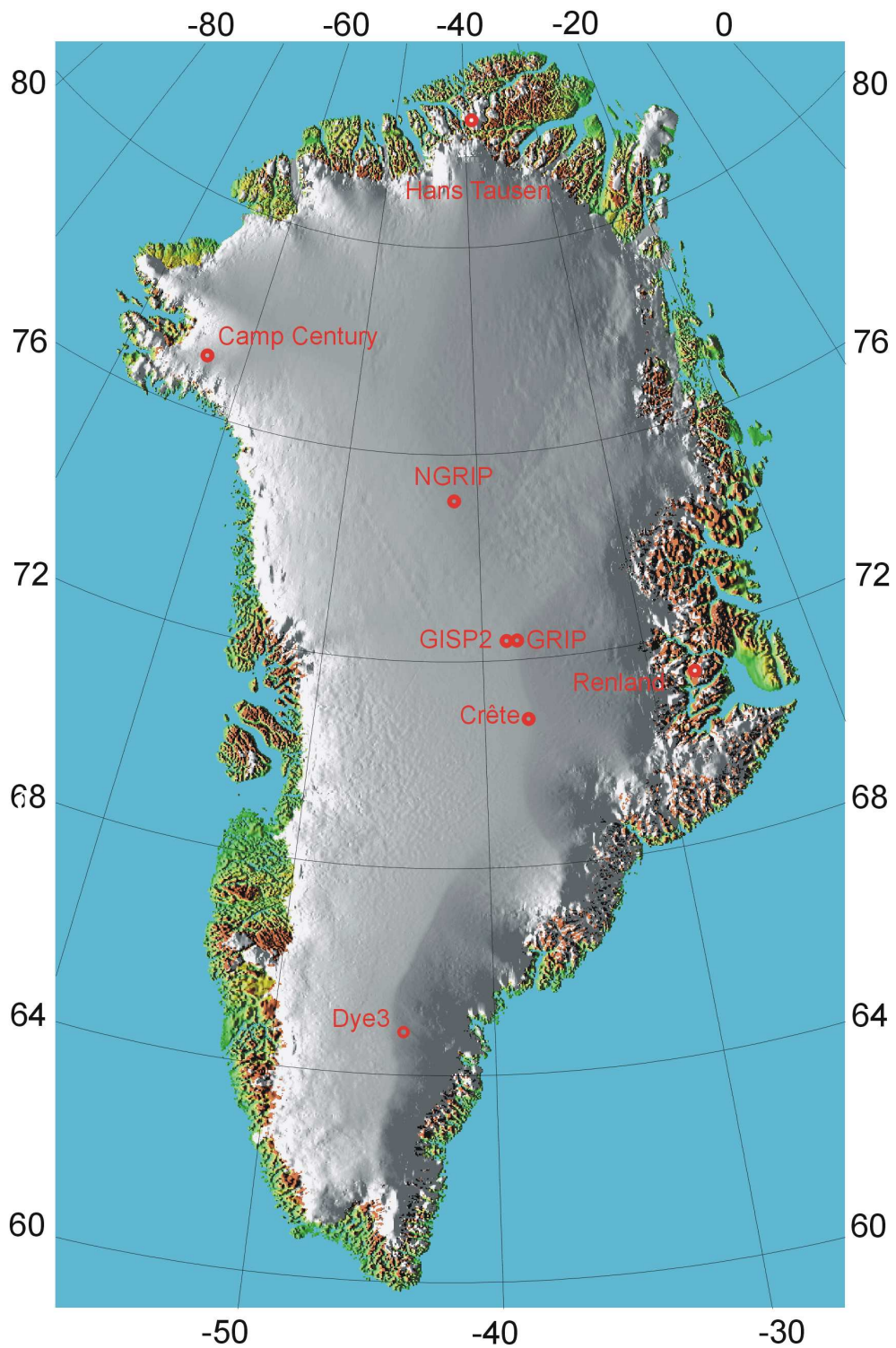


# Indhold

<b>Indledning</b>	<b>2</b>
<b>1 Datering af iskerner</b>	<b>7</b>
1.1 Iskerneforskningens fødsel . . . . .	7
1.2 Datering af is ved modellering af isens flydning . . . . .	10
1.2.1 Nye-modellen . . . . .	10
1.2.2 Dansgaard-Johnsen-modellen . . . . .	11
1.2.3 Mere avancerede modeller . . . . .	12
1.3 Datering af is med stratigrafiske metoder . . . . .	14
1.3.1 Datering ved hjælp af reference-horisonter . . . . .	14
1.3.2 Datering ved identifikation af årlag . . . . .	15
1.3.3 Dateringen af GISP2-iskernen . . . . .	19
1.3.4 Behovet for en ny stratigrafisk tællemetode . . . . .	22
<b>2 Dynamisk Dekorrelation af iskernedata</b>	<b>23</b>
2.1 Behandling af dybdeserier som tidsserier . . . . .	24
2.2 Indledende formalisme . . . . .	26
2.3 Principalkomponent Analyse - PCA . . . . .	26
2.4 Dynamisk Dekorrelation . . . . .	28
2.4.1 Dynamisk Dekorrelation for uendelige serier . . . . .	28
2.4.2 Permutationer og skalering i egenværdi-opløsningen af $\mathbf{Q}$ . . . . .	31
2.4.3 Dynamisk Dekorrelation for endelige serier . . . . .	32
2.4.4 Anvendelse af flere samtidige værdier af $\tau$ . . . . .	34
<b>3 Data</b>	<b>37</b>
3.1 Valg af datasæt . . . . .	37
3.2 Manglende datapunkter . . . . .	41
3.3 Regularisering af data . . . . .	42
3.4 Fordelingsfunktioner for data . . . . .	44
3.4.1 Fordelingen af ECM-data fra Site A og D . . . . .	46

<b>4</b>	<b>Dynamisk Dekorrelation i praksis</b>	<b>51</b>
4.1	Synkronisering af årlige toppe . . . . .	51
4.2	Valg af $\tau$ -værdi . . . . .	52
4.2.1	Hvilke $\tau$ virker bedst ? . . . . .	53
4.2.2	Valg af $\tau$ ved anvendelse af formodninger om kildeseriernes natur . . . . .	54
4.2.3	Valg af $\tau$ ved maksimering af separationen af $c_S(\tau)$ -værdierne . . . . .	55
4.3	Dynamisk Dekorrelation : trin for trin . . . . .	56
<b>5</b>	<b>Resultater og diskussion</b>	<b>57</b>
5.1	Et simpelt mål for tælbarhed . . . . .	57
5.1.1	Filtrering . . . . .	58
5.1.2	Metodens begrænsninger . . . . .	59
5.2	Resultater for Site D datasættet . . . . .	59
5.3	Resultater for NGRIP-datasættet . . . . .	65
5.4	Resultater for CFA-datasættet fra Summit . . . . .	73
<b>6</b>	<b>Konklusion</b>	<b>79</b>
<b>A</b>	<b>Definition af <math>\delta</math>-notation</b>	<b>83</b>
<b>B</b>	<b>Diffusion og tilbagediffusion</b>	<b>85</b>
<b>C</b>	<b>Artikel i Annals of Glaciology</b>	<b>91</b>
	<b>Litteratur</b>	<b>101</b>
	<b>Figurfortegnelse</b>	<b>105</b>



Kort over Grønland med angivelse af nogle af de steder, der gennem tiden er boret iskerner.

Billedet er stillet til rådighed af KMS.



# Indledning

De store iskapper, der dækker Grønland og Antarktis, udgør et fantastisk arkiv med information om fortidens klima. Siden 1966, hvor det for første gang lykkedes at bore gennem den grønlandske Indlandsis, har en lang række projekter søgt at få adgang til denne klimainformation ved at udbore iskerner fra Indlandsisen og den antarktiske iskappe. Isen indeholder oplysninger om temperatur, atmosfærens sammensætning og cirkulationsforhold. Denne klimainformation kan analyseres med måling af isotopforhold og isens indhold af urenheder, men dette har dog begrænset interesse, hvis man ikke er i stand til med rimelig sikkerhed at fastslå isens alder. Desuden er viden om isens alder vigtig for studiet af gletscheres og iskappers indre dynamik.

Datering af en iskerne kan foretages på flere måder, men den mest direkte er at identificere og optælle de årlag, der findes i iskernen. Denne form for datering har den umiddelbare fordel, at den muliggør en absolut datering, hvor den opnåede tidsskala ikke baseres på data fra andre kilder. En række iskerner (eller dele heraf) er gennem de sidste 3 årtier blevet dateret ved manuel identifikation af årlag, heriblandt den grønlandske Dye 3-iskerne. De fleste af disse dateringer er foretaget ved at måle forholdet mellem koncentrationen af de to iltisotoper  $^{18}\text{O}$  og  $^{16}\text{O}$ , der i grove træk er et mål for temperaturen, da sneen i sin tid faldt. Den årlige temperatursvingning afspejles klart i isotopforholdet, der dermed udgør en temperaturproxy, der kan anvendes til identifikation af årlag. Årlagene bliver tyndere med dybden som følge af isens flydning, og diffusion svækker amplituden i den årlige svingning. På grund af en relativ høj årlig akkumulation var det alligevel muligt at identificere den årlige svingning 7-8000 år tilbage i Dye 3-iskernen, med en samlet anslået usikkerhed på blot 30 år.

For de fleste iskerner er en sådan datering dog ikke mulig, da virkningen af diffusion ved lavere årlig akkumulation gør det meget vanskeligt at identificere den årlige svingning ved hjælp af isotop-signalet alene. Et alternativ er at basere identifikation af årlag på isens indhold af urenheder eller iskernens fysiske egenskaber, herunder synlige variationer i isen. Mens koblingen til den underliggende fysiske temperatursvingning er klar ved anvendelse af isotopforhold, er det vanskeligere sikkert at identificere årlagene ved hjælp af f.eks. isens indhold af urenheder. Dette problem håndteres oftest ved at foretage multiparameter-analyse, hvor identifikationen af årlag bygger på flere dataserier fra samme iskerne. Herved har man opnået en betydelig forbedring af sikkerheden for at årlagene identificeres korrekt, men selve analysen

bliver oftest foretaget manuelt, hvilket naturligt medfører en vis subjektivitet. Desuden anvendes ved optællingen ofte kun én eller få simultane dataserier, idet alle serier kun konsulteres i tvivlstilfælde.

I denne opgave vil jeg beskrive en ny metode til multiparameter-datering af iskerner. Den grundlæggende idé er at anvende alle de tilgængelige parallelle dataserier, der indeholder et årligt signal, til at konstruere et signal, der udviser så kraftig årlig variation som muligt. Det er med andre ord tanken at samle de årlige komponenter fra flere dataserier i et enkelt signal, som derefter kan anvendes til datering. Konstruktionen af dette signal vil ske ud fra kriterier, der kan formuleres objektivt, således at resultaterne bliver reproducerbare. Det er håbet, at det producerede årlige signal udviser en så klar årlig svingning, at det efterfølgende vil være muligt at udvikle en egentlig dateringsmetode, hvor antallet af årlag kan bestemmes med god præcision ud fra en række objektive kriterier. Opgavens fokus er dog konstruktionen af det ”optimale årlige signal”, mens selve optællingen af årlag kun berøres overfladisk.

Det skal her bemærkes, at den anvendte metode til at konstruere det samlede årlige signal ikke er ny. Metoden blev beskrevet første gang i 1994, og har siden haft en del forskellige anvendelser, men det er så vidt vides første gang at metoden anvendes til analyse af iskernedata.

I del 1 vil jeg beskrive nogle hovedtræk i iskernefysikkens historie og præsentere en række eksisterende metoder til datering. I del 2 vil jeg beskrive de metoder, der bruges til at samle komponenterne af det årlige signal fra dataserierne, mens del 3 beskriver på hvilken baggrund, jeg har udvalgt dataserierne, og hvordan de klargøres til analysen. Del 4 beskæftiger sig med den praktiske udførelse af analysen og resultaterne præsenteres og vurderes i del 5. I del 6 samles trådene, og jeg giver nogle bud på mulige næste skridt mod en objektiv absolut dateringsmetode.

# Del 1

## Datering af iskerner

Jeg vil i dette kapitel omtale forskellige metoder til datering af iskerner. Det er metoder, der er udviklet siden tresserne, hvor den første iskerne fra den grønlandske Indlandsis blev analyseret. Først vil jeg dog redegøre for den historiske baggrund for iskerneforskningen.

### 1.1 Iskerneforskningens fødsel

I halvtredserne arbejdede Willi Dansgaard med måling af isotopforhold ved hjælp af et massespektrometer, der var indkøbt af Biofysisk Laboratorium med henblik på primært biologiske og fysiologiske eksperimenter. Ved noget Willi Dansgaard selv beskriver som en pludselig indskydelse, begyndte han at spekulere over hvordan forholdet mellem koncentrationen af iltisotoper i regnvand egentlig var. Da han havde et velegnet instrument til at besvare spørgsmålet, samlede han en række vandprøver i ølflasker på græsplænen under et længere regnskyl og målte derefter isotopforholdet i prøverne [*Dansgaard, 2000*]. For ilt gælder, at langt den største del af ilt-atomerne er isotopen  $^{16}\text{O}$ , mens størstedelen af de resterende er  $^{18}\text{O}$ . Isotopindholdet i en prøve angives oftest som en relativ afvigelse i forhold til indholdet i en standardprøve, og denne afvigelse kaldes prøvens  $\delta^{18}\text{O}$ -værdi. En negativ  $\delta^{18}\text{O}$ -værdi svarer til, at der er mindre af den tunge  $^{18}\text{O}$ -isotop i prøven end i standardprøven. Se i øvrigt appendix A for en mere fyldig og formel definition af  $\delta$ -værdier. Det opsamlede regnvand stammede fra en stor koldfronts passage over Danmark, og Willi Dansgaard opdagede en systematisk variation af  $\delta^{18}\text{O}$ -værdierne i løbet af frontens passage [*Dansgaard, 1953*]. Han tolkede dette, som at nedbørens  $\delta^{18}\text{O}$ -værdi var direkte koblet til temperaturen på nedbørens dannelsessted (kondensationstemperaturen), en iagttagelse der i hovedtræk har vist sig at holde stik.

Ansporet af dette resultat overvejede han om man kunne bruge isotopinformation fra Indlandsisen som kilde til oplysninger om palæoklimaet. I [*Dansgaard, 1954*] beskriver han første gang sin idé<sup>1</sup> :

---

<sup>1</sup>Isotopforhold blev på det tidspunkt angivet absolut og ikke som delta-værdier. I den citerede artikel betegnes forholdet med  $a_f$ .

In certain areas on the Greenland Ice Cap is a distinct layer formation caused by melting in the summer season. On the supposition that the character of the circulatory processes, in all essentials, have not varied over a long period of time, the above, in the opinion of this author, offers the possibility by measurements of the  $a_f$  in these layers of ice to determine climatic changes over a period of time of several hundred years of the past. ...

An investigation will be undertaken as soon as the opportunity offers.

Denne idé kan opfattes som en "arbejdstitel" for de næste mange års arbejde med iskerner på Københavns Universitet, selvom ideen med datering ved hjælp af optælling af årlige smeltelag aldrig blev en realitet.

For at undersøge sammenhængen mellem  $\delta^{18}\text{O}$ -værdier og temperaturer yderligere skaffede Willi Dansgaard i første omgang flodvandsprøver fra store dele af kloden gennem Ø.K.s net af kontorer og en håndfuld (smeltede) isprøver fra EGIG-traversen, der i 1959 krydsede Midtgrønland. Flodvandsprøverne tjente primært til at verificere at der *var* forskel på isotopforholdene afhængig af nedbørssted, og at denne forskel i nogen grad kunne knyttes sammen med temperaturen. De flodvandsprøver der stammede fra tempererede og polare egne viste, sammen med prøverne fra indlandsisen, en noget klarere sammenhæng mellem  $\delta$ -værdi og temperatur, men prøverne var ikke opsamlet under tilstrækkeligt kontrollerede forhold til, at der umiddelbart kunne udledes en entydig sammenhæng.

I starten af 1960'erne begyndte man at indsamle regnvandsprøver fra hele verden (undtagen i østblokken, hvor man var bange for at prøverne kunne misbruges militært) i et samarbejde mellem World Meteorological Organisation og den internationale atomenergi-kommission, IAEA. Formålet var blandt andet at spore radioaktive isotoper fra atombombeprovsprængninger, men det lykkedes Willi Dansgaard at overtale IAEA's direktør til at afgive en del af de indsamlede prøver til måling af  $\delta^{18}\text{O}$  og  $\delta\text{D}$  i København<sup>2</sup>, hvorved der på ret kort tid blev tilvejebragt datagrundlag for en mere grundig undersøgelse af sammenhængen mellem årsmiddeltemperatur og isotopsammensætningen af regnvand. Den fundne sammenhæng mellem temperatur og  $\delta^{18}\text{O}$  viste en klar sammenhæng for data fra atlantiske kyststationer og polare egne, og viste at en ændring på 1 ‰ svarede til en temperaturændring på omtrent 1,4 - 1,5°C [*Dansgaard*, 1964].

Det skulle senere vise sig, at sammenhængen 1 ‰  $\sim$  1,4 - 1,5°C ikke har generel gyldighed, heller ikke for tempererede og polare egne<sup>3</sup>. Sammenhængen blev fundet ved sammenligning af sammenhørende ( $\delta^{18}\text{O}$ , T)-værdier fra en lang række målestationer og repræsenterer derfor en rumlig sammenhæng. Ved tolkning af paleotemperatursignalet fra iskerner skal man iste-

<sup>2</sup> $\delta\text{D}$  læses delta-deuterium og betegner jævnfør definitionen af  $\delta^{18}\text{O}$  i appendix A den relative afvigelse af isotopforholdet  $^2\text{D}/^1\text{H}$  i forhold til en standardprøve.

<sup>3</sup>Det skal for retfærdighedens skyld siges, at Willi Dansgaard også selv var meget forsigtig med at tillægge den fundne sammenhæng generel gyldighed.



det anvende en sammenhæng mellem  $\delta^{18}\text{O}$  og temperatur, der har gyldighed tilbage i tiden. Denne kalibrering er behæftet med en betydelig usikkerhed, blandt andet fordi en sådan sammenhæng meget vel kan tænkes kun at være gyldig for en given region i et givet tidsrum. Især når klimaet skifter mellem afgørende forskellige klimatiske tilstande, som det f.eks. skete i forrige istid med skift mellem kolde stadialer og varme interstadialer, er det usikkert hvordan forholdet mellem  $\delta^{18}\text{O}$  og temperatur udvikler sig.

Ved hjælp af inversion af borehulstemperaturen fra GRIP-borehullet lykkedes det Dorthe Dahl-Jensen m.fl. at rekonstruere paleotemperaturer direkte fra den målte borehulstemperaturprofil ved anvendelse af oplysninger og antagelser om isens flydning og den geotermiske flux [Dahl-Jensen et al., 1998, Johnsen et al., 1995]. Denne uafhængige metode til bestemmelse af paleotemperaturer viser, at der ved sidste istids kulmination (LGM - Last Glacial Maximum) på Summit var  $23 \pm 2^\circ$  koldere end i dag.  $\delta^{18}\text{O}$ -kurverne fra både GRIP [Johnsen et al., 1997] og GISP2 [Grootes & Stuiver, 1997] viser LGM-værdier der er ca.  $7 - 8\text{‰}$  lavere end i dag, svarende til en sammenhæng mellem  $\delta^{18}\text{O}$  og temperatur på omtrent  $3^\circ\text{C}/\text{‰}$ .

På samme tid som Willi Dansgaard nåede frem til omtalte sammenhæng mellem  $\delta^{18}\text{O}$ -værdier og temperatur, var det amerikanske forsvar i gang med at bore en iskerne i Nordvestgrønland. På basen *Camp Century*, der var gravet ned under Indlandsisens overflade ca. 200 hundrede kilometer fra Thule Air base, borede *U.S. Army Cold Regions Research and Engineering Laboratory* en iskerne gennem hele iskappen. Det tog adskillige år og man måtte anvende flere forskellige boremetoder, før man i 1966 nåede bunden efter at have gennemboret ialt 1390 meter is.

Willi Dansgaards idé fra 1954 blev til virkelighed : Willi Dansgaard fik lov til at måle en  $\delta^{18}\text{O}$ -profil på hele iskernen. Analysen af disse data, der primært blev udført af Willi Dansgaard og Sigfus J. Johnsen, blev præsenteret i to artikler i henholdsvis *Journal of Glaciology* og *Science* i 1969 [Dansgaard & Johnsen, 1969, Dansgaard et al., 1969]. I disse artikler gøres det klart, at en dateringsmetode må udvikles før den paleoklimatiske information i iskernen kan udnyttes fuldt ud. Willi Dansgaard og Sigfus Johnsen gør derefter rede for, at det er muligt at datere de øvre dele af iskernen ved at måle detaljerede  $\delta^{18}\text{O}$ -profiler og heri finde den årlige svingning, hvor hvert maksimum svarer til høje sommertemperaturer. I de nedre dele af kernen har diffusion og udtynding af årlagene imidlertid gjort det umuligt at gennemføre denne analyse, hvorfor andre metoder må tages i brug. De introducerer derfor en flydemodel (der i løbet af 1970'erne bliver kendt som Dansgaard-Johnsen modellen, eller blot DJ-modellen), hvor isens alder tilnærmelsesvis bestemmes ud fra simple antagelser om isens bevægelse. Med disse to artikler satte Willi Dansgaard og Sigfus Johnsen dagsordenen for de næste mange års bestræbelser på at datere iskerner. I det følgende vil jeg omtale disse bestræbelser, der naturligt lader sig opdele i *stratigrafiske metoder*, hvor årlag identificeres og optælles, og *modelleringsmetoder*, hvor man på mere eller mindre sofistikeret vis regner sig frem til isens alder ved at beregne isens flydemønster.

## 1.2 Datering af is ved modellering af isens flydning

J. F. Nye var en pioner i bestræbelserne på at bestemme en ismasses alder ved hjælp af overvejelser om isens flydning. Allerede i de tidlige 1950'ere gjorde han sig overvejelser om den indre dynamik i iskapper, i første omgang ved at behandle isflydning som plastisk flydning [Nye, 1951]. Nye var primært interesseret i at bestemme de indre spændings- og hastighedsfelter i gletschere og iskapper (f.eks. [Nye, 1957]), og brugte ikke selv sine resultater til at datere is med.

### 1.2.1 Nye-modellen

Nye konkluderede ud fra sine beregninger af hastighedsfeltet i iskapper, at den horisontale hastighed er nogenlunde uafhængig af dybden på størstedelen af iskappen og kun ændrer sig væsentligt nær bunden, hvor den dominerende del af deformationen sker ved forskydning. Han lavede derfor en simpel model for flydningen af is, ved at antage at den vertikale deformationshastighed er konstant langs en lodret linie i iskappen, hvilket for plan flydning<sup>4</sup> er ækvivalent med at den horisontale hastighed er uafhængig af dybden [Nye, 1963].

Vi antager i det følgende, at flydning kan opfattes som plan, således at der ingen flydning sker i  $y$ -retningen. Den horisontale flyderetning benævnes  $x$  og regnes fra isdeleren mod isranden og højden over den flade bund kaldes  $z$ . Hastighederne i  $(x, y, z)$ -retningerne benævnes  $(u, v, w)$ . Se figur 1.1, hvor koordinatsystemet er illustreret.

