bolhar-Nordenkamp, R. C. Leegood og S. P. Long, red., 'Photosynthesis and production in a changing environment - a field and laboratory manual', Chapman & Hall, kapitel 3, side 36–46.
- Bird, A. C. [1991], Principles of remote sensing: Electromagnetic radiation, reflectance and emissivity, i A. S. Belward og C. R. Valenzuela, red., 'Remote Sensing and Geographical Information Systems for Resource Management in Developing Countries', nr. 1 i 'Eurocourses', Kluwer, Dordrecht, kapitel 1, side 1–17.
- Boegh, E., Houborg, R., Bienkowski, J., Braban, C. F., Dalgaard, T., van Dijk, N., Dragosits, U., Holmes, E., Magliulo, V., Schelde, K., Tommasi, P. D., Vitale, L., Theobald, M. R., Cellier, P. og Sutton, M. A. [2013], 'Remote sensing of LAI, chlorophyll and leaf nitrogen pools of crop- and grasslands in five european landscapes', *Biogeosciences* **10**, 6279–6307.
- Boegh, E., Soegaard, H., Broge, N., Hasager, C., Jensen, N., Schelde, K. og Thomsen, A. [2002], 'Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture', *Remote Sensing of Environment* **81**, 179–193.

- Brady, N. C. og Weil, R. R. [2008], *The nature and properties of soils*, 14 udgave, Harlow : Prentice Hall.
- Broge, N. H. [2003], Prediction of green canopy area index and canopy chlorophyll density of homogenous canopies from measurements of spectral reflectance in the visible and near-visible domain, PhD thesis, Danish Institute of Agricultural Sciences, Department of Agricultural Systems og University of Copenhagen, Institute of Geography.
- Campbell, G. S. [1986], 'Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution', *Agricultural and Forest Meteorology* **36**, 317–321.
- Campbell, G. S. [1990], 'Derivation of an angle density function for canopies with ellipsoidal leaf angle distributions', *Agricultural and Forest Meteorology* **49**, 173–176.
- Casa, R., Baret, F., Buis, S., Lopez-Lozano, R., Pascucci, S., Palombo, A. og Jones, H. G. [2010], 'Estimation of maize canopy properties from remote sensing by inversion of 1-D and 4-D models', *Precision Agric* **11**, 319–334.
- Ciganda, V. S., Gitelson, A. A. og Schepers, J. [2012], 'How deep does a remote sensor sense? expression of chlorophyll content in a maize canopy', *Remote Sensing of Environment* **126**, 240–247.
- Danmarks Meteorologiske Institut [2014], 'Vejrarkiv', Website. <http://www.dmi.dk/vejr/arkiver/vejrarkiv>.
- Danmarks Miljøundersøgelser [2000], *AIS - Areal Informations Systemet*, Miljø- og Energiministeriet.
- de Wit, C. T. [1965], *Photosynthesis of leaf canopies*, Centre for Agricultural Publications and Documentation, Wageningen.
- Den Store Danske [2014a], 'Blad', http://www.denstoredanske.dk/Natur_og_milj\T1\o/Botanik/Plantemorfologi/blad.
- Den Store Danske [2014b], 'Lambert-Beers lov', <http://www.denstoredanske.dk/index.php?sideId=114015>.
- Den danske jordklassificering* [n.d.]. DJF Geodata.
- Dudley, R. M. [2002], *Real analysis and probability*, nr. 74 i 'Cambridge studies in advanced mathematics', 2 udgave, Cambridge University Press.
- Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidet, L. P., Ustin, S. L., le Maire, G. og Jacquemoud, S. [2008], 'PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments', *Remote Sensing of Environment* **112**, 3030–3043.

- Gitelson, A. A., Gritz, Y. og Merzlyak, M. N. [2003], 'Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves', *Journal of Plant Physiology* **160**, 271–282.
- Gitelson, A. A., Kaufman, Y. J. og Merzlyak, M. N. [1996], 'Use of a green channel in remote sensing of global vegetation from EOS-MODIS', *Remote Sensing of Environment* **58**, 289–298.
- Gitelson, A. A., na, A. V., Ciganda, V., Rundquist, D. C. og Arkebauer, T. J. [2005], 'Remote estimation of canopy chlorophyll content in crops', *Geophysical Research Letters* **32**. L08403.
- Goel, N. S. [1988], 'Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data', *Remote Sensing Reviews* **4**, 1–212.
- Hoffer, R. M. [1978], Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data, i P. H. Swain og S. M. Davis, red., 'Remote sensing – the quantitative approach', McGraw-Hill, New York, kapitel 5, side 227–289.
- Houborg, R. og Anderson, M. C. [2009], 'Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at regional scales', *Journal of Applied Remote Sensing* **3**.
- Houborg, R., Soegaard, H. og Boegh, E. [2007], 'Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data', *Remote Sensing of Environment* **106**, 39–58.
- Huang, W., Wang, Z., Huang, L., Lamb, D. W., Ma, Z., Zhang, J., Wang, J. og Zhao, C. [2011], 'Estimation of vertical distribution of chlorophyll concentration by bi-directional canopy reflectance spectra in winter wheat', *Precision Agriculture* **12**, 165–178.
- Huete, A., Justice, C. og Liu, H. [1994], 'Development of vegetation and soil indices for MODIS-EOS', *Remote Sensing of Environment* **49**, 224–234.
- Huete, A. R. [1988], 'A soil-adjusted vegetation index (SAVI)', *Remote Sensing of Environment* **25**, 295–309.
- Huete, A. R., Didan, K. og van Leeuwen, W. [1999], MODIS vegetation index (MOD13), Algorithm theoretical basis document, Technical report, University of Arizona. Version 3.1.