Vi antager at iskappen er i ligevægt, og at akkumulationsraten mellem isdeleren og en afstand  $x_0$  er givet ved konstanten  $a$ <sup>5</sup>.  $x_0$  skal være tilstrækkelig lille til at vi befinder os i iskappens akkumulationsområde, og der derfor ingen afsmeltning er fra toppen. Den samlede tilførsel af is som følge af nedbør i området mellem isdeleren og  $x_0$  er da  $a \cdot x_0$ . Idet vi anvender Nyes antagelse om at den horisontale hastighed  $u$  er uafhængig af dybden, fås at ismængden, der forlader området gennem det lodrette snit i  $x_0$  er  $u(x_0) \cdot H(x_0)$ , hvor  $H(x)$  er iskappens højdeprofil<sup>6</sup>. Da vi har antaget at iskappen er i ligevægt og antager at der ingen bundsmeltning er, må der gælde at  $a \cdot x_0 = u(x_0) \cdot H(x_0)$ . Ved at anvende kontinuitetsligningen<sup>7</sup>  $\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0$  fås

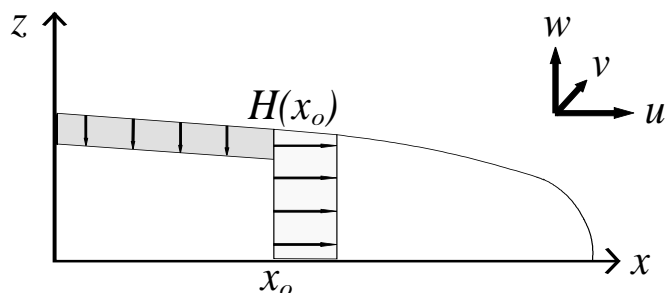
$$\left. \frac{\partial w}{\partial z} \right|_{x_0} = - \left. \frac{\partial u}{\partial x} \right|_{x_0} = - \frac{a}{H(x_0)} \quad (1.1)$$

<sup>4</sup>Ved plan flydning (eller 2D-flydning) forstås at flydningshastigheden kan indeholdes i et plan. Set fra oven svarer det til at to nærliggende ispartiklers banekurver er parallelle. Plan flydning er en rimelig antagelse for størstedelen af den indre del af den grønlandske indlandsis, hvor isen flyder stort set enten øst eller vest ud fra en nord-syd-gående delelinie, der kaldes isdeleren.

<sup>5</sup>Akkumulationsrater måles i glaciologien oftest i (centi)meter is pr. år, idet nyfalden snes densitet varierer meget.

<sup>6</sup>Sne der ligger på en iskappes overflade, kaldes firn, og den proces hvor sne bliver til is, kaldes firnifikation.  $H(x)$  er korrigeret for at sneen har lavere densitet end is ved at firnlaget tænkes erstattet af et lag is med samme masse.

<sup>7</sup>Leddene  $\frac{\partial v}{\partial y}$  udgår for plan flydning. Desuden anvendes at is er et inkompressibelt materiale.



Figur 1.1: Illustration af det anvendte koordinatsystem og Nye-modellens ligevægts-antagelser. Den årlige pålejring (stribet areal  $a \cdot x_0$ ) flyder ud gennem det lodrette snit i  $x_0$ , hvor den horisontale hastighed antages at være uafhængig af dybden (gråtonet areal  $u(x_0) \cdot H(x_0)$ ).

Heraf fås  $w(z) = \frac{-a \cdot z}{H(x_0)}$ , hvor randbetingelsen  $w(0) = 0$  udtrykker at der ingen bundsmeltning er. Da  $w = \frac{dz}{dt}$  kan alderen  $t(z)$  af is i højden  $z$  over bunden findes ved integration :

$$t = \int_{H(x_0)}^z \frac{1}{w(z')} dz' = \frac{H(x_0)}{a} \ln \left( \frac{H(x_0)}{z} \right) \quad (1.2)$$

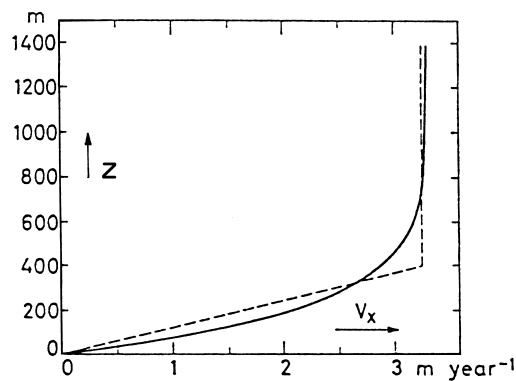
Denne meget simple sammenhæng mellem dybde og alder ses stadig i dag anvendt ved datering af is fra toppen af iskerner. Sålænge man ikke nærmer sig bunden er det en rimelig tilnærmelse, der til gengæld ikke nødvendiggør kendskab til andre fysiske størrelser end istykkelse og akkumulation. En af svaghederne ved Nye-modellen er, at antagelsen om uændret horisontal hastighed langs en lodret linie medfører at hele gletscheren glider som en blok over underlaget, hvilket kun med rimelighed opfyldes af iskapper, der har smelte-temperaturer langs bunden. I så fald vil der være bundsmeltning, hvilket er i modstrid med de anvendte antagelser. Man kan dog nemt modificere ligning 1.2 så bundsmeltning indregnes, hvilket dog indfører en parameter, der kan være vanskelig at bestemme en god værdi for.

### 1.2.2 Dansgaard-Johnsen-modellen

Som beskrevet på side 11 indførte Willi Dansgaard og Sigfus Johnsen under analysen af Camp Century-iskernen i [*Dansgaard & Johnsen, 1969*] en flydemodel (herefter omtalt som DJ-modellen), der tager højde for nogle af Nye-modellens svagheder. Udgangspunktet var en beregning af den horisontale hastighedsprofil i en iskappe ved anvendelse af Glens flydelov<sup>8</sup> [*Glen, 1955*] under anvendelse af en række forsimplede antagelser. Figur 1.2 viser den horisontale flydehastighed som funktion af dybden for Camp Century og den tilnærmede hastighedsprofil, som anvendes i DJ-modellen.

Denne tilnærmelse kan virke grov, men man bør huske, at hastighedsprofilen bestemt ved Glens lov i sig selv er en tilnærmelse. Glens flydelov tager nemlig ikke højde for isens defor-

<sup>8</sup>Glens lov udtrykker at deformationshastigheden for polykrystallin is er proportional med den påtrykte spænding opløftet til en passende eksponent  $n$ , der typisk sættes til 3. Proportionalitetskonstanten afhænger blandt andet af temperaturen.



Figur 1.2: Horizontal flydehastighed beregnet ved integration af Glens lov (optrukket linie), og DJ-modellens forsimplede profil (stiplet). Fra [Dansgaard & Johnsen, 1969].

mationshistorie, men beskriver den del af flydningen, der kaldes "secondary creep", hvilket ikke er en realistisk situation i den nedre del af iskappen<sup>9</sup>. Ud fra den tilnærmede horisontale hastighedsprofil beregnes den vertikale hastighedsprofil på samme måde som skitseret for Nye-modellen:

$$w(z) = \begin{cases} \frac{-r \cdot z^2}{2h} & \text{for } 0 < z \leq h \\ -r \left( z - \frac{h}{2} \right) & \text{for } h < z \leq H(x_0) \end{cases} \quad \text{med } r = \frac{a}{H(x_0) - \frac{h}{2}} \quad (1.3)$$

hvor  $h$  er den såkaldte *knækhøjde*, der er den højde over bunden, hvor den horisontale flydehastighed går fra at være lineært voksende til at være konstant (på figur 1.2 er  $h = 400\text{m}$ ). På samme måde som for Nye-modellen i ligning 1.2 kan isens alder som funktion af dybden beregnes ved integration  $t = \int_{H(x_0)}^z \frac{1}{w(z')} dz'$ . Forskellen på de resulterende aldre for Nye-modellen og DJ-modellen er lille i den øverste halvdel af iskappen, men begynder at blive betydelig når vi nærmer os knækhøjden. Nær bunden er forskellen enorm: 20 meter over bunden på Camp Century borestedet forudsiger Nye-modellen en alder på ca. 16.500 år, mens DJ-modellen giver en alder på langt over 100.000 år.

### 1.2.3 Mere avancerede modeller

DJ-modellen kan generaliseres på to oplagte måder. For det første kan man vælge en horisontal hastighedsprofil, hvor iskappen opdeles i et antal lag, i hvilke den horisontale hastighed varierer lineært, de såkaldte sandwich-modeller [Reeh, 1989]. Ved anvendelse af mange sådanne lag kan hastighedsprofilen tilnærmes vilkårligt godt til den ønskede form. Man kan også indføre bundglidning, hvor bundglidningshastigheden typisk fastlægges som en fast procentdel  $f_b$  af overfladehastigheden. Det afledte udtryk for den vertikale hastighed bliver da en anelse mere

<sup>9</sup>Is, der deformeres, ændrer struktur, idet de enkelte iskrystallers optiske akser drejes mod kompressionsaksen (se f.eks. [Azuma & Higashi, 1985]). Situationen kompliceres yderligere af, at der foregår rekryttering. Skal man i detaljer beskrive flydning af is, der har undergået kraftig deformation, må man derfor holde regnskab med de enkelte ismassers bevægelse gennem iskappen og den deformation, de allerede har oplevet. For hver ismasse må man derefter anvende en flydemodel, der er tilpasset de lokale forhold (se f.eks. [Azuma, 1994]).

kompliceret, end hvad vi så i ligning 1.3, men det ændrer ikke på, at man kan opstille analytiske udtryk for såvel hastigheder som isens alder. Modeller, der anvender mere komplicerede udtryk for den horisontale hastighed, men stadig bygger på antagelsen om plan flydning, har dog ikke nydt den store bevågenhed, idet DJ-modellen giver rimelig præcise resultater med et minimum af parametre.

Nye-modellen såvel som DJ-modellen har dog den oplagte svaghed, at det ikke er en realistisk antagelse at akkumulationsraten er konstant (hverken over større geografiske områder eller over tid), og at de færreste - om nogen - iskapper befinder sig i en ligevægtstilstand. I [Dahl-Jensen *et al.*, 1993] videreudvikles en DJ-model, så der tages højde for variation i den årlige akkumulation. Udgangspunktet er en almindelig DJ-model, hvorfra udtyndingsraten som funktion af dybden bestemmes. Parametrene i DJ-modellen er tilpasset, så isens modelerede alder stemmer overens med resultatet fra en stratigrafisk datering af Holocæn-delen af kernen (se afsnit 1.3). I den øverste del af kernen er der god korrelation mellem  $\delta^{18}\text{O}$ -værdierne og de observerede årlagstykkelser, der er korrigeret for udtynding ved hjælp af den omtalte udtyndingsrate. Det sandsynliggøres, at den heraf fundne sammenhæng mellem  $\delta^{18}\text{O}$  og akkumulation er gyldig længere tilbage i tiden, hvorved den målte  $\delta^{18}\text{O}$ -profil kan anvendes som en kontinuert akkumulations-profil. En tidsskala for resten af GRIP-kernen beregnes med en model, der dog stadig bygger på samme antagelser som en DJ-model med undtagelse af den variable akkumulationsrate.

En anden tilgang er de såkaldte flydelinie-modeller, der beskriver isens flydning fra isdeleren langs et snit mod isranden. Der kan tages højde for divergens og konvergens i flydningen langs linien, men ofte antages plan flydning. Det er også muligt at inddrage overflade- såvel som bundtopografi, hvilket naturligvis kan give meget store ændringer i dateringen af de nederste dele af en iskappe. Et meget vigtigt element i disse modeller er, at det er muligt at anvende forskellige flydelove i forskellige dybder, således at man f.eks. kan tage højde for det eksperimentelle faktum, at istidsis er 2-3 gange lettere at deformere end is fra Holocæn. Modeller af denne type kræver omfattende numeriske beregninger og stor viden om de fysiske forhold omkring borestedet for at kunne beskrive flydningen realistisk. I [Reeh, 1988] og [Hvidberg, 1996] bestemmes på denne måde en modelleret aldersprofil, mens [Dahl-Jensen, 1989] ikke eksplicit bestemmer aldersprofilen, selvom det er muligt at beregne alderen som funktion af dybden på et givet sted ud fra det beregnede flydemønster. I [Reeh, 1989] sammenlignes estimer af præcisionen af en moderat omfattende flydelinie-model og en DJ-model. Usikkerheden ved datering af 1000 år gammel is er her angivet til ca. 3 år for flydelinie-modeller og 30 år for DJ-modellen.

Egentlige 3-dimensionale flydemodeller, der modellerer hele iskapper på en gang, anvendes ikke endnu til datering af iskerner<sup>10</sup>. Blot at få disse modeller til at reproducere indlandsisens

---

<sup>10</sup>Vostok-iskernen er blevet dateret ved en 3-dimensional flydemodel [Petit *et al.*, 1999], men modellen tvinges til at producere givne aldre ved henholdsvis 1534 og 3254 meters dybde, og den konstruerede tidsskala er derfor ikke opnået ved egentlig flyde-modellering, men ved en avanceret form for interpolation.

nuværende udbredelse og højde-forhold har vist sig at være en ambitiøs opgave, og det er derfor ikke oplagt at anvende disse modellers resultater i dateringsøjemed. Desuden er den vertikale opløsning i disse modeller oftest ikke fin nok til, at de umiddelbart kan bruges til datering.

### 1.3 Datering af is med stratigrafiske metoder

Stratigrafiske metoder baserer sig som navnet angiver på strata - det vil sige lag - i isen. Ved lag forstås til dels synlige mørke eller "tågede" lag i den ellers klare is, men ligeså ofte refereres til årlag, der fastlægges ved måling af isens indhold af urenheder eller ioner samt ved måling af isotopforhold. Stratigrafiske metoder tjener i hovedtræk 2 formål :

- Årlag, defineret ud fra enten visuel stratigrafi eller fra systematiske variationer i  $\delta^{18}\text{O}$ , elektrisk ledningsevne eller koncentration af forskellige urenheder, kan anvendes til en absolut datering af iskerner ved simpelthen at tælle antallet af årlag fra toppen af kernen. Dye 3-iskernen fra Sydgrønland blev dateret ca. 8000 tilbage i tiden ved hjælp af det årlige  $\delta^{18}\text{O}$ -signal. Dateringen af GISP2-iskernen baserede sig på visuel stratigrafi og vil blive omtalt særskilt i afsnit 1.3.3.
- Karakteristiske lag, der kan stamme fra f.eks. vulkanudbrud, radioaktivt nedfald fra kernevåbenafprøvning eller fra markante klimatiske begivenheder, kan sommetider findes i flere iskerner (og måske endda i havsedimentkerner). Sådanne lag, der antages at være deponeret samtidigt flere steder, kaldes reference-horisonter, og kan i nogle tilfælde bruges til absolut datering, og i andre tilfælde bruges til at knytte to signaler sammen tidsmæssigt (relativ datering). Som eksempel herpå kan nævnes, at dateringen for Dye 3-iskernen blev overført til GRIP-iskernen ved at anvende genkendelige vulkanske lag [*Hammer et al.*, 1999].

Såvel optælling af årlag som reference-horisonter anvendes desuden til at kalibrere flydemodeler. Eksempelvis kan parametrene  $a$  og  $h$  i en Dansgaard-Johnsen flydemodel (se afsnit 1.2.2) tilpasses, så de modellerede årlagstykkelser passer med årlagstykkelserne i et antal optalte sekvenser fra forskellige dybder. På samme måde kan man anvende en velkendt og veldateret reference-horizont som f.eks. overgangen mellem Yngre Dryas og Holocæn til at tilpasse parametrene i en flydemodel [*Hammer et al.*, 1978, *Schwander et al.*, 2001].

#### 1.3.1 Datering ved hjælp af reference-horisonter

Veldaterede reference-horisonter er væsentlige, idet de gør det muligt at vurdere præcisionen af andre dateringsmetoder. Desværre er alle reference-horisonter, der er dateret præcist med uafhængige metoder, unge ud fra et iskerne-daterings synspunkt. Fra Island har man optegnelser om vulkanudbrud fra omkring år 1000, og dateringen heraf må vurderes som værende

sikker. Et af de ældste veldaterede vulkanudbrud er Vesuvs udbrud i år 79. Tilstedeværelsen af vulkanske lag i iskerner kan forholdsvis nemt bestemmes ved måling af den elektriske ledningsevne (ECM), idet syre fra de vulkanske gasser giver meget stor elektrisk ledningsevne [Hammer, 1980]. Islandske vulkanudbrud er meget velegnede som reference-horisonter, idet de har afsat nemt identificerbare aftryk i isen på Grønland, mens vulkanudbrud på mellem-breddegrader ikke altid har givet et let genkendeligt signal i iskernerne. I nogle tilfælde kan vulkanudbrud identificeres i kraft af udbruddets placering i serien af kendte vulkanudbrud, andre gange må den kemiske sammensætning af signalet analyseres nærmere for at fastslå hvilken vulkan, der kan være tale om.

Ældre reference-horisonter kan normalt ikke dateres absolut med uafhængige metoder og kan derfor ikke direkte bruges til iskerne-datering, men kan med stor fordel anvendes til relativ datering, hvor en tidsskala overføres fra en iskerne til en anden ved hjælp af et antal reference-horisonter, der genfindes i begge iskerner. Også enkeltstående smeltelag i iskerner, der stammer fra boresteder hvor temperaturen normalt er under frysepunktet året rundt, kan anvendes til sammenligning af tidsskalaer fra forskellige iskerner.

Til datering af helt ung is kan man anvende en reference-horizont, der stammer fra henfaldsprodukter fra afprøvning af kernevåben. I 1953 begyndte man at prøvesprænge kernevåben ved åbne sprængninger, hvor store mængder radioaktivt materiale blev frigivet til atmosfæren. Man kan ved måling af  $\beta$ -aktiviteten af isprøver konstatere en stigning i aktiviteten i 1953, efterfulgt af en meget kraftig top i 1954. Efter en pause blev prøvesprængningerne genoptaget i 1961-62, hvilket giver en meget markant top i  $\beta$ -aktiviteten i 1963 [Clausen & Hammer, 1988]. Disse reference-horisonter kan anvendes til at bestemme dybden af årlagene fra år 1954 og 1963, hvilket er et godt udgangspunkt for datering ved optælling af årlag, idet de øverste 10 - 20 årlige svingninger i  $\delta^{18}\text{O}$ -kurven kan være vanskelige at identificere, da de højfrekvente sub-årlige svingninger endnu ikke er forsvundet ved diffusion (se appendix B). Også ved bestemmelse af akkumulationsrater ud fra shallow-kerner<sup>11</sup> kan denne metode være af stor praktisk betydning.

### 1.3.2 Datering ved identifikation af årlag

Datering af iskerner alene ved hjælp af synlige lag og variationer i isens struktur er kun i ét tilfælde (se afsnit 1.3.3) blevet anvendt til datering af iskerner, der rækker mere end nogle hundrede år tilbage, men er i nogen grad blevet anvendt på shallow-kerner [Langway, Jr., 1967]. Til dybde-iskerner måles istedet langs kernen isotopforhold, koncentration af udvalgte ioner samt elektrisk ledningsevne, der alle udviser et årligt signal. Jeg vil her kort redegøre for årsagerne til nogle udvalgte parametres årlige variation i is fra den grønlandske Indlandsis, men det skal understreges, at listen er ufuldstændig.

- $\delta^{18}\text{O}$ -værdien afhænger af kondensationstemperaturen ved nedbørens dannelse, og den

---

<sup>11</sup>Shallow-kerner er korte iskerner, der stammer fra de øverste lag af tykke iskapper. Ordet *shallow* kan ikke umiddelbart oversættes til et dækkende dansk ord.

årlige temperatursvingning afspejles derfor i en målt  $\delta^{18}\text{O}$ -profil. Temperaturen svinger naturligvis ikke jævnt i løbet af året, men de højfrekvente variationer diffunderer væk i løbet af de første ca. 10 år [Johnsen, 1977], og efterlader en tydelig årlig svingning i  $\delta^{18}\text{O}$ -værdien. Hvis den årlige akkumulation er højere end ca. 20 cm bevares den årlige svingning i isen efter firnifikation, og svingningen kan i så fald følges tusinder af år tilbage i tiden. Diffusion svækker signalet både under firnifikationen og efterfølgende i isen, og er behandlet særskilt i appendix B.

- Isens elektriske ledningsevne for jævnstrøm (ECM-signalet) er et udtryk for isens indhold af syrer. Målingen kan udføres ved at måle strømmen når to elektroder pålagt et spændingsfald på 1250V trækkes over en ren isoverflade<sup>12</sup>. Det målte signal er et udtryk for  $[\text{H}^+]$  i isen, og det viser sig at surheden er størst om sommeren (i de perioder hvor vulkanske kilder ikke bidrager). Dette kan forklares ved, at der ved biologisk aktivitet i havene dannes  $\text{H}_2\text{S}$ , som ved fotokemiske processer oxideres til svovlsyre. Da både dannelseshastigheden for  $\text{H}_2\text{S}$  og mængden af det nødvendige lys er størst om sommeren, fås et højere syreindhold om sommeren end om vinteren [Hammer, 1980]. På figur 1.3 ses det klart, at der er god korrelation mellem  $\delta^{18}\text{O}$  og ECM-signalet. ECM kan måles med en opløsning på et par millimeter, hvilket er væsentlig bedre, end hvad de fleste andre analyse metoder har kunnet præstere. På grund af den gode opløsning kunne den årlige svingning i Byrd-iskernen følges helt ned til 1900 meters dybde, hvor årlagstykkelsen er omkring 1 cm [Hammer et al., 1994].
- Isens dielektriske egenskaber (DEP) under påvirkning af vekselspænding afhænger både af isens surhed og af indholdet af primært  $\text{Cl}^-$  og  $\text{NH}_4^+$ , og er en slags vekselspændingsudgave af ECM [Moore et al., 1994, Wilhelms et al., 1998]. Isens kapacitans og ledningsevne måles for et bredt udvalg af frekvenser og resultaterne er på mange måder parallelle med ECM-resultater. Da DEP måles uden at kernen påvirkes mekanisk (og derfor heller ikke skal renses), er metoden velegnet til brug i felten.
- Støvindholdet, og dermed indholdet af blandt andet  $\text{Mg}^{2+}$  og  $\text{Ca}^{2+}$ , i isen er størst om foråret. Store dele af det støv, der findes i den grønlandske Indlandsis, stammer formentlig fra støvstorme i Gobi-ørkenen, der primært forekommer i det tidlige forår [Biscaye et al., 1997]. De nærmere detaljer omkring transporten af støv til Indlandsisen er ikke vel afklarede, og der søges stadig efter en forklaring på de store ændringer i støvindhold, der f.eks. kan observeres ved omslag mellem stadialer og interstadialer i istiden [Fuhrer et al., 1999].
- Indholdet af  $\text{Na}^+$  og  $\text{Cl}^-$  i isen er maksimalt sent på vinteren. En mulig forklaring herpå

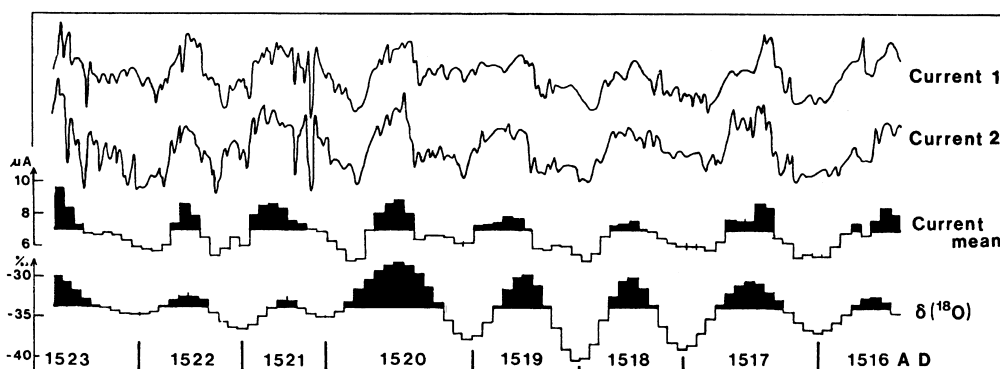
---

<sup>12</sup>Iskappeis kan i praksis stort set ikke lede jævnstrøm. Den strøm, der løber, skyldes forskydninger af ladningerne i isen og er en transient strøm. Elektroderne skal derfor bevæges over isoverfladen med en passende hastighed.

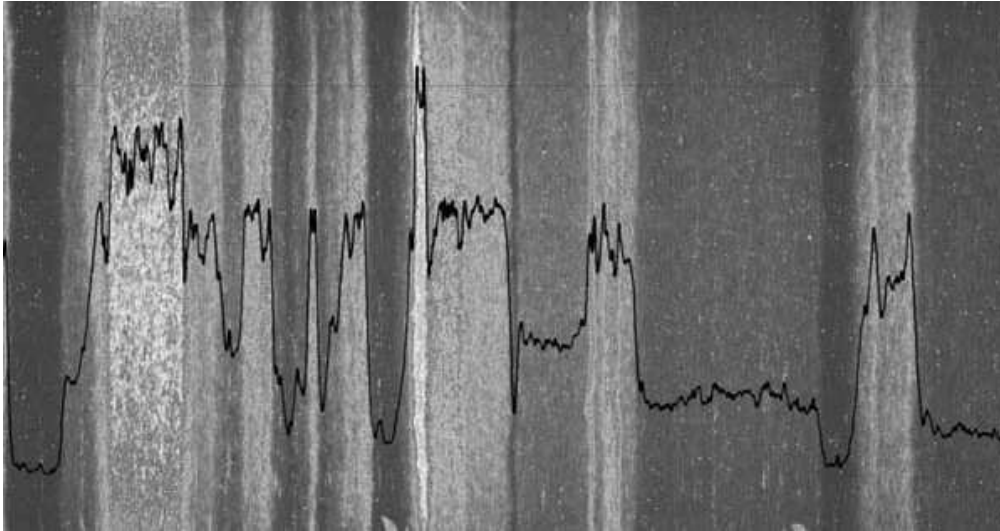


er, at de dominerende vejrsystemer bringer atlantiske luftmasser med stort indhold af havsalt tæt på Grønland, hvor der i forbindelse med voldsomme vinterstorme sker en opblanding med de højere luftlag, der bevæger sig ind over Indlandsisen. Om sommeren har størstedelen af luften over indlandsisen kontinental oprindelse, hvorfor indholdet af  $\text{Na}^+$  og  $\text{Cl}^-$  er lavere [Steffensen, 1988].

- Koncentrationerne af  $\text{SO}_4^{2-}$  og  $\text{NO}_3^-$  udviser også årlige svingninger. Blandt de mulige sulfat-kilder er nedbrydning af organisk materiale, vulkaner samt antropogene bidrag, mens nitrats mulige kilder er talrige. Hverken kilder eller transportmekanismer er velkendte [Clausen & Langway, Jr., 1989].
- I isen findes synlige tågede bånd, der menes at repræsentere et årligt signal. Tågede bånd blev anvendt til datering af GISP2-iskernen, på trods af at det ikke klart blev verificeret, at de tågede bånd udgør et årligt signal (afsnit 1.3.3). Sammenhængen mellem den visuelle stratigrafi og det årlige signal i andre dataserier fra NGRIP-iskernen analyseres i øjeblikket. Søren Wedel Nielsen har i samarbejde med blandt andet Joseph Kipfstuhl udviklet en linie-scanner, der kan foretage en præcis digital affotografering af det visuelle signal i iskerner. NGRIP-iskernen blev scannet på denne måde i felten i boresæson 2000, hvor man nåede ned til dybder på omkring 2900 meter. På figur 1.4 ses et eksempel på resultatet af en sådan scanning. Søren Wedel Nielsen arbejder i øjeblikket på at omsætte informationen i det scannede signal til en enkelt numerisk dataserie, som eksemplificeret ved den superponerede kurve på figur 1.4. Der er scannet kerner i et dybdeinterval, der svarer til det meste af forrige istid, og det er håbet at denne ultra-højopløste dataserie kan anvendes til datering. Dette forudsætter dog, at det bliver muligt klart at vise, at de lyse bånd repræsenterer et årligt signal.



Figur 1.3: ECM-signal og  $\delta^{18}\text{O}$ -sekvens for et kort stykke af Crête-iskernen. Øverst 2 målinger fra samme stykke af iskernen, derunder gennemsnittet af de to målinger, plottet med en opløsning, der svarer til opløsningen af  $\delta^{18}\text{O}$ -kurven, der ses nederst.  $\delta^{18}\text{O}$ -signalet er tilbagediffunderet (se appendix B). Det ses, at ECM-målingen i det store hele er reproducerbar, og at ECM-signalet udviser god korrelation med  $\delta^{18}\text{O}$ -værdien. Fra [Hammer, 1980].



Figur 1.4: Scannet billede af et udsnit af *bag* 3202 (iskerner udskæres i *bags* af 55 cm) fra NGRIP-iskernen, svarende til dybden 1760,55 - 1761,10 meter. Kurven oven på scanningsbilledet viser et forsøg på at konvertere billedet til en enkelt numerisk dataserie. Billedet er venligst stillet til rådighed af Søren Wedel Nielsen.

Ved at udnytte de årlige variationer er det i princippet muligt at datere en iskerne ved at tælle antallet af årlag fra toppen og nedefter. Der er dog en række forhold der gør denne fremgangsmåde vanskelig eller ligefrem umulig :

- Det må godtgøres, at den svingning, man anvender til dateringen, i realiteten er et årligt signal. For iltisotopforhold er det veldokumenteret, at  $\delta^{18}\text{O}$ -signalet er koblet til kondensationstemperaturen, der udviser en klar årlig variation. For nogle parametres vedkommende formoder man, at den observerede svingning er en årlig svingning, men formodningen bygger oftest på at koncentrationen korrelerer med f.eks.  $\delta^{18}\text{O}$  i udvalgte prøver, og i mindre grad på at årsagen til den årlige variation kan forklares.
- Det må defineres, hvilke karakteristika der definerer et årlag. En typisk måleserie, hvad enten den udtrykker iltisotopforhold eller koncentrationer af ioner, indeholder toppe, der ikke kan henføres til den årlige svingning. Årsagerne hertil er mangfoldige, men bidrag fra vulkanudbrud, naturlige svingninger i temperatur på kort tidsskala eller simpel støj er eksempler. Variationer i den årlige akkumulation vanskeliggør desuden identifikationen af årlag på baggrund af toppenes varighed. Jeg har ikke fundet et eneste ord i litteraturen omkring hvordan et årlag defineres, hvilket meget vel kan skyldes at en sådan definition ikke lader sig udtrykke objektivt.
- På trods af et signal vides at indeholde et årligt signal, kan enkelte årlag være utydelige eller endog helt forsvundet på grund af diffusion (se appendix B), udtynding af lag på grund af isens flydning eller fordi det var markant mindre nedbør det pågældende år. Der er dog en hvis sandsynlighed for, at et utydeligt lag vil være tydeligere i en anden

dataserie, og det er derfor af essentiel betydning af anvende flere parametre samtidigt for at opnå en mere sikker identifikation af alle årlag. F.eks. vil et år med meget lidt nedbør give et svagt signal i  $\delta^{18}\text{O}$ -profilen, mens koncentrationen af tørdeponerede kemiske sporstoffer vil være upåvirket eller endda forhøjet. I mange eksempler på stratigrafisk datering har man derfor også anvendt multi-parameter-datering, men ofte foretages dateringen på baggrund af den årlige svingning i én dataserie, mens de andre parametre kun anvendes i tvivlstilfælde.

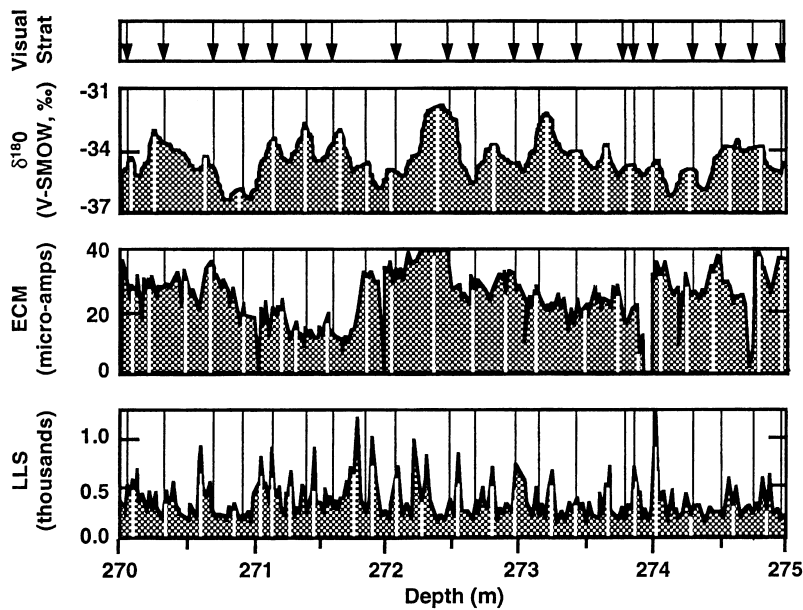
- [Rempel et al., 2001] har foreslået, at de betydelige temperaturgradienter, der findes i en iskappes indre, sammen med tilstedeværelsen af små mængder flydende vand langs iskrystallernes grænser, kan give anledning til en forskydning af de opløselige urenheder i forhold til den omkringliggende is. Denne forskydning har ingen praktisk betydning for de datasæt, jeg anvender i denne opgave, da målingernes opløsning er for grov og serierne stammer fra lille dybde, og jeg vil derfor ikke her omtale dette problem yderligere. Ved anvendelse af data fra den nederste halvdel af Indlandsisen, samt ved anvendelse af højopløste dataserier fra f.eks. CFA-målinger, kan denne process vise sig at være meget afgørende.

I det næste afsnit vil jeg beskrive en konkret anvendelse af stratigrafiske metoder. GISP2-iskernen blev dateret ved at årlagene blev identificeret ved hjælp af visuel stratigrafi. Dateringen blev i varierende grad understøttet af andre dataserier. Fremgangsmåden illustrerer nogle af de problemer, der i praksis forekommer ved stratigrafisk datering, især hvad angår reproducerbar identifikation af årlag ved multiparameter-datering. Samtidig er det den af de grønlandske iskerner, der er dateret længst tilbage i tiden ved hjælp af stratigrafiske metoder.

### 1.3.3 Dateringen af GISP2-iskernen

Som nævnt i indledningen til dette afsnit er det vanskeligt at datere iskerner ud fra identifikation af synlige årlag, men metoden blev ikke desto mindre anvendt til datering af den amerikanske GISP2-iskerne, der blev boret tæt ved og samtidig med den europæiske GRIP-iskerne. Principperne i dateringen er beskrevet i [Alley et al., 1997] og [Meese et al., 1997], og man skelner mellem datering af is fra henholdsvis Holocæn og den sidste istid, Wisconsin-istiden.

For Holocæn bygger metoden på en antagelse om, at de ændringer i primært boble-størrelse og -hyppighed, der kan observeres i kernen, er udtryk for en årstidsvariation, der kan bruges som basis for en absolut stratigrafisk datering. Årsagen til variationen er, at der i løbet af sommeren dannes et såkaldt *depth-hoar*-lag (hoar = rimfrost) på grund af solens kraftige påvirkning. Dette lag er kendetegnet ved større kornstørrelse og lavere densitet end vinter-sneen, der har mindre og mere regelmæssige krystaller. Firnifikation bevarer dette mønster på en sådan måde, at sommerlagene genfindes i isen som lag med færre og større luftbobler, der ifølge [Alley et al., 1997] klart kan identificeres i hele Holocæn-delen af iskernen. Selv i



Figur 1.5: Bestemmelse af årlag i GISP2-iskernen på basis af visuel stratigrafi og 3 målte signaler : støvindhold målt med Laserlys-spredning (LLS), ECM-signalet og isens  $\delta^{18}\text{O}$ -værdi. Bemærk det manglende visuelle årlag i dybden 271,85 meter, og den manglende  $\delta^{18}\text{O}$ -top ved ca. 273 meters dybde. Fra [Meese et al., 1997].

de dele af kernen, hvor luftboblerne har været opløst som ”gitter-luft”<sup>13</sup>, kan den årlige variation genfindes. Identifikationen af årlag skete primært på baggrund af denne variation i boble-størrelse og -hyppighed, men også ECM-signalet, måling af støvindhold (ved hjælp af laserlys-spredning) og detaljerede  $\delta^{18}\text{O}$ -sekvenser blev anvendt til at støtte dateringen og til at dokumentere, at boble-variationerne er et årligt signal. På figur 1.5 ses en sekvens, hvor dateringen er gennemført ved hjælp af 4 parametre. Bemærk, at der ved ca. 271,85 meters dybde findes et lag med stærk indikation af et årlag i både ECM,  $\delta^{18}\text{O}$  og støv-signalet, der ikke blev fundet som visuelt lag. Til gengæld findes et visuelt årlag i ca. 273 meters dybde, der ikke kan ses i  $\delta^{18}\text{O}$ -signalet, muligvis på grund af at et tyndt årlag vil være forsvundet fra  $\delta^{18}\text{O}$ -profilen ved diffusion. Figuren viser klart, at en præcis datering ikke kan forventes ved enkelt-parameter-datering. [Alley et al., 1997] vurderer selv præcisionen af dateringen af Holocæn-delen af iskernen til 1% eller bedre.

I istidsis er boble-signalet ikke så tydeligt, så her anvendes variationer i støvindholdet i isen indirekte. Ved måling på tidlig Holocæn-is konkluderes det, at der er sammenfald mellem højt støvindhold og forekomsten af synlige tågede bånd (cloudy bands). Disse synlige bånd bruges derefter til stratigrafisk datering (nogle steder igen støttet af støvmålinger og ECM) tilbage til ca. 50.000 år før nu med en anslået usikkerhed på højst 5%. Dateringen føres helt tilbage til

<sup>13</sup>Når trykket bliver tilstrækkeligt højt presses luftens molekyler ind i isens krystalgitter, og boblerne forsvinder derfor tilsyneladende. Processen - der på engelsk kaldes *clathrate-formation* - kan sammenlignes med gittervand i kemiske forbindelser. Under opbevaring af den borede kerne ved atmosfæretryk dukker boblerne op igen, ifølge [Alley et al., 1997] (og referencerne heri) på omtrent samme sted og med samme størrelsesfordeling.

Eem-tidens slutning ca. 110.000 år før nu, hvor præcisionen anslås til 20% [Meese *et al.*, 1997]. Under dateringen af især perioden, der ligger før 50.000 år før nu sammenlignes resultatet løbende med SPECMAP-tidsskalaen, der er afledt af havsediment-kerner, og man kan overveje i hvor høj grad dateringen af istids-isen kan siges at være uafhængig af SPECMAP-tidsskalaen. Kritikken af den fundne tidsskala kan sammenfattes i følgende :

- Selvom det sandsynliggøres at de synlige variationer i boble-størrelse og -hyppighed er årlige signaler, kan det meget vel tænkes at der i løbet af een sommer kan dannes mere end et lag, der muligvis senere kan tolkes som årlag. Tilsvarende er det ikke velunderbygget at tågede bånd er et "rent" årligt signal. Dateringen af istidsisen kan i øvrigt ikke reproduceres, idet de synlige lag forsvinder under opbevaringen af kernen og ikke er affotograferet eller på anden måde gengivet i sin helhed.
- Det illustreres på graferne i [Meese *et al.*, 1997], at der generelt tælles for mange årlag hvis ECM-signalet og støvindholdet alene anvendes til datering. En kritisk granskning af figur 1.5 og de andre eksemplificerede dataserier fra [Meese *et al.*, 1997] viser, at der generelt ikke er klare kriterier for, hvad der kendetegner et årlag i GISP2-dateringen, men at det tilsyneladende afgøres fra gang til gang hvad, der klassificeres som et år.
- 4 forskellige personer foretog optællingen, og i de fleste dele af kernen blev årlagene optalt af mindst to personer, dog ikke uafhængigt af hinanden, idet den anden person ofte kun verificerede den optælling, den første person havde foretaget. Det er derfor ikke oplagt, at definitionen af hvad der kendetegner et årlag, har været uændret i løbet af optællingen.
- Kun dele af kernen er dateret med multiparameter-metoder.

GISP2-tidsskalaen har da også været udsat for en del kritik. Dateringen af GRIP-iskernen giver en tidsskala, der ikke stemmer godt overens med GISP2-tidsskalaen, hvilket naturligvis er uheldigt når de to iskerner stammer fra praktisk talt samme sted på Indlandsisen, og derfor må repræsentere parallelle klimatiske signaler.

I [Wagner *et al.*, 2001] beskrives en metode til bestemmelse af årlagstykkelser, der er helt uafhængig af isens flydemønster. Nukliderne  $^{36}\text{Cl}$  og  $^{10}\text{Be}$  dannes i atmosfæren på grund af den kosmiske stråling. Intensiteten af den kosmiske stråling varierer på grund af variationer i Jordens magnetiske felt, der kan rekonstrueres ud fra blandt andet havsedimentkerner. Den heraf afledte dannelseshastighed antages at være proportional med koncentrationen, svarende til at  $^{36}\text{Cl}$  og  $^{10}\text{Be}$  har korte residenstider i atmosfæren. Den beregnede koncentration af  $^{36}\text{Cl}$  og  $^{10}\text{Be}$  sammenlignes med det målte indhold af nukliderne i GRIP-iskernen, hvorved der opnås en beregnet akkumulationsprofil, der passer godt med den akkumulationsprofil, der blev fundet ved modellering [Johnsen *et al.*, 1995], men afviger markant fra den akkumulationsprofil, der kan afledes af ovenstående datering af GISP2-iskernen. Dette understøtter påstanden om at GISP2-tidsskalaen ikke kan betragtes som sikker.

### 1.3.4 Behovet for en ny stratigrafisk tællemetode

Eksemplet illustrerer klart, at selv et ambitiøst dateringsprojekt som GISP2-dateringen bygger på et grundlag, der i høj grad kan diskuteres. På samme måde kan man betvivle andre iskerne-dateringer, der dog ofte i højere grad end GISP2-eksemplet kan reproducere. Et af hovedproblemerne er, at det ikke er oplagt at antallet af talte år i en del af iskernen kan reproducere ved to tællinger foretaget af samme person, ligesom det er yderst tvivlsomt om definitionen af et årlag er uafhængig af hvem, der foretager optællingen. Det vil derfor være ønskeligt hvis det var muligt at udvikle en metode til stratigrafisk datering, der giver reproducerbare resultater. Det vil naturligvis være ideelt hvis dateringen kunne baseres på objektive kriterier, men det er ikke klart hvordan sådanne kriterier skal opstilles. Som et realistisk gennemførligt alternativ vil jeg derfor i det følgende arbejde frem mod en dateringsmetode, der giver reproducerbare resultater, men som ikke er objektiv. Jeg vil søge at begrænse subjektiviteten til, at der undervejs vil være nogle principielle valg, der kan diskuteres, og nogle parametre, hvis værdier skal fastsættes. Det vil dermed være muligt at undersøge dateringens følsomhed overfor ændringer i antagelse og parametre, ligesom metodens resultater bliver uafhængige af hvem der udfører analysen.

## Del 2

# Dynamisk Dekorrelation af iskerndata

I dette kapitel vil jeg beskrive den metode, jeg har brugt til at udtrække information fra flere tidsserier fra samme iskerne. Målet er ud fra et antal dataserier fra samme dybdeinterval i samme iskerne at udtrække et signal, der udviser en mere tydelig årlig svingning end hvad de individuelle dataserier udviser. Først beskrives kort principalkomponent analyse (der oftest kaldes Principal Component Analysis eller blot PCA), der ofte anvendes til at finde underliggende mønstre i dataserier, og derefter beskrives en variant af Independent Component Analysis, der kaldes Dynamisk Dekorrelation. Sidstnævnte metode er et af de centrale elementer i denne opgave.

Fælles for PCA og Dynamisk Dekorrelation er antagelsen om, at de målte dataserier kan tænkes frembragt ved linearkombination af et antal underliggende kilde-serier. Man kunne forstille sig, at kilde-serierne kunne være

- En serie med stærk årlig variation, f.eks. koblet til den fysiske temperatursvingning.
- Et antal periodisk svingende signaler med forskellige perioder. Disse kunne være koblet til f.eks. variationer i solindstrålingen, solpletcyklen eller til svingninger i atmosfærens og havstrømmenes cirkulationsmønstre, eller repræsentere temperatursvingninger på korte tidsskalaer.
- Sporadiske signaler, der repræsenterer enkeltbegivenheder og ikke periodiske fænomener. Bidrag af produkter fra vulkanudbrud er typiske sporadiske bidrag.
- En eller flere støjkilder. Støj kan både skyldes måleproceduren, men kan også være såkaldt *depositional noise*, der f.eks. kan skyldes variationer i sneens indhold af sporstoffer i løbet af en nedbørsbegivenhed eller ujævn fordeling af sneen på overfladen [Steffensen et al., 1996].

Desuden kan indholdet af kemiske sporstoffer i isen tænkes rent fysisk at stamme fra forskellige kildeområder eller -processer, der kunne give anledning til en opsplitning i endnu flere kildeserier [Ditlevsen & Marsh, 1998].

Med henblik på datering vil det naturligvis være af stor interesse at kunne udtrække et underliggende årligt signal. Det er håbet, at et sådan signal vil udvise større klarhed i den årlige svingning end de målte serier udviser enkeltvis. Man kan også formulere det således, at vi søger den af de underliggende serier, der indeholder den bedst mulige årlige svingning.

Selvom jeg ovenfor antydede, at de underliggende serier kunne repræsentere reelle fysiske svingninger, behøver dette ikke at være tilfældet. Man kan også gå mere pragmatisk til værks og lade de målte serier være linearkombinationer af en række kildeserier, der blot opfattes som matematiske konstruktioner uden umiddelbar fysisk betydning. Uanset hvilken tilgang man anvender, er antagelsen dog den samme; nemlig at der kan findes en kildeserie med mere udtalt årlig svingning, end den der findes i de enkelte dataserier.

## 2.1 Behandling af dybdeserier som tidsserier

Den information om klimaet, man kan hente i en iskerne, er lagret ved en kombination af mange processer. Isotopforholdet i sneen bevares i isen, selvom diffusionsprocesser udviser de hurtige variationer (se appendix B), men også luftens isotopforhold er repræsenteret i isprøver, da små luftbobler indkapsles mellem iskrystallerne under firnifikationsprocessen. Desuden indeholder isen ioner og andre urenheder, der både kan stamme fra nedbør og kan være tørdeponeret.

De dataserier, der anvendes i dette speciale, stammer alle fra målinger på selve isen og ikke fra isens luftindhold. Isotopforholdene i isen er som nævnt bestemt af nedbørens isotopforhold, så de målte  $\delta^{18}\text{O}$ -værdier repræsenterer således kun information fra de tidsrum, hvor der var nedbør, og mellemliggende perioder vil ikke være dækket af data. Når man betragter en målt  $\delta^{18}\text{O}$  dataserie, er det derfor vigtigt at notere sig, at det målte signal repræsenterer en temperaturproxy, der kun indeholder data fra de tidspunkter, hvor der var nedbør, og man kan derfor ikke direkte aflede f.eks. den årlige middeltemperatur af  $\delta^{18}\text{O}$ -data, selvom man kendte sammenhængen mellem  $\delta^{18}\text{O}$ -værdien i nedbøren og temperaturen på nedbørstidspunktet præcist. Det målte signal repræsenterer nærmere en proxy for gennemsnittet af de temperaturer, der var mens nedbøren faldt, vægtet med nedbørsmængden.

Deponering af ioner og urenheder på isoverfladen sker som tidligere nævnt ikke kun under nedbør, og indholdet af ioner og urenheder vil derfor være mindre direkte koblet til nedbørsfordelingen.

Ved analyse af iskerner måles de forskellige parametre (isotopforhold, ionkoncentrationer, støvindhold, ledningsevne) som funktion af dybden, og datasierne repræsenterer derfor ikke tidsserier, men ”dybdeserier”. Ved anvendelse af f.eks. PCA er der ingen problemer forbundet hermed, da PCA ikke anvender tidslige egenskaber ved signalet, og den varierende opløsning er derfor uden betydning. Når man derimod ønsker at anvende værktøjer fra tidsserieanalyse, må



man derfor opfatte dataserierne som tidsserier med varierende (og ukendt) tidslig opløsning. Dynamisk Dekorrelation anvender egenskaber ved signalernes auto- og krydskorrelationsfunktioner, der indeholder tidslig information. Ideelt set burde dybdeserierne derfor konverteres til tidsserier, hvilket dog kræver nøjagtig information om fordelingen af den årlige akkumulation. [Meeker *et al.*, 1995] omtaler problemet med varierende tidslig opløsning i iskernesignaler, og anviser et udvalg af metoder til konvertering af dybdeserier til tidsserier. Fra et daterings-synspunkt er disse metoder dog af meget begrænset interesse, idet konverteringen kun kan foretages, hvis man har nøjagtige viden om fordelingen af den årlige akkumulation, hvilket forudsætter at dataserien allerede er dateret. Vi må således acceptere at anvende dybdeserier som tidsserier, og blot huske på at de anvendte dataseriers opløsning er ikke-uniform og kun i middel kendt.