- Huete, A. R., Liu, H. Q., Batchily, K. og van Leeuwen, W. [1997], 'A comparison of vegetation indices over a global set of TM images for EOS-MODIS', *Remote Sensing of Environment* **59**, 440–451.
- International Center for Remote Sensing Education* [2011], http://www.r-s-c-c.org/sites/default/files/film_spectralcurves.jpg.
- Jacquemoud, S., Féret, J.-B., François, C. og Broadbent, E. [2013], 'PROSPECT + SAIL = PROSAIL', <http://teledetection.ipgp.jussieu.fr/prosail/>.
- Jacquemoud, S. og Baret, F. [1990], 'PROSPECT: A model of leaf optical properties spectra', *Remote Sensing of Environment* **34**, 75–91.
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C. og Ustin, S. L. [2009], 'PROSPECT+SAIL models: A review of use for vegetation characterization', *Remote Sensing of Environment* **113**, S56–S66.
- Jiang, Z., Huete, A. R., Didan, K. og Miura, T. [2008], 'Development of a two-band enhanced vegetation index without a blue band', *Remote Sensing of Environment* **112**, 3833–3845.
- LAI-2000* [n.d.], http://envsupport.licor.com/images/env/product_list_photos/LAI-2000_lg.jpg.
- Larsen, J. [2006a], *Basisstatistik*, nr. 435 i 'IMFUFA tekster', 2 udgave, Roskilde Universitetscenter.
- Larsen, J. [2006b], *Sandsynlighedsregning & statistik for matematikstuderende*, nr. 452 i 'IMFUFA tekster', Roskilde Universitetscenter.
- LiCor [1992], *LAI-2000 Plant Canopy Analyzer*.
- MATLAB[®] [2013], *version 8.2.0.701 (R2013b)*, The MathWorks Inc., Natick, Massachusetts.
- Nilson, T. [1971], 'A theoretical analysis of the frequency of gaps in plant stands', *Agricultural Meteorology* **8**, 25–38.
- QGIS Development Team [2013], *QGIS Geographic Information System version 2.0.1*, Open Source Geospatial Foundation.
URL: <http://qgis.osgeo.org>
- R Development Core Team [2013], *R: A Language and Environment for Statistical Computing version 3.1.0*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <http://www.R-project.org>

- Richards, J. A. og Jia, X. [2006], *Remote Sensing Digital Image Analysis – an Introduction*, 4 udgave, Springer, kapitel 1, side 1–9.
- Rosenberg, N. J. [1974], *Microclimate: The Biological Environment*, John Wiley & Sons.
- Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S. og Martonchik, J. V. [2006], 'Reflectance quantities in optical remote sensing – definitions and case studies', *Remote Sensing of Environment* **103**, 27–42.
- Spectrum Technologies [n.d.a], 'SPAD - 502 Plus Chlorophyll Meter', <http://www.specmeters.com/nutrient-management/chlorophyll-meters/chlorophyll/spad502p/>.
- Spectrum Technologies [n.d.b], *SPAD 502 Plus Chlorophyll Meter - product manual*.
- Suits, G. H. [1972], 'The calculation of the directional reflectance of a vegetative canopy', *Remote Sensing of Environment* **2**, 117–125.
- Tjur, T. [2002a], *Forelæsningsnoter i sandsynlighedsregning til brug ved undervisningen på MØK-studiets 2. år*, 5 udgave, Statistikgruppen, Handelshøjskolen i København.
- Tjur, T. [2002b], *Forelæsningsnoter i statistik til brug ved undervisningen på MØK-studiets 2. år*, 3 udgave, Statistikgruppen, Handelshøjskolen i København.
- Tucker, C. J. og Sellers, P. J. [1986], 'Satellite remote sensing of primary production', *International Journal of Remote Sensing* **7**(11), 1395–1416.
- Verhoef, W. [1984], 'Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model', *Remote Sensing of Environment* **16**, 125–141.
- Verhoef, W. [1985], 'Earth observation modeling based on layer scattering matrices', *Remote Sensing of Environment* **17**, 165–178.
- Verhoef, W. [1997], *Theory of radiative transfer models applied in optical remote sensing of vegetation canopies*, PhD thesis, Wageningen Agricultural University.
- Verhoef, W. og Bach, H. [2003], 'Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models', *Remote Sensing of Environment* **87**, 23–41.
- Viña, A., Gitelson, A. A., Nguy-Robertson, A. L. og Peng, Y. [2011], 'Comparison of different vegetation indices for the remote assessment of green leaf area index of crops', *Remote Sensing of Environment* .

- Wang, J., Wang, Z., Uchida, S. og Huang, W. [2009], 'Relative discrimination of planophile and erectophile wheat types using multi-temporal spectrum measurements', *Japan Agricultural Research Quarterly* **43**(2), 157—166.
- Wang, W.-M., Li, Z.-L. og Su, H.-B. [2007], 'Comparison of leaf angle distribution functions: Effects on extinction coefficient and fraction of sunlit foliage', *Agricultural and Forest Meteorology* **143**, 106–122.
- Welles, J. M. og Norman, J. M. [1991], 'Instrument for indirect measure of canopy architecture', *Agronomy Journal* **83**, 818–825.
- Wikipedia [2014], 'Spherical coordinate system', http://en.wikipedia.org/wiki/File:3D_Spherical.svg.
- Winterhalter, L., Mistele, B. og Schmidhalter, U. [2012], 'Assessing the vertical footprint of reflectance measurements to characterize nitrogen uptake and biomass distribution in maize canopies', *Field Crops Research* **129**, 14–20.