I nogle tilfælde vil man på forhånd vide, at en given dataserie strækker sig hen over flere forskellige klimatiske situationer som f.eks. kolde stadialer med lav gennemsnitlig akkumulation og varmere interstadialer med højere gennemsnitlig akkumulation<sup>1</sup>. I så fald skal man være ekstra opmærksom på, at det er en dybdeserie man har målt, og enten forsøge sig med en konvertering af dybdeserien til en tidsserie, eller opdele datasættet og foretage analysen separat for hver klimatiske situation. Ved en opdelingen er det desuden muligt at tage højde for at sammensætningen af det målte signal fra forskellige kilder meget vel kan blive ændret fra én klimatiske periode til en anden.

Isens flydning betyder, at årlagenes tykkelse i middel må forventes at aftage indenfor en given dataserie (under forudsætning af at serien ikke, som ovenfor beskrevet, indeholder is fra perioder med markant forskellig middel-akkumulation). Hver prøve vil dermed i den nedre ende af serien i middel svare til et længere tidsrum end i den øvre ende, hvilket for lange serier vil kræve en korrektion, hvis dybdeserierne skal analyseres med tidsserie-metoder.

For med rimelighed at kunne anvende metoder, der anvender tidslige informationer fra dataserierne direkte på dybdeserier, må man derfor sikre sig følgende

- Data skal stamme fra en periode hvori der ikke forekommer voldsomme ændringer i akkumulationen. Man kan derfor ikke umiddelbart analysere dataserier, der strækker sig over både stadialer og interstadialer, med mindre dataserierne opdeles tilsvarende.
- Dataserierne skal opdeles i mindre stykker, således at udtyndingen af årlag på grund af isens flydning indenfor hver serie ikke ændrer den gennemsnitlige årlagstykkelse væsentligt. Den maksimale serielængde afhænger dermed af isens flydemønster på det relevante borested. Et overslag ved anvendelse af en DJ-model med typiske parameterverdier for Centralgrønland viser, at årlagstykkelsen kun ændres med et par procent indenfor en 50 meter lang dataserie.

---

<sup>1</sup>Under forrige istid forekom et antal meget abrupte klimatiske svingninger. De kolde perioder kaldes stadialer, mens de varme perioder, der typisk er af kortere varighed, kaldes interstadialer.

- Data skal stamme fra et sted, hvor nedbøren tidsligt er nogenlunde jævnt fordelt, set i forhold til datasættets tidslige opløsning, således at hver måling med rimelig tilnærmelse kan tænkes at svare til et konstant tidsrum. [Shuman et al., 1995] gør rede for at dette er en rimelig antagelse for Centralgrønland.

Ovenstående skal som nævnt være opfyldt for at man overhovedet kan anvende metoder, der udnytter tidslig information i dataserierne. Desuden skal man holde sig for øje, at den linearkombination af kildeserier vi søger ikke nødvendigvis er stationær, det vil sige konstant over tid. Dette er endnu en grund til, at de metoder, jeg vil beskrive i det følgende, kun bør anvendes på relativt korte dele af en iskerne. Analyse af en hel iskerne vil derfor enten skulle gennemføres ved en egentlig opdeling af data, eller ved at analysere data i et "vindue", der flytter sig gennem de fulde dataserier.

## 2.2 Indledende formalisme

Vi arrangerer de  $N$  målte dataserier - hver af længde  $T$  - som rækker i en  $N \times T$ -matrix  $\mathbf{X}$ . Antagelsen fra side 23 om at dataserierne fremkommer som linearkombinationer af nogle kildeserier kan da formuleres som

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (2.1)$$

hvor matricen  $\mathbf{S}$  indeholder de  $N$  kildeserier, og hvor linearkombinationernes koefficienter er samlet i matricen  $\mathbf{A}$ , der kaldes *blandingsmatricen*. Her har vi antaget at antallet af dataserier er identisk med antallet af kildeserier, hvilket vil blive diskuteret siden hen. Opgaven går ud på at rekonstruere de underliggende kildeserier (det vil sige sættet  $S$ , der er repræsenteret ved matricen  $\mathbf{S}$ ) ud fra vor viden om de målte data  $X$  (repræsenteret ved matricen  $\mathbf{X}$ ). Jeg vil i resten af opgaven forudsætte, at dataserierne såvel som kildeserier er normaliseret, så deres middelværdi er nul, samt skalere dataserierne således, at spredningen er 1.

Som problemet er formuleret i ligning 2.1, er det oplagt, at der skal være mere information til rådighed, før  $\mathbf{A}$  og  $\mathbf{S}$  kan bestemmes entydigt, idet kun  $\mathbf{X}$  kendes.

## 2.3 Principalkomponent Analyse - PCA

En mulig vej ud af problemet med utilstrækkelig information er at antage, at kildeserierne er indbyrdes orthogonale betraget som vektorer i det højdimensionale datarum, hvilket i dataserie-terminologi svarer til at kildeserierne er ukorreleerede<sup>2</sup>. Den første vektor vælges som den retning i data-rummet, hvor data udviser den største varians. Den næste basisvektor vælges vinkelret herpå i retningen, der maksimerer den resterende varians og så fremdeles. Man har derved beskrevet dataserierne i  $\mathbf{X}$  ved hjælp af en basis, hvor størstedelen af datasættets

<sup>2</sup>I det følgende vil jeg skifte mellem at beskrive data- og kildeserierne som tidsserier og som vektorer, idet hver beskrivelsesmåde har fordele og ulemper.

varians kan beskrives ved hjælp af de første basisvektorer (de principale komponenter). Hvis man fjerner bidraget fra de sidste basisvektorer taber man således ikke nogen stor del af variansen i datasættet, og PCA kan derfor med fordel anvendes til at reducere dimensionaliteten af store data-sæt.

PCA producerer som nævnt pr. definition et sæt af kildeserier, der er orthogonale. Det er ikke sandsynligt, at dette er tilfældet for de kildeserier, vi formoder at iskernedata er opbygget af, og PCA kan derfor ikke forventes at give et godt bud på sættet af kildeserier  $S$ . Desuden fokuserer PCA på at maksimere variansen i de første principalkomponenter, hvilket ikke behøver at være sammenfaldende med ønsket om at finde et stærkt årligt signal.

I denne opgave vil jeg derfor ikke anvende PCA til at finde kildeserierne, men istedet til at fjerne eventuel redundans i dataserierne. De nærmere detaljer vil blive beskrevet udførligt senere.

Man kan simpelt foretage PCA ved at diagonalisere korrelationsmatricen  $\mathbf{C}_X$ , der på plads  $(i, j)$  har korrelationskoefficienten mellem dataserie  $i$  og  $j$ . Der findes dog en numerisk mere robust metode, der også kan anvendes, hvis  $\mathbf{C}_X$  er singular (eller beregningsmæssigt tæt på at være det), svarende til at en eller flere af dataserierne er indbyrdes afhængige. Denne metode gør brug af en SVD-opløsning (singular value decomposition) af datamatrixen  $\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\top$ , hvor  $\mathbf{U}$  er en  $N \times N$  unitær matrix,  $\mathbf{V}$  er en  $T \times T$  unitær matrix og  $\mathbf{\Lambda}$  er en  $N \times T$  matrix med ikke-negative værdier - de singulære værdier - i "diagonalen" og nul ellers. Opløsningen er entydig, bortset fra permutationer af de singulære værdier og de tilhørende søjler i  $\mathbf{U}$  og  $\mathbf{V}$ , hvis de singulære værdier er forskellige.

Man identificerer matrixen  $\mathbf{U}$  som en blandingsmatrix, der angiver hvordan, dataserierne dannes ved linearkombination af kildeserierne, der opfattes som produktet  $\mathbf{\Lambda}\mathbf{V}^\top$ . Da  $\mathbf{V}$  er unitær, er kildeserierne varianser angivet i diagonalen i matrixen  $\mathbf{\Lambda}$ .

Hvis opløsningen giver singulære værdier der er nul, er der redundans i data. Er en eller flere af de singulære værdier markant lavere end de resterende værdier, kan man vælge at sætte den/dem lig nul. Dermed reduceres antallet af kildeserier, der bidrager til dataserierne, uden at der tabes væsentlig information. Denne procedure reducerer antallet af dimensioner, og sikrer at der tabes så lidt som muligt i forhold til det originale signal (i  $\mathcal{L}^2$ -forstand) [*Bishop*, 1995, side 310 - 313].

Det er almindeligt at man efterfølgende beskærer matrixerne i SVD-opløsningen, således at kun det øvre venstre hjørne af  $\mathbf{\Lambda}$  beholdes, mens de rækker og søjler der kun indeholder nuller stryges. Tilsvarende må der fjernes søjler fra højre side af  $\mathbf{U}$  og  $\mathbf{V}$ , således at kun de søjler, der bidrager til  $\mathbf{X}$ , bibeholdes.

## 2.4 Dynamisk Dekorrelation

Når man forlader antagelsen om indbyrdes orthogonale kildeserier, må man søge anden information for entydigt at kunne bestemme blandingsmatricen og kildeserierne. Metoder, der anvender den mindre vidtgående antagelse om at kildeserierne er statistisk uafhængige, kaldes Indterpende Component Analyses (ICA), men denne antagelse giver ikke i sig selv nok information til, at man kan bestemme kildeserierne. Yderligere antagelser må vælges ud fra den konkrete fysiske sammenhæng, der studeres. I [Hyvärinen, 1999] findes en oversigt over mulige antagelser, og en kort diskussion af hvornår disse antagelser kan tænkes opfyldt.

Som beskrevet på side 23 antager vi, at kildeserierne er forskellige (kvasi-)periodiske signaler, sporadiske elementer samt støj. Karakteristisk for disse serier er, at de har meget forskellige autokorrelationsfunktioner, og det vil derfor være naturligt at udnytte denne information til at bestemme  $S$ . Ideen med at anvende den tidslige information fra autokorrelationsfunktionerne blev introduceret af [Molgedey & Schuster, 1994] og er siden blevet yderligere udviklet af blandt andet Lars Kai Hansen fra DTU [Hansen et al., 2000]. Den grundlæggende idé er, at man antager at kildeseriernes autokorrelationsfunktioner  $c(\tau)$  er forskellige for en værdi af  $\tau$ , der dog ikke i analysen kan beregnes, men må vælges ud fra andre kriterier. Med denne information er det muligt entydigt at bestemme kildeserierne. Metodens grundprincipper står klart frem for uendeligt lange dataserier, og jeg vil derfor starte med dette tilfælde, omend dets praktiske betydning er begrænset.

### 2.4.1 Dynamisk Dekorrelation for uendelige serier

Vi betragter igen de  $N$  målte dataserier og danner korrelationsmatricen  $\mathbf{C}_X(\tau)$ , der som ovenfor beskrevet på den  $(i, j)$ 'te plads indeholder korrelationskoefficienten  $c_{X_{ij}}(\tau)$  for dataserierne  $x_{i,t}$  og  $x_{j,t}$ ,  $t \in \mathbb{Z}$  :

$$c_{X_{ij}}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T+1} \sum_{t=-T}^T x_{i,t} x_{j,t+\tau} \quad (2.2)$$

For at fjerne eventuel redundans i data foretages nu en principalkomponent analyse. Den ovenfor skitserede SVD-baserede metode til at udføre PCA kan ikke anvendes, idet matricen  $\mathbf{X}$  har uendeligt mange søjler, så istedet diagonaliseres  $\mathbf{C}_X(0)$  ved, at der bestemmes en diagonal matrix  $\mathbf{L}$  og en orthogonal matrix  $\mathbf{P}$  så  $\mathbf{C}_X(0) = \mathbf{P}\mathbf{L}\mathbf{P}^\top$ . Er nogle af egenverdierne i  $\mathbf{L}$  nul, fjernes de tilsvarende søjler i  $\mathbf{P}$ , hvorved matricen  $\tilde{\mathbf{P}}$  dannes. Denne kan derefter bruges til at projicere data ned på et underrum, der har maksimal rang :

$$\tilde{\mathbf{X}} = \tilde{\mathbf{P}}^\top \mathbf{X} \quad (2.3)$$

Antallet af kildeserier blev tidligere af hensyn til argumentationens overskuelighed sat til at være identisk med antallet af dataserier. Lad os tænke os at de  $\tilde{N}$  lineært uafhængige dataserier i matricen  $\tilde{\mathbf{X}}$  kan frembringes som linearkombinationer af et antal  $M$  lineært uafhængige kildeserier. Der må gælde, at  $\tilde{N} \leq M$ , idet man ikke kan danne flere end  $M$  lineært uafhængige

serier ud fra de  $M$  kildeserier. Er  $M = \tilde{N}$ , er vi i den ovenfor beskrevne situation, hvor man med antagelsen om forskellige autokorrelationsfunktioner for kildeserierne kan bestemme de  $M$  kildeserier fra samme antal uafhængige dataserier. Er  $M > \tilde{N}$ , er problemet underbestemt, og man kan på grund af manglende information ikke bestemme mere end  $\tilde{N}$  kildeserier.  $\tilde{N}$  serier er desuden tilstrækkeligt til at udspænde rummet, der udspændes af de  $\tilde{N}$  serier i  $\tilde{\mathbf{X}}$ . man kan derfor uden tab af generalisation skære antallet af kildeserier ned fra  $M$  til  $\tilde{N}$ , hvorved man er tilbage ved den beskrevne situation. At lade  $S$  bestå af  $\tilde{N}$  serier er dermed ikke en indskrænkelse i forhold til den generelle situation, men en konsekvens af de grundlæggende antagelser.

Vi definerer  $\mathbf{C}_{\tilde{\mathbf{X}}}(\tau)$  i analogi med definition 2.2, og danner nu den såkaldte *kvotientmatrix*

$$\mathbf{Q} = \mathbf{C}_{\tilde{\mathbf{X}}}(\tau)\mathbf{C}_{\tilde{\mathbf{X}}}(0)^{-1} \quad (2.4)$$

Ved at anvende den at  $\mathbf{X} = \mathbf{A}\mathbf{S}$  sammen med definitionen på  $\tilde{\mathbf{X}}$  fås nu at  $\tilde{\mathbf{X}} = \tilde{\mathbf{P}}^\top \mathbf{A}\mathbf{S}$  og dermed at  $\mathbf{C}_{\tilde{\mathbf{X}}}(\tau) = \tilde{\mathbf{P}}^\top \mathbf{A}\mathbf{C}_S(\tau)\mathbf{A}^\top \tilde{\mathbf{P}}$ . Indsættes dette resultat i definitionen på  $\mathbf{Q}$ , fås

$$\mathbf{Q} = (\tilde{\mathbf{P}}^\top \mathbf{A})\mathbf{C}_S(\tau)\mathbf{C}_S(0)^{-1}(\tilde{\mathbf{P}}^\top \mathbf{A})^{-1} \quad (2.5)$$

En af de grundlæggende antagelser om kildeserierne i  $S$  er, at de er statistisk uafhængige. Krydskorrelationskoefficienterne for to kildeserier  $c_{S_i}$  og  $c_{S_j}$ ,  $i \neq j$  fremkommer jævnfør ligning 2.2 som grænseværdien

$$c_{S_{ij}}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T+1} \sum_{t=-T}^T s_{i,t} s_{j,t+\tau} \quad (2.6)$$

Kildeserierne er ifølge antagelsen statistisk uafhængige, og har middelværdi 0. Uafhængigheden af sandsynlighedsfordelingerne for  $s_{i,t_0}$  og  $s_{j,t_0}$  for  $i \neq j$ ,  $t_0 \in \mathbb{Z}$  betyder at summen i ligning 2.6 er en sum af led der i gennemsnit er 0, hvorfor også grænseværdien er nul. Dette betyder, at  $\mathbf{C}_S(\tau)$ , og derfor specielt  $\mathbf{C}_S(0)$ , er diagonale matricer, hvorved kvotientmatricen kan skrives som

$$\mathbf{Q} = (\tilde{\mathbf{P}}^\top \mathbf{A}) \begin{bmatrix} \frac{c_{S_{11}}(\tau)}{c_{S_{11}}(0)} & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & 0 & \frac{c_{S_{\tilde{N}\tilde{N}}}(\tau)}{c_{S_{\tilde{N}\tilde{N}}}(0)} \end{bmatrix} (\tilde{\mathbf{P}}^\top \mathbf{A})^{-1} \quad (2.7)$$

For kompakthedens skyld indføres  $\Sigma \equiv \mathbf{C}_S(\tau)\mathbf{C}_S(0)^{-1}$ , hvorved vi kan skrive

$$\mathbf{Q} = (\tilde{\mathbf{P}}^\top \mathbf{A}) \Sigma (\tilde{\mathbf{P}}^\top \mathbf{A})^{-1} \quad (2.8)$$

Produktet i ligning 2.7 (eller 2.8) minder meget om en egen værdi-opløsning af  $\mathbf{Q}$ , og det er netop denne egenskab ved  $\mathbf{Q}$ , der er hjørnестenen i Dynamisk Dekorrelation. Ved at udnytte denne tilsyneladende lighed kan man bestemme  $\tilde{\mathbf{P}}^\top \mathbf{A}$  (og dermed  $\mathbf{A}$ , idet  $\tilde{\mathbf{P}}$  kendes). Metodens grundtræk er som følger :

- $\mathbf{Q}$  beregnes ud fra data og ved diagonalisering bestemmes en egenværdi-opløsning

$$\mathbf{Q} = \mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1} \quad (2.9)$$

Lad os indtil videre tænke os at egenværdierne i  $\mathbf{\Omega}$  alle er reelle og forskellige.

- Ved at sammenligne ligning 2.7 med ligning 2.9 identificeres  $\tilde{\mathbf{P}}^\top \mathbf{A}$  med  $\mathbf{\Phi}$ . Der gælder dog ikke lighed mellem  $\tilde{\mathbf{P}}^\top \mathbf{A}$  og  $\mathbf{\Phi}$  idet der kan være forskel på skaleringen af de enkelte matricer i  $\mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$  og højresiden af ligning 2.7, selvom begge produkter er lig med  $\mathbf{Q}$ . Desuden kan søjlerne i  $\mathbf{\Phi}$  samt de tilhørende egenværdier i  $\mathbf{\Omega}$  permuteres, uden at det påvirker produktet. Der er ikke taget højde for hverken permutation eller skalering i de følgende simple udtryk for  $\mathbf{A}$  og  $\mathbf{S}$ .

- Da  $\tilde{\mathbf{P}}$  er orthogonal, er

$$\mathbf{A} = \left(\tilde{\mathbf{P}}^\top\right)^{-1} \mathbf{\Phi} = \tilde{\mathbf{P}} \mathbf{\Phi} \quad (2.10)$$

- Når blandingsmatricen kendes, kan kildeserierne bestemmes. Da  $\mathbf{A}$  ikke nødvendigvis er invertibel anvendes at  $\tilde{\mathbf{X}} = \tilde{\mathbf{P}}^\top \mathbf{A} \mathbf{S}$ , hvilket giver at  $\mathbf{S} = \left(\tilde{\mathbf{P}}^\top \mathbf{A}\right)^{-1} \tilde{\mathbf{X}}$ . Ved at bruge resultatet fra ligning 2.10 findes at

$$\mathbf{S} = \left(\tilde{\mathbf{P}}^\top \mathbf{A}\right)^{-1} \tilde{\mathbf{X}} = \mathbf{\Phi}^{-1} \tilde{\mathbf{X}} = \mathbf{\Phi}^{-1} \tilde{\mathbf{P}}^\top \mathbf{X} \quad (2.11)$$

Problemet med skalering og permutationer er af rent teknisk art, og vil blive omtalt særskilt i afsnit 2.4.2. Beviset for at der altid vil eksistere en egenværdi-opløsning af  $\mathbf{Q}$  med reelle egenværdier kan findes i [Hansen *et al.*, 2000]. Antagelsen om forskellige egenværdier i  $\mathbf{\Omega}$  skal derimod gøres til genstand for diskussion i det følgende.

Så længe egenværdierne er forskellige, er egenværdi-opløsningen entydigt fastlagt (igen: bortset fra permutation af egenværdier og tilhørende søjler), og blandingsmatricen kan dermed bestemmes entydigt ved ligning 2.10. Er to egenværdier derimod ens, vil de tilhørende egenvektorer udspænde et underrum, hvori enhver vektor<sup>3</sup> vil være egenvektor med den pågældende egenværdi. Da de nævnte egenvektorer netop er søjlerne i matricen  $\mathbf{\Phi}$ , er denne ikke entydigt bestemt, hvis egenværdierne i  $\mathbf{\Omega}$  er degenererede, hvorfor hverken blandingsmatrix eller kildeserier kan bestemmes entydigt.

Er to egenværdier tæt på at være identiske, opstår der numeriske problemer med egenværdi-opløsningen, idet små fejl i den beregnede kvotientmatrix, forårsaget af f.eks. måleusikkerheder eller støj, giver meget forskellige egenværdi-opløsninger.

Det er her, at den grundlæggende antagelse om kildeseriernes forskellige autokorrelationsfunktioner<sup>4</sup> kommer ind i billedet : ved at vælge en værdi af  $\tau$ , hvor autokorrelationsfunktionerne

<sup>3</sup>Husk, at egenvektorer og kildeserier er det samme i denne sammenhæng

<sup>4</sup>Det er de formelt set de *normerede* autokorrelationsfunktioner  $\frac{c_{S_{ii}}(\tau)}{c_{S_{ii}}(0)}$ , der skal være forskellige, men jeg vil i det følgende udelade normeringen, der er helt triviell.

for de forskellige kildeserier er tilstrækkeligt forskellige, kan vi sikre entydigheden af de beregnede kildeserier. Det er dog desværre ikke umiddelbart muligt a priori at udpege en velegnet værdi for  $\tau$ , hvorfor en hvis grad af "trial and error" må påregnes før man finder et godt valg af  $\tau$ .

### 2.4.2 Permutationer og skalering i egenværdi-opløsningen af $\mathbf{Q}$

Som beskrevet er der ikke lighed i matematisk forstand mellem de enkelte matricer i ligning 2.8 og matricerne i egenværdi-opløsningen  $\mathbf{Q} = \mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$ . Specielt vil søjlerne i den ved diagonalisering producerede matrix  $\mathbf{\Phi}$  have norm 1, mens produktet  $\mathbf{P}^T \mathbf{A}$  ikke generelt vil have denne egenskab. Desuden kan man permutere søjlerne i  $\mathbf{\Phi}$  og de tilhørende egenværdier i  $\mathbf{\Omega}$  uden at det ændrer det samlede resultat. I det følgende vil jeg vise hvordan der kan opnås lighed mellem de enkelte matricer i de to udtryk  $\mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$  og  $(\tilde{\mathbf{P}}^T \mathbf{A}) \mathbf{\Sigma} (\tilde{\mathbf{P}}^T \mathbf{A})^{-1}$ , der begge er lig med  $\mathbf{Q}$ .

Vi definerer først en permutationsmatrix  $\mathbf{\Pi}$  og en skaleringsmatrix  $\mathbf{\Xi}$ .  $\mathbf{\Pi}$  defineres ved at rækkerne i en enhedsmatrix af samme dimension som  $\mathbf{\Omega}$  (det vil sige  $\tilde{N} \times \tilde{N}$ ) permuteres. Det er oplagt at  $\mathbf{\Pi}$  er invertibel og at  $\mathbf{\Pi}\mathbf{\Pi}^{-1} = \mathbf{\Pi}^{-1}\mathbf{\Pi} = \mathbf{I}$ . Skaleringsmatricen  $\mathbf{\Xi}$  defineres som en  $\tilde{N} \times \tilde{N}$  diagonal matrix  $\mathbf{\Xi} = \text{diag}([\xi_1, \xi_2, \dots, \xi_{\tilde{N}}])$  med skaleringskonstanter  $\xi_i \neq 0, i \in \{1, \dots, \tilde{N}\}$ . Da den inverse af en diagonal matrix findes ved at erstatte diagonal-elementerne med disses reciprokke værdier, og idet produktet af diagonale matricer bestemmes ved elementvis multiplikation af diagonal-elementerne, er  $\mathbf{\Xi}\mathbf{\Omega}\mathbf{\Xi}^{-1} = \mathbf{\Omega}$ , hvor vi har anvendt at  $\mathbf{\Omega}$  er diagonal. Da produktet i ligning 2.8 er lig med egenværdi-opløsningen  $\mathbf{Q} = \mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$  har vi at

$$\begin{aligned} (\tilde{\mathbf{P}}^T \mathbf{A}) \mathbf{\Sigma} (\tilde{\mathbf{P}}^T \mathbf{A})^{-1} &= \mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1} \\ &= \mathbf{\Phi}\mathbf{\Xi}\mathbf{\Omega}\mathbf{\Xi}^{-1}\mathbf{\Phi}^{-1} \\ &= \mathbf{\Phi}\mathbf{\Xi}\mathbf{\Pi}\mathbf{\Pi}^{-1}\mathbf{\Omega}\mathbf{\Pi}\mathbf{\Pi}^{-1}\mathbf{\Xi}^{-1}\mathbf{\Phi}^{-1} \\ &= (\mathbf{\Phi}\mathbf{\Xi}\mathbf{\Pi}) \mathbf{\Pi}^{-1}\mathbf{\Omega}\mathbf{\Pi} (\mathbf{\Phi}\mathbf{\Xi}\mathbf{\Pi})^{-1} \end{aligned} \quad (2.12)$$

I analogi med fremgangsmåden i afsnit 2.4.1 identificeres matricerne på højre og venstre side nu med hinanden. Det nye er i denne forbindelse, at vi ved at vælge  $\mathbf{\Pi}$  og  $\mathbf{\Xi}$  kan opnå lighed mellem matricerne :

$$\tilde{\mathbf{P}}^T \mathbf{A} = \mathbf{\Phi}\mathbf{\Xi}\mathbf{\Pi} \quad \text{og} \quad \mathbf{\Sigma} = \mathbf{\Pi}^{-1}\mathbf{\Omega}\mathbf{\Pi} \quad (2.13)$$

Herefter kan udtryk analoge til ligning 2.10 og 2.11 opstilles og kildeserierne kan dermed bestemmes.

Det skal indskydes, at denne procedure ingen praktisk betydning har, da kildeseriernes rækkefølge under alle omstændigheder er vilkårlig, og en skalering af serierne er uden praktisk betydning. Jeg synes imidlertid at metodens grundlæggende princip, som netop udtrykkes matematisk stringent i ligning 2.13, illustreres bedst på denne måde.

### 2.4.3 Dynamisk Dekorrelation for endelige serier

Lad os nu forlade tilfældet med uendeligt lange dataserier til fordel for den situation, der forekommer ved analyse af virkelige datasæt. Da dataserierne nu har endelig længde  $T$ , er matricen  $\mathbf{X}$  af dimension  $N \times T$ , og vi kan derfor foretage en SVD-opløsning :

$$\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\top \quad (2.14)$$

Af  $\mathbf{\Lambda}$  vil det fremgå om der er redundans i data, idet der i bekræftende fald er et eller flere af  $\mathbf{\Lambda}$ 's diagonal-elementer, der er nul. Er dette tilfældet beskæres de tre matricer, således at de rækker og søjler der ikke bidrager til det samlede produkt fjernes. For at undgå at indføre endnu et sæt symboler vil jeg lade matricerne i ligning 2.14 betegne de beskårne matricer.

Matricen  $\mathbf{U}$  fra ligning 2.14 spiller samme rolle som  $\tilde{\mathbf{P}}$  gjorde i tilfældet med uendelige dataserier. Herved opnås et meget simpelt udtryk for de projicerede data, idet vi i analogi med ligning 2.3 definerer  $\tilde{\mathbf{X}} = \mathbf{\Lambda}\mathbf{V}^\top$ .

Korrelationsfunktionerne  $c_{\tilde{X}_{ij}}(\tau)$  er ikke veldefinerede for endelige dataserier, idet der er problemer i seriernes ender. For at illustrere dette problem betragtes først to uendelige dataserier  $\tilde{x}_{i,t}$  og  $\tilde{x}_{j,t}$  :

$$\begin{array}{cccccc} \dots & \tilde{x}_{i,-2} & \tilde{x}_{i,-1} & \tilde{x}_{i,0} & \tilde{x}_{i,1} & \tilde{x}_{i,2} & \dots \\ \dots & \tilde{x}_{j,-2} & \tilde{x}_{j,-1} & \tilde{x}_{j,0} & \tilde{x}_{j,1} & \tilde{x}_{j,2} & \dots \end{array} \quad (2.15)$$

Korrelationskoefficienten  $c_{\tilde{X}_{ij}}(0)$  beregnes ved at elementer, der står over hinanden, multipliceres, hvorefter gennemsnittet af alle disse produkter beregnes. Når  $c_{\tilde{X}_{ij}}(\tau)$  beregnes, flyttes en af serierne  $\tau$  pladser til siden hvorefter der multipliceres og tages gennemsnit som ovenfor beskrevet. I tilfældet  $\tau = 2$  kan man illustrere parringen af led således :

$$\begin{array}{cccccc} \dots & \tilde{x}_{i,0} & \tilde{x}_{i,1} & \tilde{x}_{i,2} & \tilde{x}_{i,3} & \tilde{x}_{i,4} & \dots \\ \dots & \tilde{x}_{j,-2} & \tilde{x}_{j,-1} & \tilde{x}_{j,0} & \tilde{x}_{j,1} & \tilde{x}_{j,2} & \dots \end{array} \quad (2.16)$$

Når serierne derimod har endelig længde  $T$  vil en tilsvarende opskrivning give

$$\begin{array}{cccccc} \tilde{x}_{i,1} & \tilde{x}_{i,2} & \tilde{x}_{i,3} & \tilde{x}_{i,4} & \dots & \tilde{x}_{i,T-1} & \tilde{x}_{i,T} \\ & & \tilde{x}_{j,1} & \tilde{x}_{j,2} & \dots & \tilde{x}_{j,T-3} & \tilde{x}_{j,T-2} & \tilde{x}_{j,T-1} & \tilde{x}_{j,T} \end{array} \quad (2.17)$$

hvoraf det ses, at de 2 første elementer i  $\tilde{x}_{i,t}$  og de to sidste i  $\tilde{x}_{j,t}$  ikke umiddelbart kan danne par med et element fra den anden serie. Man kan bestemme en tilnærmet værdi til  $c_{\tilde{X}_{ij}}(\tau)$  ved at se bort fra disse uparrede værdier, men vi vil her anvende en anden tilnærmelse til  $c_{\tilde{X}_{ij}}(\tau)$ , hvor vi istedet for at rykke serierne  $\tau$  pladser i forhold til hinanden vil permutere den ene serie cyklisk  $\tau$  pladser. Ovenstående eksempel, hvor  $\tau = 2$ , kan da illustreres således :

$$\begin{array}{cccccc} \tilde{x}_{i,1} & \tilde{x}_{i,2} & \tilde{x}_{i,3} & \dots & \tilde{x}_{i,T-1} & \tilde{x}_{i,T} \\ \tilde{x}_{j,T-1} & \tilde{x}_{j,T} & \tilde{x}_{j,1} & \dots & \tilde{x}_{j,T-3} & \tilde{x}_{j,T-2} \end{array} \quad (2.18)$$



Denne tilnærmelse må antages at være god så længe  $\tau \ll T$ . For at formulere beregningen af de tilnærmede korrelationskoefficienter ved matrix-regning indføres matricen  $\tilde{\mathbf{X}}_\tau$ , der på plads  $(n, t)$  har elementet, der står på plads  $(n, (t + \tau - 1)_T + 1)$  i  $\tilde{\mathbf{X}}$ , idet  $(t + \tau - 1)_T$  betegner værdien af  $t + \tau - 1$  modulo  $T$ . I sin korteste form er definitionen således  $\tilde{\mathbf{X}}_\tau(n, t) = \tilde{\mathbf{X}}_{(n, (t + \tau - 1)_T + 1)}$ . Ved anvendelse af  $\tilde{\mathbf{X}}_\tau$  kan den omtalte tilnærmelse til  $c_{\tilde{X}_{ij}}(\tau)$  skrives som

$$\tilde{c}_{\tilde{X}_{ij}}(\tau) = \frac{1}{T} \tilde{\mathbf{X}} \tilde{\mathbf{X}}_\tau^\top \quad (2.19)$$

I modsætning til  $\tilde{\mathbf{C}}_X(\tau)$  vil den ved ligning 2.19 bestemte korrelationsmatrix  $\tilde{\mathbf{C}}_{\tilde{X}}(\tau)$  ikke generelt være symmetrisk, hvilket giver problemer med kvotientmatrixens egenværdiopløsning, der ikke kan garanteres udelukkende at have reelle egenværdier. Derfor defineres en ”symmetriseret” udgave af  $\tilde{\mathbf{C}}_{\tilde{X}}(\tau)$  givet ved

$$\hat{\mathbf{C}}_{\tilde{X}}(\tau) = \frac{1}{2T} \left( \tilde{\mathbf{X}}_\tau \tilde{\mathbf{X}}^\top + \tilde{\mathbf{X}} \tilde{\mathbf{X}}_\tau^\top \right) \quad (2.20)$$

som vi anvender til at beregne den estimerede kvotientmatrix  $\hat{\mathbf{Q}}$

$$\begin{aligned} \hat{\mathbf{Q}} &= \hat{\mathbf{C}}_{\tilde{X}}(\tau) \hat{\mathbf{C}}_{\tilde{X}}(0)^{-1} \\ &= \frac{1}{2} \left( \tilde{\mathbf{X}}_\tau \tilde{\mathbf{X}}^\top + \tilde{\mathbf{X}} \tilde{\mathbf{X}}_\tau^\top \right) \left( \tilde{\mathbf{X}} \tilde{\mathbf{X}}^\top \right)^{-1} \end{aligned} \quad (2.21)$$

$\hat{\mathbf{Q}}$  er oplagt symmetrisk, og det kan vises, at den også er positiv definit, hvilket garanterer eksistensen af en egenværdiopløsning med udelukkende reelle egenværdier [Hansen et al., 2000]. Ved at anvende at  $\tilde{\mathbf{X}} = \mathbf{\Lambda} \mathbf{V}^\top$  og at  $\mathbf{V}$  er en unitær matrix fås

$$\begin{aligned} \hat{\mathbf{Q}} &= \frac{1}{2} \mathbf{\Lambda} \left( \mathbf{V}_\tau^\top \mathbf{V} + \mathbf{V}^\top \mathbf{V}_\tau \right) \mathbf{\Lambda} \left( \mathbf{\Lambda} \mathbf{V}^\top \mathbf{V} \mathbf{\Lambda} \right)^{-1} \\ &= \frac{1}{2} \mathbf{\Lambda} \left( \mathbf{V}_\tau^\top \mathbf{V} + \mathbf{V}^\top \mathbf{V}_\tau \right) \mathbf{\Lambda}^{-1} \end{aligned} \quad (2.22)$$

hvor vi har defineret  $\mathbf{V}_\tau$  ved permutation i analogi med  $\tilde{\mathbf{X}}_\tau$ . Ved hjælp af denne ligning kan  $\hat{\mathbf{Q}}$  umiddelbart beregnes fra de kendte datasier.

Vi vil nu i overensstemmelse med fremgangsmåden i afsnit 2.4.1 opskrive andre udtryk for  $\hat{\mathbf{Q}}$  for gennem sammenligning med ligning 2.22 at kunne bestemme den estimerede blandingsmatrix og derved beregne kildeserierne. Den for uendeligt lange serier anvendte projektionsmatrix  $\tilde{\mathbf{P}}$  er som tidligere beskrevet erstattet af  $\mathbf{U}$  fra ligning 2.14. Ligning 2.8 ændres derfor for endelige datasier til

$$\hat{\mathbf{Q}} = \left( \mathbf{U}^\top \mathbf{A} \right) \tilde{\mathbf{\Sigma}} \left( \mathbf{U}^\top \mathbf{A} \right)^{-1} \quad (2.23)$$

Da vi kun har antaget, at kildeserierne er statistisk uafhængige, og ikke ukorreleerede som i PCA, vil de ikke-diagonale elementer ikke generelt være nul, som det var tilfældet for uendeligt lange serier. Dog vil de i de fleste tilfælde være små i sammenligning med diagonal-elementerne, så længe serielængden  $T$  er af anseelig størrelse. En række numeriske forsøg med tilfældigt genererede normaliserede serier viste, at ved en serielængde på  $T = 100$  elementer havde de

ikke-diagonale elementer i  $\tilde{\Sigma} = \tilde{\mathbf{C}}_S(\tau)\tilde{\mathbf{C}}_S(0)^{-1}$  en middelværdi i størrelsesordenen  $10^{-3}$  med en spredning på omkring 0,1, og at spredningen (ikke overraskende) varierede som  $T^{-\frac{1}{2}}$  for andre serielængder. Til trods for at  $\tilde{\Sigma}$  kun er tilnærmelsesvis diagonal, vil vi beregne en estimeret blandingsmatrix  $\hat{\mathbf{A}}$  og matricen med de estimerede kildeserier  $\hat{\mathbf{S}}$  ved at identificere matricerne i opløsningen i ligning 2.23 med en egen værdi-opløsning  $\hat{\mathbf{Q}} = \mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$ . Herved identificeres  $\mathbf{U}^T\hat{\mathbf{A}}$  med  $\mathbf{\Phi}$ , hvoraf det ses at

$$\hat{\mathbf{A}} = \mathbf{U}\mathbf{\Phi} \quad (2.24)$$

I analogi med 2.10 bestemmes kildeserierne ud fra ligningen  $\tilde{\mathbf{X}} = \mathbf{U}^T\mathbf{A}\mathbf{S}$ , hvor både  $\mathbf{A}$  og  $\mathbf{S}$  erstattes af deres estimerede varianter

$$\begin{aligned} \hat{\mathbf{S}} &= (\mathbf{U}^T\hat{\mathbf{A}})^{-1}\tilde{\mathbf{X}} \\ &= \mathbf{\Phi}^{-1}\tilde{\mathbf{X}} \\ &= \mathbf{\Phi}^{-1}\mathbf{\Lambda}\mathbf{V}^T \end{aligned} \quad (2.25)$$

Igen er det vigtigt at diagonal-elementerne i  $\tilde{\Sigma}$  har forskellige værdier, for at de beregnede kildeserier kan bestemmes entydigt (dog stadig med undtagelse af skalering og permutation). Man må derfor sikre sig at  $\tau$  vælges således at diagonal-elementerne er forskellige, hvilket er muligt ifølge den grundlæggende antagelse om at kildeseriernes autokorrelationsfunktioner er forskellige for mindst én værdi af  $\tau$ . Der findes dog ikke nogen måde til på forhånd at afgøre, hvilken værdi af  $\tau$  der er optimal, men for en given værdi af  $\tau$  kan man betragte diagonal-elementerne i matricen  $\mathbf{\Omega}$ , der fremkommer ved egen værdi-opløsningen af  $\hat{\mathbf{Q}}$ . Fås ikke tilstrækkeligt forskellige værdier, vælges en anden værdi af  $\tau$ , og analysen gentages. Da beregningsarbejdet i analysen er overkommeligt ved hjælp af (især matrix-baserede) ”programmeringssprog” som f.eks. IDL eller MatLab, er en sådan ”trial and error” procedure praktisk gennemførlig. I afsnit 4.2.1 beskrives i detaljer hvordan jeg forsøgte at vælge  $\tau$  på denne måde, og netop betydningen af valget af  $\tau$  vil blive diskuteret indgående i de følgende kapitler.

#### 2.4.4 Anvendelse af flere samtidige værdier af $\tau$

Som beskrevet er valget af  $\tau$  af essentiel betydning for metodens resultater. Kun når egen værdi-opløsningen  $\mathbf{\Phi}\mathbf{\Omega}\mathbf{\Phi}^{-1}$  af  $\mathbf{Q}$  (eller  $\hat{\mathbf{Q}}$ ) giver en matrix  $\mathbf{\Omega}$  med forskellige egen værdier, er  $\mathbf{\Phi}$  entydigt fastlagt, hvilket naturligvis er en forudsætning for at blandingsmatricen og kildeserierne entydigt kan beregnes. I tilfælde af degenererede egen værdier bliver de kildeserier, hvis  $c_S(\tau)$ -værdier er ens, ikke separeret, men erstattes af tilfældige linearkombinationer heraf.

Lad os nu tænke os et eksempel, hvor vi er interesserede i at bestemme 3 kildeserier, men vi ikke kan finde en værdi af  $\tau$ , der giver tilstrækkeligt forskellige egen værdier for  $\mathbf{\Omega}$ . Vi tænker os dog, at der findes en værdi  $\tau_1$ , der giver god separation af kildeserie 1 i forhold til kildeserie 2 og 3, der indbyrdes ikke kan adskilles fordi deres tilhørende egen værdier i  $\mathbf{\Omega}$  er (næsten) ens. Tilsvarende tænker vi os en værdi  $\tau_2$ , hvor signal 3 separeres fra signal 1 og 2. Ved at definere kvotientmatricen som gennemsnittet af  $\hat{\mathbf{Q}}$  for de to beskrevne ”halvgode” valg kan man opnå

et samlet sæt egenverdier, der lever op til kravet om tilstrækkeligt forskellig egenverdier. Som en generalisering af dette eksempel udvides definitionen af kvotientmatricen i ligning 2.4 ved at indføre summen

$$\mathbf{Q}_{udv} = \sum_{k=1}^q \alpha_k \cdot \mathbf{C}_{\tilde{X}}(\tau_k) \mathbf{C}_{\tilde{X}}(0)^{-1} \quad (2.26)$$

hvor  $\alpha_1, \dots, \alpha_q$ ,  $\sum_{k=1}^q \alpha_k = 1$  er et sæt af vægtningskoefficienter.

For endelige serier medvirker dette at 2.22 må erstattes med

$$\hat{\mathbf{Q}}_{udv} = \frac{1}{2} \mathbf{\Lambda} \left[ \sum_{k=1}^q \alpha_k \left( \mathbf{V}_{\tau_k}^{\top} \mathbf{V} + \mathbf{V}^{\top} \mathbf{V}_{\tau_k} \right) \right] \mathbf{\Lambda}^{-1} \quad (2.27)$$

Den udvidede definition af kvotientmatricen medfører naturligvis tilsvarende ændringer i udtrykkene for f.eks.  $\mathbf{\Sigma}$ , men da disse udtryk ikke anvendes til eksplicite beregninger, medtages de ikke her. I selve analysen bestemmes både blandingsmatrix og kilde serier på samme måde som ved anvendelse af kun en værdi for  $\tau$ .

At anvende flere forskellige værdier af  $\tau$  samtidigt kan dermed i nogle tilfælde forbedre analysens resultater og giver mulighed for at lette den gjorte antagelse om, at kilde serierne skal have forskellige autokorrelationsfunktioner for mindst en værdi af  $\tau$ . man kan nøjes med at antage, at der findes et sæt værdier af  $\tau$  og  $\alpha$ , der gør at  $\mathbf{Q}$  (eller  $\hat{\mathbf{Q}}$ ) som defineret i ligning 2.26 (og 2.27) har tilstrækkeligt forskellige egenverdier.



## Del 3

# Data

Konklusionen på afsnit 2.4 var, at Dynamisk Dekorrelation kan udføres ved at konstruere en håndfuld matricer og diagonalisere disse. At gennemføre Dynamisk Dekorrelation af et datasæt i praksis er dog en noget mere omfattende affære. I dette afsnit vil jeg beskrive hvilke krav der må stilles til data, og hvordan de valgte datasæt klargøres til Dynamisk Dekorrelation, mens selve analysen beskrives i det efterfølgende afsnit.

### 3.1 Valg af datasæt

Datasæt, der skal anvendes til stratigrafisk datering, må have en tilstrækkelig høj opløsning til, at alle årlag, der findes i iskernen, er repræsenteret i datasættet. Det er vanskeligt at opstille et generelt kriterium for om opløsningen er "tilstrækkelig høj", men man bør ved opstilling af et sådan kriterium være opmærksom på følgende :

- Hvis målinger opsamles med konstant frekvens  $\nu$ , kan man kun med sikkerhed identificere svingninger med frekvens mindre end  $\frac{\nu}{2}$ , den såkaldte Nyquist-frekvens, idet der ellers kan opstå *aliasing*<sup>1</sup>. Er man interesseret i det årlige signal, må man derfor som absolut minimum have mere end 2 målinger pr. år.

Risikoen for aliasing er mindre ved af anvendelse data, der som iskernedata indeholder en svingning med varierende frekvens, men mindst 2 målinger pr. gennemsnitlig årlags-tykkelse er stadig et absolut mindstekrav for overhovedet at kunne identificere den årlige svingning.

- I iskernedata forekommer der årlagstykkelser, der afviger markant fra gennemsnittet. Opløsningen af målingerne skal være høj nok til at også et årlag, der f.eks. kun er halvt så tykt som gennemsnittet, bliver repræsenteret i data som et årlag. En opløsning på

---

<sup>1</sup>Der findes mig bekendt ikke noget dansk ord, der har samme betydning som det engelske *aliasing*. Hvis en svingning har frekvensen 1Hz og man foretager målinger med frekvensen 1,5Hz (der altså er større end svingningens frekvens men mindre end Nyquist-frekvensen), vil målingernes værdier svinge med frekvensen  $\frac{1}{1.5\text{Hz}/1\text{Hz}-1} = 2\text{Hz}$ . At svingningen dermed tillægges en forkert frekvens benævnes aliasing.

4 målinger pr. gennemsnitlig årlagstykkelse må derfor være absolut minimum hvis også årlag med halv tykkelse i forhold til gennemsnittet skal repræsenteres i dataserien.

- Det årlige signal for f.eks.  $\text{Cl}^-$  er ikke en harmonisk svingning, men består af 1 eller 2 årlige toppe, der rejser sig over et lavt baggrundsniveau. Det er derfor vigtigt at målingerne foretages med en opløsning, der sikrer at en top med høj koncentration, der f.eks. kun har en varighed på få måneder, også repræsenteres i data som et tydeligt maksimum.

Her er det nødvendigt at skelne mellem analysemetoder, der giver ægte diskrete værdier fordi målingen repræsenterer signalets værdi i et antal målepunkter, og metoder hvor iskernen skæres i små stykker, der smeltes og måles, hvorved hver målt værdi repræsenterer signalets gennemsnitlige værdi i et interval. Herved bliver dataserien naturligvis også diskret, men kun som følge af midling af et kontinuert signal.

I førstnævnte tilfælde løber man en risiko for at en spids top falder mellem to målepunkter og derfor slet ikke registreres, mens toppen i sidstnævnte tilfælde altid er med, men bliver opblandet med resten af prøven og derfor svækkes ved fortynding, evt. endda fordelt i to prøver hvis toppen falder sammen med grænsen mellem to prøver. Den nødvendige opløsning ved måling af ægte diskrete signaler afhænger af varigheden af omtalte årlige top, mens den tilstrækkelige opløsning for middelværdimetoder i højere grad afhænger af den samlede stofmængde, der findes i en top sammenlignet med baggrundsniveauet.

- Har man flere parallelle dataserier fra samme iskerne, kan man tillade sig at arbejde med en anelse dårligere opløsning, da signalernes indbyrdes forskellighed mindsker risikoen for at et årlag overses på grund af målingernes begrænsede opløsning.

Sammenfattende må man sige, at en opløsning på 4 målinger pr. gennemsnitlig årlagstykkelse er den teoretiske undergrænse for at registrere det årlige signal i dataserier fra iskerner, og at 5-6 målinger pr. gennemsnitlig årlagstykkelse er en rimelig undergrænse i praksis, når man arbejder med flere parallelle serier, der er målt med middelværdimetoder.

Styrken ved Dynamisk Dekorrelation er metodens evne til at samle den årlige komponent fra flere forskellige dataserier i en enkelt komponent. Datasættet må derfor indeholde flere serier fra samme dybdeinterval, der indeholder elementer af det årlige signal. Dette krav, sammenholdt med behovet for minimum 5-6 målinger pr. gennemsnitlig årlagstykkelse, medfører, at udvalget af anvendelige dataserier er begrænset.

I de iskerner, der er analyseret på Københavns Universitets geofysiske afdeling, er der som hovedregel målt isotopforhold med en opløsning på 5 cm gennem hele kernens længde på afdelingens massespektrometer, selvom opløsningen til tider er bedre. Indholdet af ioner er typisk målt med ionkromatografi med en opløsning på 5 cm, men langt fra alle kerner har været underkastet denne analyse, der indebærer et meget stort prøverensningsarbejde. Fra NGRIP1-iskernen er både  $\delta^{18}\text{O}$  og ionkoncentrationer tilgængelige i 5 cm opløsning i de øverste 350

meter, men med en gennemsnitlig akkumulation på 19,5 cm is år<sup>-1</sup> [Dahl-Jensen *et al.*, 2001], er datasættets opløsning kun tilstrækkelig høj i den øverste halvdel af firnlaget, hvor årlagstykkelsen er væsentlig højere da sneen kun er delvis komprimeret på dette stadium. Til gengæld omfatter ionkromatografi-datasættet koncentrationen af ikke mindre end 12 ioner (i  $\mu\text{ækv./kg}$ ), hvilket er et usædvanligt højt antal parametre. Efter råd af H. B. Clausen har jeg valgt en sekvens på ca. 7 meter fra omkring 30 meters dybde, hvor opløsningen er lige over 5,5 prøver pr. år i gennemsnit, og således akkurat god nok. H. B. Clausen har dateret toppen af NGRIP1-iskernen ud fra identifikation af årlag i  $\delta^{18}\text{O}$  og ionkromatografi-data og har forsynet mig med en tidsskala, der dækker den valgte sekvens. Denne tidsskala vil blive brugt til vurdering af resultaterne fra Dynamisk Dekorrelation og er meget sikker, da endepunkterne af serien er 2 veldaterede referencehorisonter, nemlig et klart identificerbart smeltelag i 1889 og udbruddet af Katmai i 1912.

Støvindhold og især ECM-data findes i tilstrækkelig opløsning for flere iskerner, men da der sjældent findes andre parallelle serier fra samme iskerner, er antallet af parallelle serier i underkanten til Dynamisk Dekorrelation. Ved forundersøgelserne til GRIP-boringen på Midtgrønland blev der boret en række shallow-kerner nær Crête, der kendes som Site A, B, C, D, E, F og G, eller *alfabetserien*. Fra disse kerner findes  $\delta^{18}\text{O}$  og ECM-data, men J. P. Steffensen har desuden målt koncentrationen af 3 udvalgte ioner samt støvindhold (i masseenheder) for perioden 1889 - 1912 i Site A og D kernerne [Steffensen, 1988]. De 6 parallelle serier udgør et velegnet datasæt til Dynamisk Dekorrelation på grund af den gode opløsning, der skyldes relativ høj akkumulation samt at prøverne er taget fra firn-laget. Det skal tilføjes, at det er et tilfælde at Site A/D og NGRIP dataserierne netop dækker den samme tidsperiode, idet jeg begyndte analysen af NGRIP datasættet, før jeg kendte til eksistensen af Site A/D datasættene. Datasættene fra Site A/D minder meget om hinanden, og jeg har analyseret de to datasæt helt parallelt. Resultaterne er næsten identiske, hvorfor jeg har valgt kun at vise resultater fra Site D. En oversigt over hvilke data, der er tilgængelige for Site D findes i tabel 3.1.

Dataserie	Dybdeinterval (m)	Opløsning (cm)
$\delta^{18}\text{O}$	1,16 - 100,06	3 - 8
ECM	1,21 - 100,05	1
$[\text{Cl}^-]$ , $[\text{SO}_4^-]$ , $[\text{NO}_3^-]$ , Støv	42,40 - 51,15	5

Tabel 3.1: Tilgængelige dataserier fra Site D.

Det segment, hvor der findes 5 cm ion-koncentrationsdata og støvdata, dækker perioden 1889 - 1910 [Steffensen, 1988]. Denne datering baserer sig på manuel optælling af årlag i  $\delta^{18}\text{O}$ -profilen, og er efterfølgende valideret ved hjælp af smeltelaget fra 1889 og Katmai-udbruddet i 1912. Tidsskalaen må derfor klassificeres som velunderbygget, og jeg har derfor brugt denne tidsskala som reference ved vurdering af kvaliteten af resultaterne fra Dynamisk Dekorrelation.

Efter boring af kernen blev der skåret prøver til  $\delta^{18}\text{O}$ -måling med en prøvestørrelse svarende til 8 prøver pr. år. Størrelsen af prøverne blev valgt ud fra en forudgående model-datering, og prøvestørrelsen varierer derfor fra 8 cm i toppen til 3-4 cm i bunden af kernen. Da Dynamisk Dekorrelation kræver at dataserierne er målt med samme dybde-inddeling har jeg foretaget en *resampling* af  $\delta^{18}\text{O}$ -dataserier til 5 cm inddeling ved hjælp af 3.ordens *spline*-interpolation. ECM-værdierne er omregnet til 5 cm inddeling ved midling af 1 cm dataserien.

Mens der som nævnt blev målt ionkoncentrationer ved ionkromatografi i NGRIP1-iskernen, har man anvendt Continuous Flow Analyses (CFA) til måling af ion-koncentrationer i NGRIP2-iskernen. Denne målemetode tillader en opløsning på ned til under 1 cm, men efterbehandlingen af resultaterne herfra er ikke tilendebragt, og det har derfor ikke været muligt at anvende CFA data fra NGRIP i denne opgave. Istedet har jeg anvendt data fra en CFA-profil fra en amerikansk iskerne fra Summit, Grønland, der strækker sig fra overfladen til 136,61 meters dybde. Datasættet, der endnu ikke er publiceret i sin helhed, er venligst stillet til rådighed af Manuel Hutterli, University of Arizona, Tucson, USA, og har en opløsning på 2 cm. Der mangler en del værdier i datasættet, og for samtidig at undgå at skulle korrigere for udtynding af årlag på grund af isens flydning (jævnfør. side 2.1), har jeg udvalgt korte 4 sekvenser fra datasættet. De 4 datasæt er valgt så der mangler maksimalt 2 datapunkter i træk for hver komponent, og stammer fra dybderne 47 m, 78 m, 115 m og 131 m og dækker således kernens længde nogenlunde repræsentativt. Datasættet består af 6 parallelle serier :  $[\text{NH}_4^+]$ ,  $[\text{Ca}^{++}]$ ,  $[\text{NO}_3^-]$ ,  $[\text{H}_2\text{O}_2]$ ,  $[\text{HCHO}]$  i ppbw samt smeltevandets elektriske ledningsevne i  $\mu\text{S}$ .

De vigtigste informationer om de anvendte data er samlet i tabel 3.2. Det bør nævnes at både  $\delta^{18}\text{O}$  og ECM-data er tilgængelige fra NGRIP, men jeg har valgt ikke at anvende dem for at gøre Site D og NGRIP datasættene så forskellige som muligt, og derved teste Dynamisk Dekorrelation mere alsidigt. De ubehandlede datasæt er ikke afbildet her, men NGRIP datasættet kan findes i forbindelse med beskrivelsen af regularisering som figur 3.1, Site D datasættet kan ses i resultatafsnittet i figur 5.1, og CFA-datasættet fra Summit er illustreret på figur 5.12.

I det følgende vil dataserierne konsekvent blive præsenteret på figurer uden enheder, da de før videre analyse vil blive normaliseret for at indgå med lige vægt, og absolutte værdier derfor ikke spiller nogen rolle.

Det skal afslutningsvis nævnes, at det kan virke underligt at vælge Dynamisk Dekorrelation til iskerneanalyse når så få dataserier har tilstrækkelig høj opløsning til at kunne honorere metodens krav. Hertil kan siges, at det oprindeligt var denne opgaves mål at undersøge om Dynamisk Dekorrelation kunne bruges til datering af NGRIP iskerne ved hjælp af CFA-data, men at den omtalte efterbehandling af måleresultaterne som nævnt har forsinket datasættets tilgængelighed. Det kommende datasæt fra CFA-målingerne på NGRIP-iskernen vil være ideelt til formålet, idet det årlige signal kan identificeres 35-50.000 år tilbage i tiden og består af ca. 10 serier, hvoraf de fleste indeholder elementer af det årlige signal.



Borested	Dybdeinterval (m)	Serier	Omtrentligt antal år	Prøver pr. år (gennemsnit)
Site D	42,40 - 51,15	$\delta^{18}\text{O}$ , $[\text{Cl}^-]$ , $[\text{SO}_4^{--}]$ , $[\text{NO}_3^-]$ , Støv, ECM	20	8,8
NGRIP	27,55 - 34,85	$[\text{Li}^+]$ , $[\text{Na}^+]$ , $[\text{NH}_4^+]$ , $[\text{Mg}^{++}]$ , $[\text{Ca}^{++}]$ , $[\text{F}^-]$ , [MSA], $[\text{Cl}^-]$ , $[\text{HCO}_3^-]$ , $[\text{NO}_3^-]$ , $[\text{SO}_4^{--}]$ , $[\text{K}^+]$	23	5,5
Summit serie 1	46,01 - 48,85	$[\text{NH}_4^+]$ , $[\text{Ca}^{++}]$ , $[\text{NO}_3^-]$ ,	9	15
serie 2	77,45 - 80,13	$[\text{H}_2\text{O}_2]$ , [HCHO]	11	12,5
serie 3	113,93 - 118,05		17	12
serie 4	130,13 - 132,77		11	12

Tabel 3.2: Karakteristika for de anvendte dataserier

### 3.2 Manglende datapunkter

De datasekvenser, jeg har valgt, er blandt andet valgt ud fra et ønske om kun at have få og isolerede manglende datapunkter. Ved Dynamisk Dekorrelation kan manglende værdier i dataserierne være en alvorlig hindring, da beregningen af auto- og krydskorrelationsfunktionernes værdier giver misvisende resultater. Ved beregning af korrelationsfunktionernes værdier ”parres” værdierne fra to serier, som det ses på side 32, og den værdi, der parres med en manglende dataværdi vil ikke indgå i resultatet. Da det er tilfældigt, om det er store eller små værdier, der udgår, medfører manglende datapunkter derfor en fejl i beregningerne, hvis betydning ikke kan forudsiges. Jeg har derfor valgt at indsætte anslåede dataværdier de steder, hvor der har manglet værdier.

Jeg har valgt at anvende lineær interpolation til at erstatte de manglende værdier. For Site A og D samt NGRIP har der kun været isolerede manglende værdier, hvorved middelværdien af de tilstødende værdier er anvendt som anslået værdi, mens der i CFA datasættet flere steder har været 2 på hinanden følgende manglende værdier. Jeg har ikke fundet det nødvendigt at anvende mere sofistikerede metoder til at estimere de manglende værdier, primært fordi der ikke er tale om særlig mange værdier, og fordi en mere avanceret metode også vil indføre flere parametervalg, som jeg ikke finder, at jeg har baggrund for at foretage.

Det skal nævnes, at udvikling af en mere generel metode til erstatning af manglende værdier er nødvendig, før Dynamisk Dekorrelation kan anvendes på længere sekvenser af CFA-data. CFA-dataserier mangler typisk mange værdier, ikke mindst fordi små isstykker, der stammer fra de brud af iskernen, der kommer, når iskernen knækkes under boringen, ikke kan analyseres. Også endefladerne af de enkelte isprøver fjernes inden måling, så intervaller på op til 5 - 10 cm med manglende data for alle komponenter forekommer jævnlige.

Under et seminar på IMM på DTU stiftede jeg bekendtskab med *Gaussian Mixtures*, som jeg tror vil være i stand til at estimere de manglende værdier med langt større succes end

almindelige interpolationsmetoder. En fordel ved denne metode er desuden, at man ikke kun vil kunne erstatte de manglende værdier med estimater, men at man tillige vil kunne vurdere usikkerheden heraf.

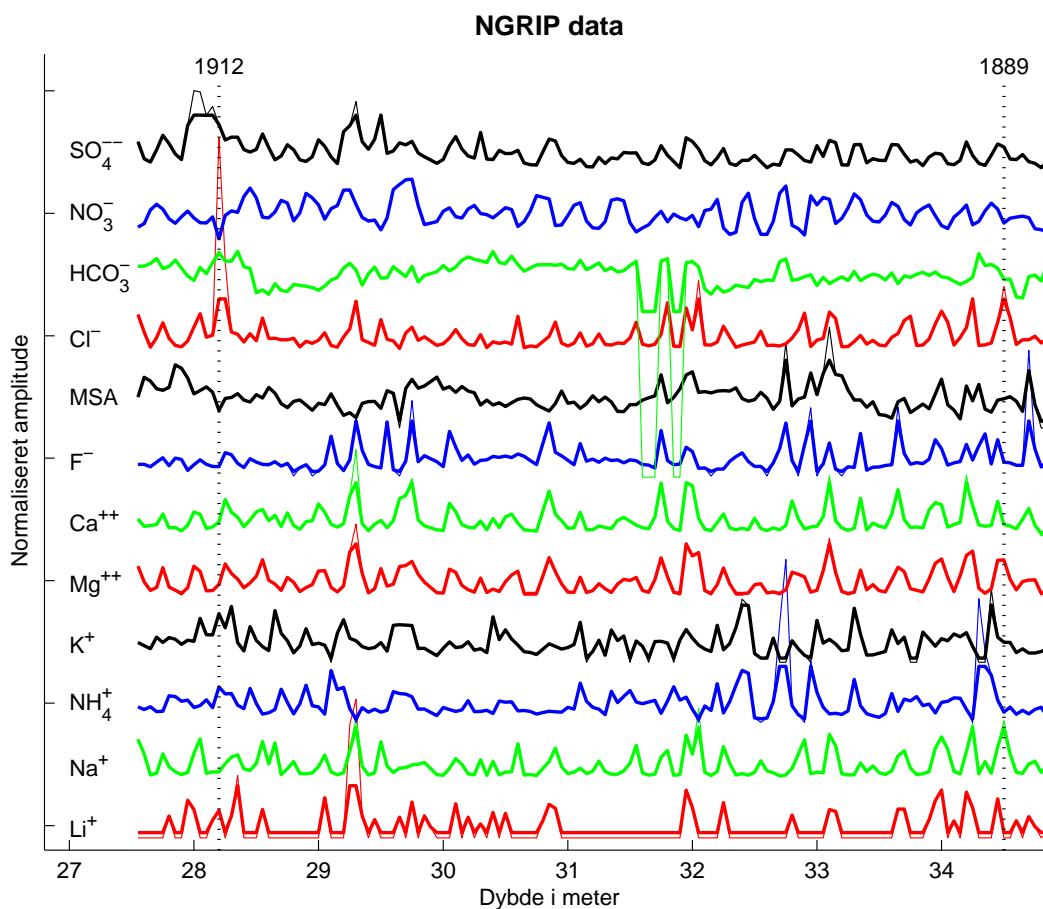
### 3.3 Regularisering af data

Et typisk datasæt indeholder som beskrevet på side 2 både periodiske og sporadiske bidrag samt støj. Man kunne håbe at Dynamisk Dekorrelation ville være i stand til at isolere bidragene fra et sporadisk signal i en enkelt kildeSerie, så den sporadiske komponent ikke ville påvirke de periodiske kildeSerier, men i praksis er dette ikke sandsynligt. I Site D datasættet findes f.eks. en klorid-top i 46,35 meters dybde, der er omtrent 70 gange højere end et typisk årligt maksimum, mens toppen stort set ikke ses i de andre dataserier. En sådan enlig top vil ved konstruktion af kildeSerierne ved Dynamisk Dekorrelation give anledning til flere problemer :

- Ved beregning af auto- og krydskorrelationsfunktionernes værdier vil der komme et meget væsentligt bidrag fra denne høje top, hvilket vil give et skævvredet billede af dataseriernes korrelationsfunktioner.
- KildeSerierne dannes rent praktisk jævnfør ligning 2.25 ved linearkombinationer af dataserierne, og toppen vil derfor dukke op i alle kildeSerier, der indeholder et bidrag fra klorid-serien. Ved konstruktion af kildeSerien der repræsenterer den årlige svingning, vil man derfor enten overhovedet ikke få bidrag (heller ikke fra den årlige komponent) fra kloridserien eller få et bidrag fra den høje top sammen med den årlige komponent.
- Ved normering af dataserierne til  $\mu = 0$  og  $\sigma = 1$  vil den høje top betyde, at resten af signalet vil blive skaleret ned, og derfor vil blive tillagt mindre vægt i analysen.

Det skal understreges, at der ikke teoretisk er noget i vejen for, at det sporadiske signal opsamles alene i en kildeSerie, men at det blot ofte ikke vil være tilfældet i praksis. Jeg har derfor valgt at fjerne de sporadiske bidrag på forhånd. Jeg har desuden valgt at reducere højden af unormalt høje årlige toppe, da de, på samme måde som beskrevet ovenfor, kan tænkes at have en negativ indflydelse på resultatets anvendelighed til datering. Denne procedure vil jeg omtale som regularisering af data.

[Meeker *et al.*, 1995] foreslår at identificere den sporadiske komponent i iskerne-data som de værdier, der adskiller sig mere end  $4\sigma$  fra et udglattet signal, der repræsenterer den langsomme trend i data. Den langsomme trend foreslås bestemt ved 3.grads *spline*-interpolation eller ved hjælp af *wavelet*-analyse. Jeg har valgt et skrapere kriterium, idet jeg har nedjusteret værdierne for alle toppe, der afviger mere end  $3\sigma$  fra middelværdien. Både spredning og middelværdi beregnes af alle data, inklusive de data, der nedjusteres. Ved nedjusteringen ændres middelværdien, og spredningen falder lidt, hvorefter nedjusteringen gentages med de nye værdier af  $\mu$  og  $\sigma$ . Resultatet er en iterativ proces, som jeg fortsætter indtil den største



Figur 3.1: Det udvalgte NGRIP datasæt. De tykke linier angiver serierne efter regularisering, hvor værdier, der afviger mere end  $3\sigma$  fra middelværdien, er nedjusteret til  $3\sigma$ -niveauet. De to årstal angiver hvornår årlagene starter (det vil sige henholdsvis vinter 1889/90 og 1912/13) i følge H. B. Clausens tidsskala, der er anvendt som reference-tidsskala.

nedjusteringen er mindre end  $10^{-4}\sigma$ . Mit valg af  $3\sigma$  som grænse kan begrundes med, at det som nævnt er vigtigt, at enkeltbegivenheder i data ikke får for stor indflydelse på det samlede resultat. For normalfordelte data (se i øvrigt afsnit 3.4) vil kun omtrent 3% af datapunkterne i gennemsnit falde udenfor  $3\sigma$ -grænsen. På figur 3.1 ses NGRIP datasættet før og efter den beskrevne regularisering er foretaget. Det ses, at det kun er relativt få justeringer, der foretages ved den beskrevne procedure, og at tydeligheden af det årlige signal på ingen måde svækkes ved nedjusteringen af de høje toppes værdier.

Da jeg her kun anvender korte dataserier, hvor den langsomme trend er ubetydelig, har jeg fundet det tilstrækkeligt at bruge datapunkternes afvigelse fra middelværdien som kriterium for nedjustering, men det vil være naturligt at bruge en udglattet kurve som udgangspunkt ved analyse af data, der udviser en tydelig langsom trend. Formentlig vil det ligeledes være formålstjenligt at fjerne denne trend inden der foretages Dynamisk Dekorrelation, selvom det meget vel kan tænkes, at en langsom trend, der er fælles for flere serier, vil blive samlet i en af de producerede kildeserier.

### 3.4 Fordelingsfunktioner for data

Nogle målte størrelser kan i sagens natur ikke antage negative værdier, mens der ikke på samme måde er en øvre grænse for størrelsen. Dette er f.eks. tilfældet for ionkoncentrationer i isprøver, der naturligvis ikke kan være negative. Omend der i teorien er en maksimal koncentration af disse ioner i is, spiller denne overgrænse i praksis ingen rolle. Man kan derfor ikke forvente at en svingning i en sådan dataserie vil udvise symmetriske udsving fra en middelværdi, mens logaritmen til koncentrationerne (i en passende enhed) må forventes at udvise en højere grad af symmetri.

Motivationen for at undersøge fordelingen af logaritmen af værdierne i nogle af dataserierne har været, at det årlige signal i de tilgængelige dataserier har væsentligt forskellige karakteristika. Nogle serier, f.eks.  $\delta^{18}\text{O}$ -signalet og til en hvis grad koncentrationerne af  $\text{NO}_3^-$  og  $\text{SO}_4^{--}$ , udviser blød årlig variation, mens andre serier, f.eks. koncentrationerne af  $\text{Li}^+$  og  $\text{Cl}^-$ , har spidse årlige toppe, der rejser sig stejlt fra et baggrundsniveau. Målet er som bekendt at samle de årlige komponenter i en enkelt kilde serie ved hjælp af Dynamisk Dekorrelation, og det var min formodning at dette bedst vil lykkes, hvis den årlige komponent i de forskellige serier ligner hinanden så meget som muligt, hvilket kan opnås ved at anvende logaritmen til dataværdierne.

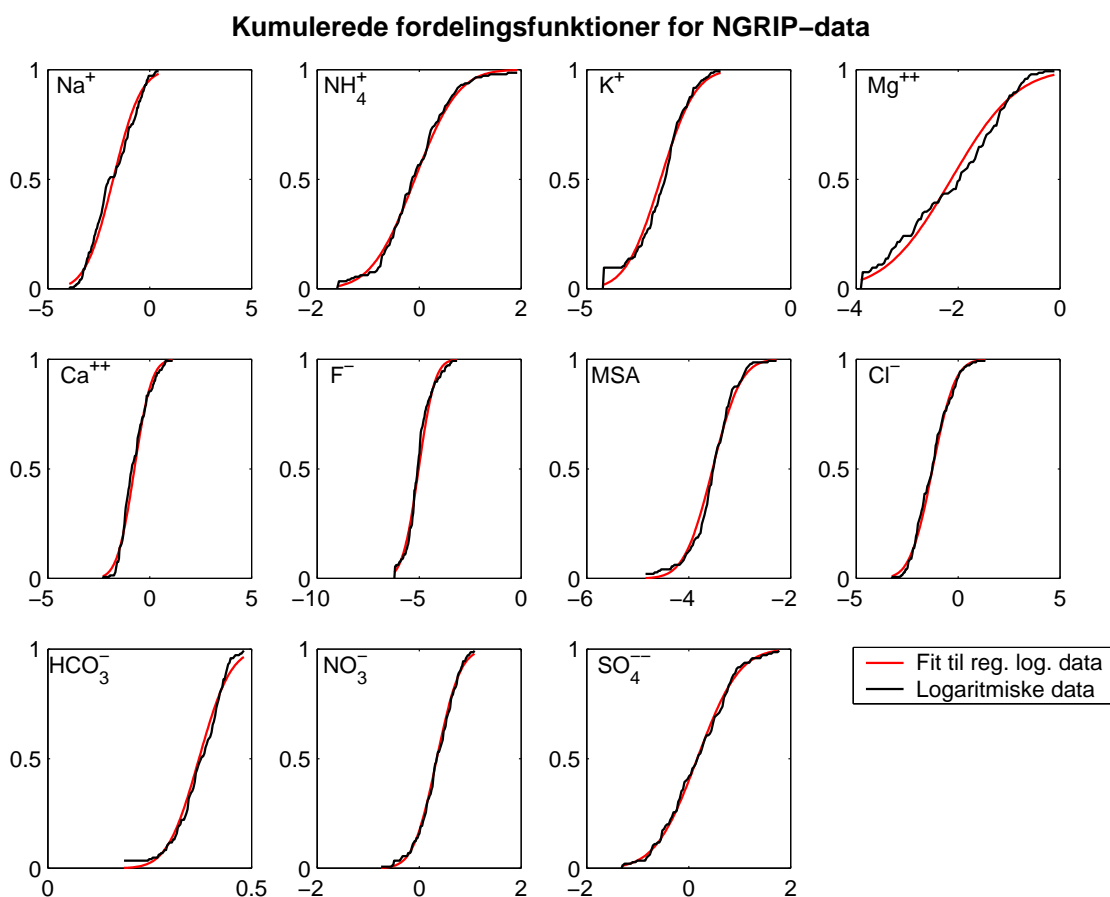
For at afgøre om logaritmen til dataværdierne svinger symmetrisk om middelværdien har jeg undersøgt fordelingen af logaritmen til dataværdierne i alle dataserier undtagen  $\delta^{18}\text{O}$ . I denne forbindelse er der to tekniske detaljer, der skal overvejes.

- I nogle serier optræder værdien ”målt nul”, hvor den målte værdi har været lavere end det anvendte apparaturs detektionsgrænse. En ”målt nul” værdi kan således repræsentere en vilkårlig værdi mellem 0 og detektionsgrænsen, men for at undgå problemer med logaritmen til disse værdier har jeg erstattet dem med en anslået detektionsgrænse for apparaturet. Denne værdi har jeg valgt ved at se på de mindste målinger, der ikke er nul, for hver serie. I tabel 3.3 ses de 3 mindste målinger for hver serie i NGRIP datasættet og den anvendte anslåede detektionsgrænse. En tilsvarende procedure er anvendt for de øvrige datasæt.

Serie	$\text{NH}_4^+$	$\text{K}^+$	$\text{Mg}^{++}$	$\text{Ca}^{++}$	$\text{F}^-$	MSA	$\text{HCO}_3^-$
Anslået detektionsgrænse	0,20	0,010	0,020	0,10	0,0020	0,007	1,00
De 3 laveste værdier i datasættet	0,24 0,26 0,28	0,010 0,010 0,016	0,024 0,025 0,025	0,14 0,18 0,18	0,0022 0,0024 0,0026	0,009 0,010 0,011	1,28 1,29 1,31

Tabel 3.3: De anslåede detektionsgrænser som erstatter målte nul-værdier. I de serier, der ikke er nævnt, var der ingen ”målt nul” værdier.

- Når man vil undersøge fordelingen af logaritmen til dataværdierne, kan man enten først tage logaritmen til værdierne og beregne den kumulerede fordelingsfunktion (*cumulated density function* eller blot *cdf*), hvorefter man tilpasser middelværdi og spredning i en normalfordeling for at opnå bedste mulig overensstemmelse mellem fordelingen af data og normalfordelingen. Denne tilpasning vil typisk foregå med mindste kvadraters metode. Alternativt kan man tilpasse parametrene i en log-normalfordeling direkte til dataværdierne, hvilket igen typisk vil ske med mindste kvadraters metode. Forskellen ligger i vægtningen af residuerne, der foretages henholdsvis før og efter logaritme-transformationen. Jeg har valgt førstnævnte metode af to årsager. For det første er normalfordelingens fordelingsfunktion i modsætning til log-normalfordelingens fordelingsfunktion let genkendelig, og for det andet giver denne metode efter min mening den mest ligelige vægtning af afvigelser for henholdsvis store og små værdier.



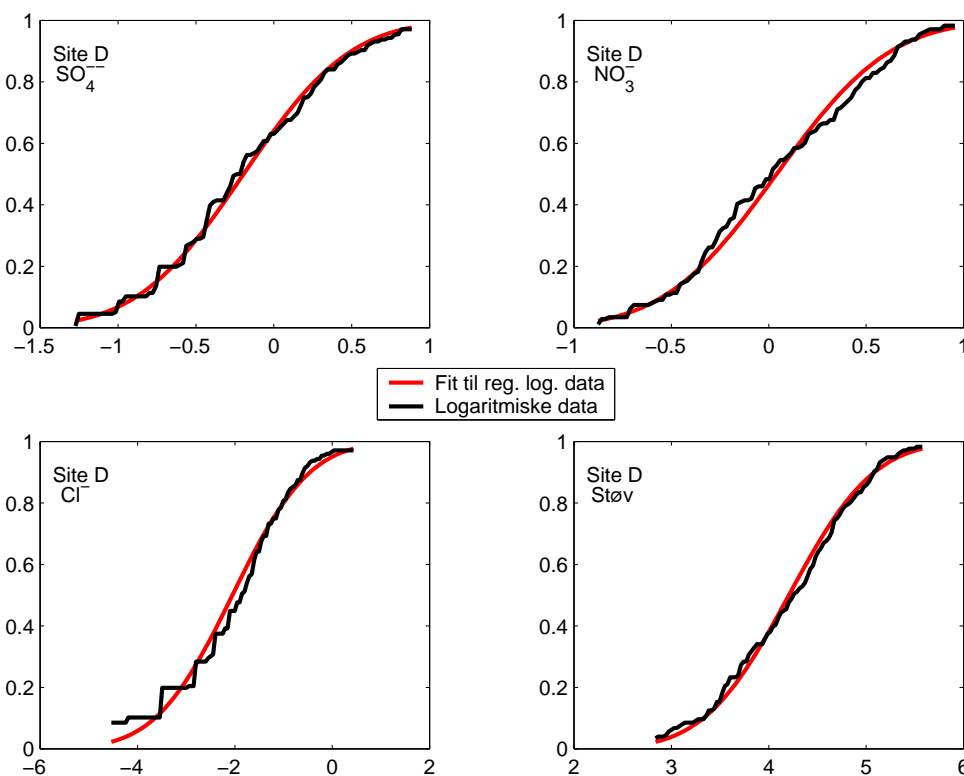
Figur 3.2: Kumuleret fordelingsfunktion (ordinat) for logaritmen til koncentrationerne (abscisse) af 11 af de 12 målte ioner i NGRIP datasættet. Til sammenligning det bedste fit af en normalfordeling til de regulariserede data. Det trin, der tydeligst observeres på grafen for K<sup>+</sup>, skyldes den omtalte procedure hvor "målt nul" er erstattet af værdien af den anslåede detektionsgrænse. Li<sup>+</sup> ikke er medtaget i analysen, da koncentrationen oftest er under detektionsgrænsen, og datasættet derfor ikke er velegnet til formålet.

For de således let modificerede datasæt har jeg bestemt de kumulerede fordelingsfunktioner. Resultatet for NGRIP data er vist på figur 3.2, sammen med det bedste fit af en normalfordeling til de regulariserede data. Bemærk, at data *ikke* er regulariseret før beregning af fordelingsfunktionerne, men at det kun er den fittede normalfordelings fordelingsfunktion, der er beregnet ud fra regulariserede data. Jeg har valgt denne procedure, da MatLabs `normfit`-rutine er relativ følsom over for enkelte datapunkter, der ligger fjernt fra gennemsnittet. Det ses, at ion-koncentrationernes logaritmisk transformerede med god tilnærmelse er normalfordelt. De afvigelser fra de fittede kurver, der ses på figuren, kan henføres til at der kun indgår 147 punkter i analysen. De tilsvarende resultater fra Site D ses på figur 3.3 og viser tilsvarende at ion-koncentrationerne og støvindholdet med god tilnærmelse kan opfattes som værende log-normalfordelte. Dette faktum er i sig selv et resultat, og peger på at det vil være rimeligt at anvende logaritmisk transformerede dataserier i de tilfælde, hvor serierne repræsenterer størrelser, der er naturligt positive. Herved opnås en mere jævn svingning om middelværdien, i hver enkelt dataserie.

For at undersøge om anvendelse af logaritmisk transformerede data har indflydelse på resultaterne af Dynamisk Dekorrelation, har jeg gennemført analysen for NGRIP-data både med almindelige og logaritmisk transformerede data. For Site D og CFA-datasættet fra Summit anvendes kun almindelige data.

### 3.4.1 Fordelingen af ECM-data fra Site A og D

ECM-serien fra Site D dækker hele kernens længde, det vil sige dybdeintervallet 1,16 - 100,06 meter. Ved beregning af de kumulerede fordelingsfunktioner for den del af Site D iskernen, hvorfra der også findes ion-koncentrationsdata, viste det sig, at ECM-data ikke kunne opfattes som log-normalfordelte, mens en sekvens fra toppen af kernen viste det modsatte. Jeg har derfor bestemt den kumulerede fordelingsfunktion individuelt for 3 stykker af serien, nemlig henholdsvis over, i og under det stykke, hvor der er udtaget prøver til ionkromatografi. Resultaterne ses på figur 3.4. De røde kurver viser normalfordelingens fordelingsfunktion, fittet til henholdsvis de rå data (stiplet) og til de dataværdier, der har log-værdi større end -1 (optrukket). Regularisering af data ved nedjustering af ekstremværdier til  $3\sigma$ -niveauet, som ellers er standard i databehandlingen, var ikke tilstrækkeligt til at give et godt fit til data-kurven, og nævnte procedure var den, der efter talrige forsøg gav den bedste overensstemmelse mellem data og den fittede kurve. Det ses, at der er en betydelig forskel på fordelingen af data i de tre intervaller. De værdier, der stammer fra dybder mindre end 42,39 meter, det vil sige fra ovenover den del af kernen, hvorfra der blev udtaget prøver til ionkromatografi, er log-normalfordelte med rimelig tilnærmelse. I intervallet 42,40 - 51,15 meter og i den nedre del af serien er værdierne log-normalfordelte, men kun hvis man ser bort fra fordelingen af de små værdier. Lad os betragte fordelingsfunktionen for intervallet 42,40 - 51,15 meter. For log-værdier ned til -0,60 (svarende til ECM-værdier på  $0,55 \mu\text{ækv H}^+$ ) passer fordelingsfunktionen



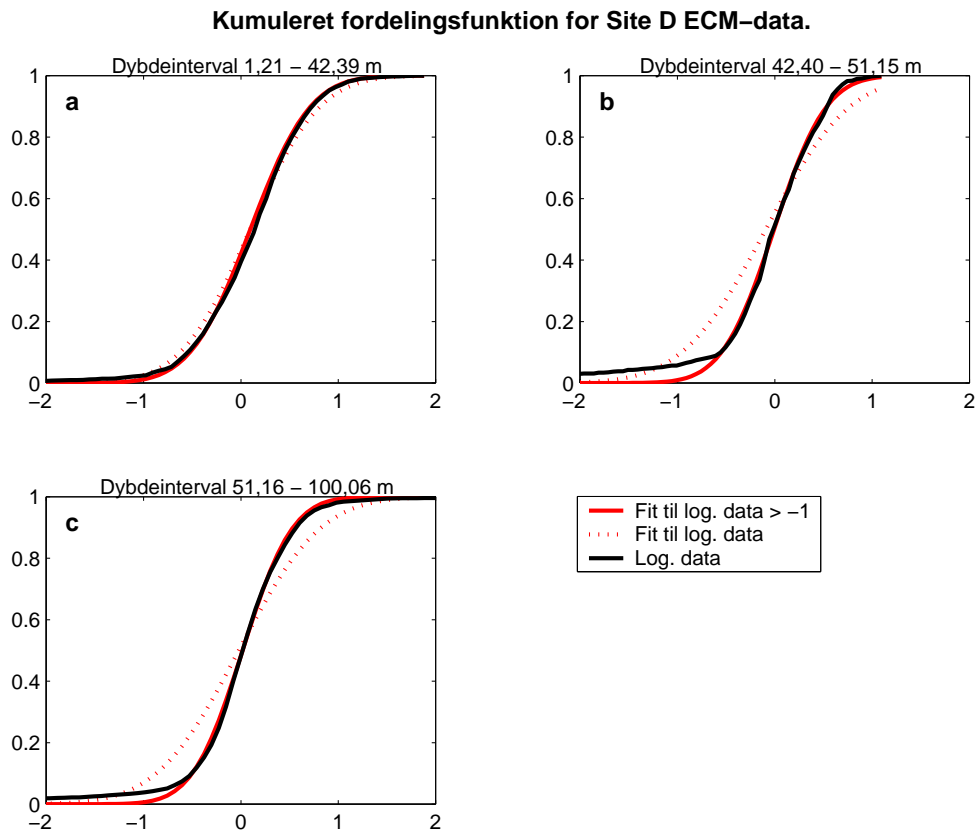
Figur 3.3: Kumuleret fordelingsfunktion (ordinat) for logaritmen til koncentrationerne (abszisse) af  $\text{SO}_4^{2-}$ ,  $\text{NO}_3^-$  og  $\text{Cl}^-$  samt støvindhold i Site D datasættet. Til sammenligning er vist normalfordelingens fordelingsfunktion, hvor  $\mu$  og  $\sigma$  er tilpasset for bedst muligt fit til de regulariserede data. Klorid-koncentrationer under detektionsgrænsen er erstattet af den anslåede detektionsgrænse for det anvendte apparatur. Trinnet til venstre på grafen for kloridkoncentrationens fordelingsfunktion skyldes denne procedure.

pænt med en normalfordeling med middelværdi  $\mu = -0,01$  og spredning  $\sigma = 0,42$ . Kommer man derimod under denne værdi, afviger værdierne's fordeling markant fra log-normalfordelingen. De 8% af værdierne, der har log-værdi under  $-0,60$  er fordelt langt mere jævnt i intervallet  $]0; 0,55]$ , end hvad der ville have været tilfældet for log-normalfordelte data. Dette ses særlig klart på det udsnit af figur 3.4 b), der er vist i figur 3.5.

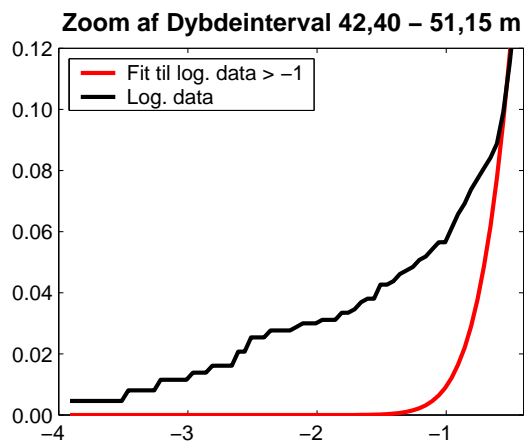
En nærmere undersøgelse af fordelingen af ECM-værdierne i de 3 dybdeintervaller i såvel Site A som Site D datasættene viser, at den omtalte skævhed i forhold til en log-normalfordeling

- er langt kraftigst for de stykker af både Site A og D kernerne, hvor der er udtaget prøver til ionkromatografi, jævnfør figur 3.5,
- er noget mindre udtalt, men stadig markant, for det dybe (51,15 - 100,06 meter) interval i Site D-serien og
- er langt svagere i toppen af de to kerner samt i bunden af Site A iskernen.

Alle ECM-data fra både Site A og D datasættene udviser altså en hvis grad af afvigelse fra log-normalfordelingen. Da de største 92% af værdierne med meget fin tilnærmelse er log-



Figur 3.4: Kumuleret fordelingsfunktion (ordinat) for logaritmen til ECM-værdien (abszisse) for Site D datasættet. En tilpasset normalfordelings fordelingsfunktion til sammenligning. Det ses at normalfordelingskurven der er fittet til de logaritmsk transformerede dataværdier (rød stiplede linie) passer meget dårligt med data, mens den kurve, der er fittet til alle data med log-værdi over -1 (rød optrukket linie), passer meget fint undtagen for små værdier.



Figur 3.5: Udsnit af figur 3.4 b), hvortil henvises for forklaring.



normalfordelt som det ses på figur 3.4, mener jeg, at der er god grund til at antage, at indholdet af  $[H^+]$  er log-normalfordelt og søge en forklaring på den observerede afvigelse for de lave værdier. En oplagt forklaring kunne være at afvigelsen opstår på grund af en ikke-lineær sammenhæng mellem  $[H^+]$  og ECM-værdi for lave værdier. [Moore *et al.*, 1994] omtaler netop problemer med kalibreringen af ECM/ $[H^+]$ -forholdet for små værdier.

Jeg har derimod svært ved at tro at den markante forskel på fordelingen af de lave værdier i de enkelte dybdeintervaller kan skyldes en reel fysisk ændring af signalet. Sammenfaldet mellem udtagning af ionkromatografi-prøver og den markante ændring i fordelingsfunktionerne for to forskellige iskerner er påfaldende, men kan naturligvis være et tilfælde. Det er dog svært at vide, hvad der rent fysisk skulle medføre at ECM-signalet i et tilfældigt på forhånd valgt 20-års interval ændres, selvom antropogene udledninger til atmosfæren begynder at få betydning netop i perioden 1889 - 1912. Årsagen kan også tænkes at være forskelle i prøvetagningsproceduren, det anvendte apparatur eller i betingelserne under prøvetagningen. Det skal nævnes her at ECM-målingerne er følsomme overfor ændringer i temperatur under prøvetagningen, ligesom kernens håndtering kan have været af betydning. Umiddelbart er det dog svært at se hvordan udtagningen af ionkromatografi-prøver skulle påvirke ECM-signalet, da ECM-målingen er foretaget før ionkromatografi-prøverne blev skåret. En afklaring af dette spørgsmål vil formentlig kræve en grundigere analyse af ECM-signalerne fra flere af kernerne fra alfabet-serien, og falder uden for dette arbejdes hovedformål.



## Del 4

# Dynamisk Dekorrelation i praksis

I dette afsnit vil jeg beskrive hvordan Dynamisk Dekorrelation foretages i praksis med vægt på hvordan  $\tau$  bestemmes. Først vil jeg dog beskrive en metode til synkronisering af det årlige signal i de enkelte serier, hvilket har vist sig at være nødvendigt. Afsnittet sluttet af med en oversigt over analysens trin fra rå data til endelige resultater.

### 4.1 Synkronisering af årlige toppe

Selvom såvel  $\delta^{18}\text{O}$  som andre iskernesignaler udviser en klar årlig svingning, er disse signaler sjældent i fase.  $\delta^{18}\text{O}$ -signalet, der som nævnt er en temperaturproxy, har maksimum sidst på sommeren samtidig med f.eks. nitratserien, mens kloridseriens årlige top oftest er placeret tidligt på foråret. Man kan derfor ikke forvente, at Dynamisk Dekorrelation vil isolere det årlige signal i en kilde serie, og man risikerer derved at flere af de producerede kilde serier vil indeholde betydelige elementer af det årlige signal, blot forskudt indbyrdes. Et yderligere problem opstår fordi disse kilde seriers autokorrelationsfunktioner vil have værdier, der ikke er klart forskellige, og Dynamisk Dekorrelation vil derfor jævnfør diskussionen i afsnit 2.4.3 ikke klart kunne separere signalerne.

For at imødegå dette problem har jeg søgt at synkronisere den årlige svingning i de enkelte serier. Signaler, hvis årlige svingning er i fase, vil have højere krydskorrelationskoefficienter, så jeg har foretaget synkroniseringen ved at forskyde serierne et antal målepunkter i forhold til hinanden og vælge den forskydning, der giver den største værdi af krydskorrelationskoefficienten. Da nogle signaler i højere grad har et årligt ”dyk” end en årlig top, har jeg tillige tilladt en spejling af serierne om nul. For at undgå at serier rykkes et eller flere hele år i forhold til hinanden, har jeg kun tilladt forskydninger på op til omkring  $3/4$  af det gennemsnitlige antal prøver på et år. Den maksimale forskydning ses i tabel 4.1. Efter at have arbejdet med metoden gennem længere tid kan jeg konstatere, at den udførte synkronisering i langt de fleste tilfælde ifølge mit skøn ser ud til at være optimal. Synkroniseringen kan naturligvis udføres på andre måder, men jeg har fundet det vigtigt at bruge en metode der er baseret på et simpelt

Borested	Prøver pr. år (gennemsnit)	Maksimal tilladt forskydning
Site D	8,8	6
NGRIP	5,5	4
Summit serie 1	15	12
Summit serie 2-4	12-12,5	9

Tabel 4.1: Den maksimale tilladte forskydning serierne imellem ved synkronisering af det årlige signal.

objektivt kriterium, og det er ikke lykkedes mig at finde et bedre kriterium.

Synkroniseringen af det årlige signal er udført i praksis ved at der udvælges en dataserie, i forhold til hvilken de andre serier enkeltvis forskydes ved maksimering af krydskorrelationskoefficienterne. For at kunne vurdere betydningen af valget af fastholdt serie har jeg for hvert datasæt foretaget Dynamisk Dekorrelation af de synkroniserede dataserier med forskellige valg af fastholdt serie. Jeg har sammenlignet de producerede serier for at finde det valg af fastholdt serie, der giver kilde signaler med klareste årlig variation. Det har vist sig at der ikke findes ét bestemt valg af fastholdt serie, der er markant bedre end de andre for NGRIP datasættet, men at synkronisering efter  $\text{NH}_4^+$ ,  $\text{K}^+$  og især  $\text{HCO}_3^-$  virkede godt. I afsnit 5.3 findes en oversigt over hvilke valg af fastholdt serie, der gav gode resultater for NGRIP datasættet. For Site D stod det derimod klart, at synkronisering efter  $\delta^{18}\text{O}$  gav klart de bedste resultater, mens  $\text{NO}_3^-$  gav de bedste resultater for CFA-datasættet fra Summit.

## 4.2 Valg af $\tau$ -værdi

Som beskrevet i afsnit 2.4.4 kan Dynamisk Dekorrelation kun adskille kildeserierne, hvis man kan vælge en værdi af  $\tau$ , for hvilken de normerede kilde signalers autokorrelationsfunktioner  $c_S(\tau)$  antager værdier, der er tilstrækkeligt forskellige<sup>1</sup>. Kan man finde en sådan værdi af  $\tau$ , kan kildeserierne bestemmes entydigt, forudsat at antagelsen om lineær blanding og indbyrdes uafhængige kildeserier er opfyldt. Man skal dog holde sig for øje, at det ikke er oplagt, at disse antagelser er opfyldt for iskernedata, og at mit mål i øvrigt ikke er nødvendigvis at adskille kildeserierne optimalt fra hinanden, men at adskille den årlige komponent optimalt fra "resten". Det er således ikke sikkert, at en  $\tau$ -værdi, der honorerer kravet om velseparerede  $c_S(\tau)$ -værdier, samtidig giver kildeserier med det bedst mulige årlige signal.

Jeg vil i det følgende beskrive hvordan jeg har prøvet at bestemme de bedst mulige  $\tau$ -værdier. Jeg vil først i afsnit 4.2.1 beskrive hvordan jeg har fundet de værdier af  $\tau$ , der giver kildeserier med godt årligt signal ud fra en vurdering af de producerede kildeserier. Derefter vil jeg i afsnit 4.2.2 og 4.2.3 beskrive 2 metoder til at vælge  $\tau$  ud fra mere teoretiske betragtninger.

<sup>1</sup>Jeg har her indført en forsimpning af notationen. Det er værdierne  $\frac{c_{S_{ii}}(\tau)}{c_{S_{ii}}(0)}$ ,  $i \in \{1, \dots, \tilde{N}\}$  (diagonalen i ligning 2.7), der skal være forskellige, men for overskuelighedens skyld vil jeg i det følgende blot omtale kilde signalernes autokorrelationsfunktioner som  $c_S(\tau)$ , og dermed udelade såvel normalisering som indices. De steder hvor indices er nødvendige for sammenhængen, nøjes jeg med ét indeks, og skriver således  $c_{S_i}(\tau)$

### 4.2.1 Hvilke $\tau$ virker bedst ?

Det ville være bekvemt at kunne opstille et objektivt kriterium for hvad der kendetegner en kildeserie med et godt årligt signal, og med dette kriterium udvælge den værdi af  $\tau$ , der producerer kildeserier med det bedst mulige årlige signal. Jeg forsøgte først at opstille et objektivt mål for kvaliteten af det årlige signal ved at Fourier-transformere kildeserierne og undersøge hvilke serier, der i frekvens-regimet får store bidrag fra svingninger med perioder nær ved 1 år. Resultaterne viste ikke nogen tydelig sammenhæng mellem klarheden af det årlige signal og signalernes Fourier-spektre, hvilket jeg tror skyldes de korte dataserier samt de for iskernedata uundgåelige og markante variationer i årlagstykkelsen. Det er ikke lykkedes mig at finde en objektiv måde at vurdere kvaliteten af det årlige signal i vilkårlige serier. Istedet har jeg manuelt udvalgt kildeserierne med klart årligt signal, og derefter opstillet en metode til at afgøre hvilke af de udvalgte serier, der er bedst. Udvalgelsesproceduren foregik i 3 faser :

1. Jeg gennemgik systematisk alle  $\tau$ -værdier i et bestemt interval. For hver værdi af  $\tau$  vurderede jeg styrken af det årlige signal i kildeserierne, og kasserede de  $\tau$ -værdier, hvor de producerede kildeserier udviste mindre markant årlig variation end de oprindelige dataserier.
2. For hver af de resterende værdier af  $\tau$  udskrev jeg en graf med den kildeserie, der havde det mest markante årlige signal. Disse grafer sammenlignede jeg på kryds og tværs, hvorefter jeg udvalgte et antal serier, som jeg mente udviste klarere årlig variation end de oprindelige dataserier .
3. Kvaliteten af det årlige signal i disse serier sammenlignede jeg som nævnt ved hjælp af en mere objektiv og reproducerbar metode, som er beskrevet i afsnit 5.1.

Udvalgelsen i fase 1 og sorteringen i fase 2 er naturligvis behæftet med en hvis usikkerhed på grund af den indbyggede subjektivitet i vurderingen, men da det med blot lidt øvelse er nemt at kende en serie med tydelig årlig variation, finder jeg det sandsynligt at alle værdier af  $\tau$ , der giver en kildeserie med markant årligt signal, er medtaget. Opdelingen af vurderingen i 3 faser er foretaget for at sikre at udvalgelsen skete efter uændrede kriterier for forskellige værdier af  $\tau$ . Hvis sammenligningen i fase 2 var blevet foretaget sekventielt under databehandlingen, som for NGRIP-datasættet strakte sig over flere dage, ville der have været en overhængende risiko for, at kriterierne for hvad der kendetegner en god serie løbende ville have ændret sig.

Ved hjælp af ovenstående princip fandt jeg for hvert datasæt en række værdier af  $\tau$ , der sikrer god separation af det årlige signal fra de øvrige serier. Jeg har ikke undervejs anvendt kendskab til værdierne af  $c_S(\tau)$ , men udelukkende udvalgt serierne ud fra et ønske om at finde det bedst mulige årlige signal i kildeserierne.

### 4.2.2 Valg af $\tau$ ved anvendelse af formodninger om kildeseriernes natur

Selvom det ikke umiddelbart er muligt a priori at vælge en værdi af  $\tau$ , der sikrer tilstrækkelig forskellige værdier af  $c_S(\tau)$  fordi kildeserierne i sig selv er ukendte, kan vi bruge vores formodninger om kildeseriernes natur til at give et bud på gode  $\tau$ -værdier<sup>2</sup>. Lad os betragte  $c_S(\tau)$  for nogle mulige kildesignaler :

- Det årlige signal vil forhåbentlig være dominerende i en af kildeserierne. Havde den årlige svingning haft konstant periode (svarende til at årlagene havde konstant tykkelse), ville  $c_S(\tau)$  for denne kildeserie have lokale maksima for hele multipla af perioden og minima der imellem. Selvom årlagenes tykkelse varierer i forhold til gennemsnittet, må man stadig formode at den omtalte kildeseriens autokorrelationsfunktion har maksimum for  $\tau = f, 2f, 3f, \dots$  og minima for  $\tau = \frac{f}{2}, \frac{3f}{2}, \frac{5f}{2}, \dots$ , hvor  $f$  er det gennemsnitlige antal prøver pr. år, men det er sandsynligt at de svingende årlagstykkelser vil gøre disse ekstrema mindre udtalte for voksende  $\tau$ . I det følgende vil jeg fokusere på minimummet ved  $\tau = \frac{f}{2}$ , der må formodes at være markant til trods for betydelige variationer i årlagstykkelsen.
- $c_S(\tau)$  for kildeserier, der indeholder svingninger med kortere periode end det årlige signal, vil også have minimum for  $\tau = \frac{f}{2}$  hvis den gennemsnitlige periode er  $\frac{f}{3}, \frac{f}{5}, \frac{f}{7}, \dots$  eller tæt herpå. Der er dog svært at forstille sig at svingninger med periode  $\frac{f}{3}$  eller  $\frac{f}{5}$  skulle repræsentere et konkret fysisk signal og af denne grund skulle være en markant komponent i kildeserierne. De endnu hurtigere svingninger med periode  $\frac{f}{7}, \frac{f}{9}, \dots$  vil med stor sandsynlighed ikke være repræsenteret, både af fysiske årsager og på grund af at datasættenes opløsning ikke er god nok til at disse signaler overhovedet findes i datasættene. Svingningen med periode  $\frac{f}{2}$  kan derimod meget vel tænkes at være repræsenteret i en kildeserie, idet nogle af dataserierne til tider har to toppe pr. år.  $c_S(\tau)$  for en svingning med periode  $\frac{f}{2}$  vil have maksimum for  $\tau = \frac{f}{2}$ . Det er således ikke sandsynligt at der i kildeserierne findes markante bidrag fra svingninger med kortere perioder end  $f$  som samtidigt har autokorrelationsfunktioner med minimum for  $\tau = \frac{f}{2}$ .
- Kildeserier der er domineret af langsomme tendenser eller af svingninger med længere periode end  $f$  vil have større  $c_S(\frac{f}{2})$ -værdier end det årlige signal.
- For en serie der er domineret af støj vil  $c_S(\tau)$  antage numerisk små værdier for alle  $\tau > 0$  (for hvid støj og  $\tau \neq 0$  vil  $c_S(\tau) \rightarrow 0$  for  $T \rightarrow \infty$ , hvor  $T$  som tidligere nævnt er serielængden).

---

<sup>2</sup>Selvom  $\tau$  rent formelt er et tidsrum, angiver  $\tau$  jævnfør side 33 i praksis det antal pladser, man forskyder to dataserier i forhold til hinanden ved udregning af  $\widehat{C}_{\widehat{X}}(\tau)$ . Værdien af  $\tau$  repræsenterer derfor et antal prøver, der hver i middel svarer til  $f^{-1}$  år, hvor  $f$  er det gennemsnitlige antal prøver pr. år. Såvel værdier af  $\tau$  som svingningers perioder vil i det følgende blive angivet som enhedsløse tal, svarende til et antal prøver. Når jeg i det følgende skriver, at en svingning har periode 7 (hvor *periode* angiver et antal prøver), svarer det dermed til en gennemsnitlig periode på  $\frac{7}{f}$  (hvor *periode* nu angiver et tidsrum).

Sammenfattende kan vi sige at  $\tau = \frac{f}{2}$  formentlig vil give god separation af det årlige signal, idet  $c_S(\tau)$  for kildeserien med dominerende årligt signal har et markant minimum ved denne værdi, og dette minimum ikke umiddelbart deles med  $c_S(\tau)$  for andre formodede kildeserier, forudsat at det årlige signal ikke er dominerende i flere kildeserier.

Alternativt kan man jævnfør afsnit 2.4.4 anvende en kombination af flere  $\tau$ -værdier, hvor det vil være nærliggende at anvende kombinationen  $\tau = \frac{f}{2}$  og  $\tau = \frac{3f}{2}$ . Hvis man forestiller os at  $c_{S_i}(\frac{f}{2})$  for kildeserien med stærk årlig variation er næsten identisk med  $c_{S_j}(\frac{f}{2})$  for en anden kildeserie, der ikke er domineret af det årlige signal, og tilsvarende at  $c_{S_i}(\frac{3f}{2})$ -værdier er næsten identisk med  $c_{S_k}(\frac{3f}{2})$ , er det usandsynligt at  $j = k$ . Med andre ord er det usandsynligt at kildeserien med stærk årlig variation deler  $c_S(\tau)$ -minima med den samme kildeserie for både  $\tau = \frac{f}{2}$  og  $\tau = \frac{3f}{2}$  med mindre at denne anden serie også domineres af det årlige signal.

### 4.2.3 Valg af $\tau$ ved maksimering af separationen af $c_S(\tau)$ -værdierne

For en given værdi af  $\tau$  produceres et sæt kildeserier. Ideelt set burde dette sæt af kildeserier være uafhængigt af valget af  $\tau$ , så længe  $c_S(\tau)$ -værdierne for de forskellige kildeserier er tilstrækkeligt forskellige og dataserierne er så lange at de ikke-diagonale elementer i  $\tilde{\Sigma}$  fra ligning 2.23 er tæt på 0. For at kunne bestemme den værdi af  $\tau$ , der separerer  $c_S(\tau)$ -værdierne bedst muligt, har jeg forsøgt at finde et objektivt mål for separation af de  $\tilde{N}$   $c_S(\tau)$ -værdier. I det følgende er  $c_S(\tau)$ -værdierne sorteret, så de udgør en voksende følge. Jeg har afprøvet en række forskellige kriterier:

**Kriterium 1 :** Minimering af det gennemsnitlige afstandskvadrat  $\mathcal{M}_2$  mellem naboværdier defineret som

$$\mathcal{M}_p = \frac{1}{\tilde{N} - 1} \sum_{n=1}^{\tilde{N}-1} (c_{S_{n+1}}(\tau) - c_{S_n}(\tau))^p$$

med  $p = 2$ .

Jeg har tillige bestemt  $\tau$  ved at minimere  $\mathcal{M}_3$  og  $\mathcal{M}_4$  og ved at maksimere  $\mathcal{M}_{-1}$  og  $\mathcal{M}_{-2}$ .

**Kriterium 2 :** Minimering af det gennemsnitlige afstandskvadrat  $\mathcal{M}_2$  mellem alle værdier defineret som

$$\mathcal{M}_p = \frac{1}{(\tilde{N} - 1) + (\tilde{N} - 2) + \dots + 1} \sum_{n=1}^{\tilde{N}-1} \sum_{m=n+1}^{\tilde{N}} (c_{S_m}(\tau) - c_{S_n}(\tau))^p$$

Jeg har ligeledes forsøgt at minimere  $\mathcal{M}_3$  og at maksimere  $\mathcal{M}_{-1}$ .

**Kriterium 3 :** Minimering af spredningen af afstande mellem naboværdier eller afstandskvadrater. Afstandene mellem naboværdierne bestemmes som  $c_{S_{n+1}}(\tau) - c_{S_n}(\tau)$  hvor  $n = 1, 2, \dots, \tilde{N} - 1$  og afstandskvadraterne  $\mathcal{M}_2$  beregnes enten kun mellem naboværdier (som i kriterium 1) eller mellem alle par (som i kriterium 2). Spredningen af værdierne bestemmes derefter på sædvanlig vis.

For alle 3 kriterier søgte jeg  $\tau$  i intervallerne  $\tau \in [1, 20]$  for Site D,  $\tau \in [1, 15]$  for NGRIP og  $\tau \in [1, 30]$  for Summit, dvs. 2 - 3 gange det gennemsnitlige antal prøver pr. år  $f$ . Valget af overgrænse for  $\tau$ -intervallerne hviler på en afvejning mellem to hensyn. På den ene side skal  $\tau$  tillades at antage tilstrækkeligt store værdier til, at forskelle i autokorrelationsfunktionerne træder tydeligt frem. Min vurdering er, at en værdi af  $\tau$  svarende til  $2f - 3f$  er tilstrækkeligt. På den anden side blev der som omtalt på side 33 ved beregning af  $\widehat{C}_{\widehat{X}}(\tau)$  anvendt en tilnærmelse, hvor dataserierne værdier blev permuteret cyklisk istedet for at blive forskudt. Den usikkerhed, tilnærmelsen medfører, vokser med  $\tau$ , og man må derfor kræve at  $\tau \ll T$ , hvor  $T$  er serielængden, der i denne opgave typisk er omkring 150. Jeg vil derfor foretrække  $\tau$ -værdier under 10, men har ladet  $\tau$  gennemløbe større intervaller for at undersøge om højere værdier, til trods for den omtalt tilnærmelse, giver gode resultater.

For hvert af de 3 kriterier har jeg bestemt de værdier af  $\tau$  der gav bedst separation af  $c_S(\tau)$ -værdierne. Desuden har jeg gentaget bestemmelsen efter at have skaleret  $c_S(\tau)$ -værdierne således at den mindste værdi er -1 og den største værdi er 1. Dette har jeg gjort ud fra en formodning om at det er den relative separation og ikke den absolutte separation af  $c_S(\tau)$ -værdierne der har betydning.

### 4.3 Dynamisk Dekorrelation : trin for trin

Jeg har på nuværende tidspunkt beskrevet alle de elementer, der indgår i databehandlingen. Lad mig her kort sammenfatte i hvilken rækkefølge jeg foretager de enkelte trin

1. Et velegnet datasæt udvælges efter de kriterier, der er omtalt i afsnit 3.1.
2. Datasættet gennemgås for manglende værdier. Manglende værdier erstattes af estimerede værdier der findes ved lineær interpolation jævnfør afsnit 3.2.
3. Serierne regulariseres jævnfør afsnit 3.3 så ingen værdier afviger mere end  $3\sigma$  fra middelværdien.
4. Hvis der skal anvendes logaritmisk transformeret data, erstattes målt-nul-værdier af den anslåede detektionsgrænse jævnfør side 44, hvorefter logaritmen beregnes.
5. Serierne normaliseres således at middelværdien er 0 og spredningen 1.
6. Serierne rykkes i forhold til en fastholdt serie jævnfør afsnit 4.1, således at det årlige signal i serierne synkroniseres.
7. Dynamisk Dekorrelation gennemføres som beskrevet i afsnit på side 32 med et passende valg af  $\tau$ , eller gennemføres successivt for en række  $\tau$ -værdier. Beregningerne i Dynamisk Dekorrelation udføres i tre trin :
  - Den estimerede kvotientmatrix  $\widehat{Q}$  beregnes ud fra data ved hjælp af ligning 2.22.
  - $\widehat{Q}$  diagonaliseres, og der fremkommer en egen værdi-opløsning  $\widehat{Q} = \Phi \Omega \Phi^{-1}$ .
  - Blandingsmatricen og kildeserierne bestemmes som beskrevet ligning 2.24 og 2.25.



## Del 5

# Resultater og diskussion

I dette afsnit vil jeg præsentere resultaterne af Dynamisk Dekorrelation anvendt på de 3 datasæt. Resultaterne er samlet i et afsnit for hvert datasæt. Da fremgangsmåden under analysen i hovedtræk har været den samme for de 3 datasæt, er forklaringerne mest udførlige i afsnit 5.2, der indeholder resultaterne fra Site D, mens de efterfølgende afsnit 5.3 om NGRIP og 5.4 om Summit er mere kortfattede.

For hvert af de 3 datasæt har jeg jævnfør diskussionen i afsnit 4.2 udvalgt et antal  $\tau$ -værdier, der giver gode kilde serier. For at vurdere hvilke af disse værdier der er bedst, vil jeg opstille et kriterium for tælbarheden<sup>1</sup> af de årlige toppe i en serie. Dette gør det muligt at sammenligne kvaliteten af det årlige signal i de forskellige kilde serier, men giver også bedre muligheder for at vurdere i hvor høj grad det producerede årlige signal er klarere end det årlige signal i de oprindelige dataserier.

### 5.1 Et simpelt mål for tælbarhed

Jeg vil definere en top som et sammenhængende antal punkter med værdier over en bestemt tærskelværdi, mens de tilstødende datapunkters værdi skal være mindre end samme tærskelværdi. En top skal med andre ord krydse - og ikke blot tangere - tærskelværdien for at blive talt med. Tilsvarende skal værdien i minimummet mellem to toppe være mindre end tærskelværdien for at de to toppe ikke tælles som en. For en regelmæssig svingning som  $\sin(2\pi t)$ ,  $t \in [0, 10]$  vil antallet af toppe være 10 hvis tærskelværdien er i intervallet  $] - 1, 1[$  og nul ellers. Jeg vil tillægge et signal ”god tælbarhed”, hvis ovenstående metode giver samme antal toppe for et bredt interval af tærskelværdier og færre toppe for tærskelhøjder, der ligger uden for dette interval. Et signal med ”dårlig tælbarhed” er i kontrast hertil et signal hvor antallet af toppe afhænger kritisk af tærskelværdien, således at der ikke findes et bredt interval af tærskelværdier der giver samme antal toppe. Et signal med toppe, der er af sammenlignelig højde, og som

---

<sup>1</sup>Med *tælbarhed* mener jeg naturligvis ikke tælbarhed i matematisk forstand, hvor en tælbar mængde er en endelig mængde eller en mængde der er ækvipotent med  $\mathbb{N}$ .

er vel adskilt af relativt dybe minima, vil på denne måde blive tillagt god tælbarhed, mens et signal med meget uens toppe, og/eller toppe der delvis flyder sammen, vil blive tillagt dårlig tælbarhed.

Resultaterne fra optællingen af toppe vil blive præsenteret på en graf hvor antallet af toppe  $p$  er afbildet som funktion af tærskelhøjden  $l$ . God tælbarhed vil svare til et bredt plateau, hvor antallet af toppe er konstant for et bredt interval af tærskelværdier. Bemærk at det ikke er et krav til en serie med god tælbarhed at det konstante antal toppe svarer til det antal årlag, der ifølge den anvendte reference-tidsskala er i den givne datasekvens, men at de to antal naturligvis helst skulle stemme overens. Tilsvarende vil et signal med dårlig tælbarhed have en  $(l, p)$ -graf uden brede konstante stykker.

### 5.1.1 Filtrering

I såvel dataserierne som i de producerede kilde serier er der foruden de årlige toppe en del mindre toppe, typisk kun bestående af 1–2 datapunkter. Da den beskrevne tællemetode ikke skelner mellem brede og smalle toppe giver disse småtoppe, der oftest kun har en amplitude på  $\frac{1}{10} - \frac{1}{5}$  af gennemsnitshøjden for de årlige toppe, et for højt antal toppe. Der kan tillige forekomme overlejlrede langsomme svingninger eller svage lineære tendenser i kilde serierne, som forhindrer en vellykket bestemmelse af antallet af årlige toppe, da toppene tælles i forhold til et konstant tærskelniveau. For at mindske disse problemer har jeg filtreret såvel dataserier som kilde serier med et båndpasfilter inden optælling af toppe. Dataserierne er *ikke* filtreret inden der er foretaget Dynamisk Dekorrelation, men kun i forbindelse med optælling af toppe. Jeg har anvendt et 10. ordens Butterworth båndpasfilter og filtreret fase-neutralt med MatLabs `filtfilt`-funktion. Pasbåndets frekvens-interval er valgt som  $[0,55; 2,2]$  år<sup>-1</sup>, men ændringer af afskæringsfrekvenserne på op til 20% ændrer ikke afgørende på resultaterne. Ved analyse af flere og længere datasæt vil det være hensigtsmæssigt at vælge pasbåndets frekvensinterval ud fra en statistisk analyse af spredningen af årlagstykkelserne, således at hverken usædvanligt tykke eller tynde årlag filtreres væk. Jeg har imidlertid ikke på det foreliggende datagrundlag fundet det muligt at foretage andet end et skøn over velegnede afskæringsfrekvenser. Tanken var, at årlag med henholdsvis cirka halv og dobbelt tykkelse i forhold til gennemsnittet skulle slippe igennem filtreringen uden betydelig svækkelse, men de konkrete afskæringsfrekvenser er valgt ud fra en konkret vurdering af hvilke afskæringsfrekvenser der fjernede de uønskede småtoppe og langsomme svingninger fra kilde serierne.

Resultater af filtreringen er præsenteret mange steder i det følgende, men på figur 5.2 ses tydeligt effekten af bortfiltrering af småtoppe, mens det på figur 5.3 ses hvordan en overlejlret langsom svingning bortfiltreres med forbedret tælbarhed til følge.

### 5.1.2 Metodens begrænsninger

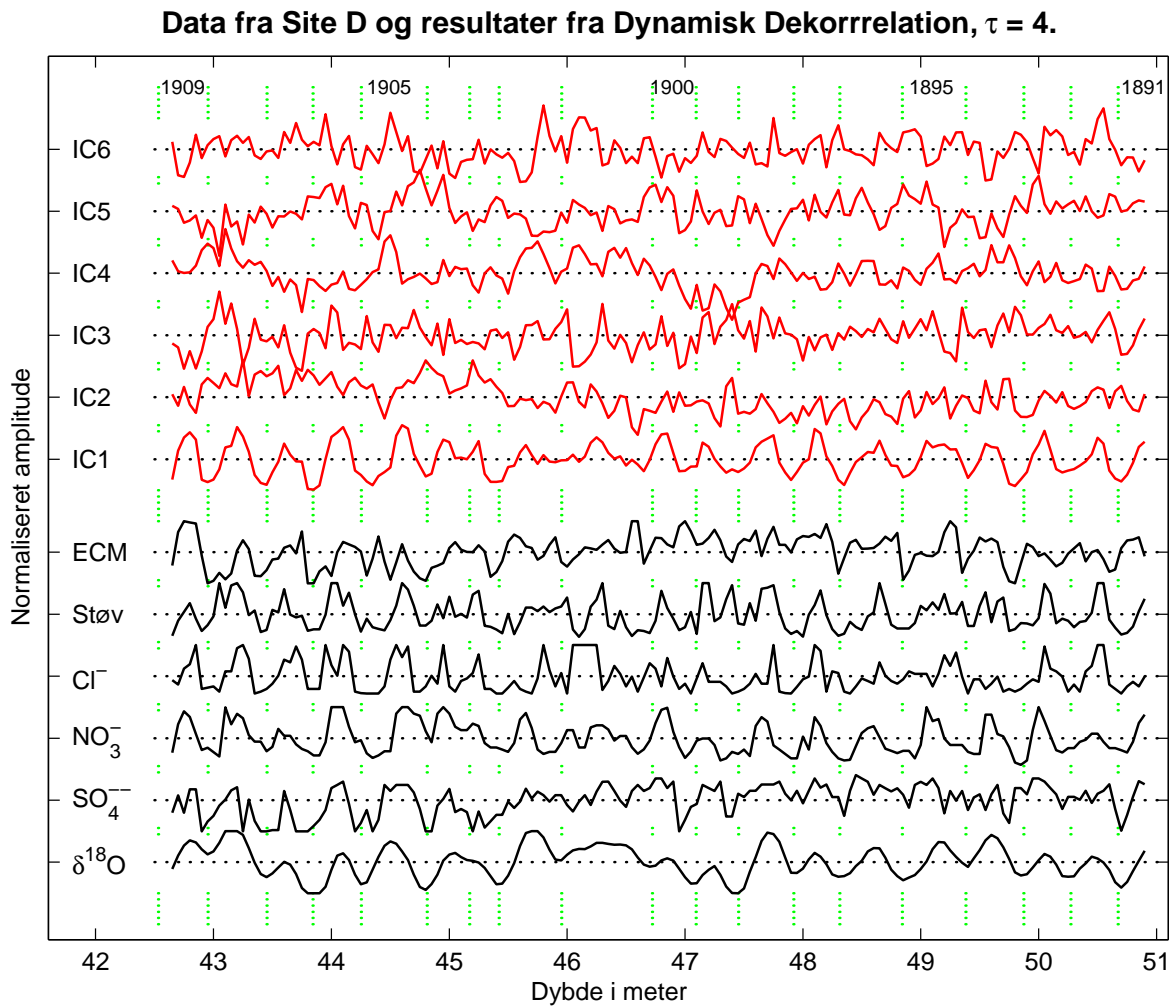
Den nævnte tællemetode anvendes i denne opgave til at vurdere kvaliteten af det årlige signal i både de oprindelige dataserier og i resultaterne fra Dynamisk Dekorrelation. Det er tanken at metoden skal give et fingerpeg om i hvilken grad Dynamisk Dekorrelation kan anvendes til dataforbedring før en egentlig datering foretages, og ikke at optællingsmetoden skal anvendes til datering i sig selv. Det er min formodning, at et signal, der får forbedret tælbarheden ved den anvendte simple tællemetode, vil være mere velegnet til dateringsformål med en vilkårlig mere avanceret tællemetode, da forbedret tælbarhed i den her anvendte betydning svarer til, at signalets årlige toppe er mere ensartede og velseparerede. Tællemetoden skal derfor ses som et skridt i retning af måling af det årlige signals styrke og kvalitet, egenskaber der er svære at beskrive objektivt.

## 5.2 Resultater for Site D datasættet

Det særlige ved Site D datasættet i forhold til de to andre datasæt i denne opgave er, at det består af fysisk meget forskellige signaler, nemlig  $\delta^{18}\text{O}$ , ECM og støv, som er målt uafhængigt af hinanden, samt de 3 ion-serier der er målt på en ionkromatograf. Dataserierne er afbildet nederst på figur 5.1. Det ses, at det årlige signal i  $\delta^{18}\text{O}$  og  $\text{NO}_3^-$ -serierne er nogenlunde klart, idet de årlige toppe er veldefinerede og nogenlunde adskilte, dog med undtagelse af årlaget 1903–1904. Som beskrevet i afsnit 4.1 har det også vist sig, at synkronisering af det årlige signal relativt til  $\delta^{18}\text{O}$  giver det bedste resultat, hvilket formentligt skyldes regelmæssigheden af svingningen i dette signal. Figur 5.1 viser foruden datasættet de 6 producerede kildeserier ved anvendelse af  $\tau = 4$ .

Jeg har ladet  $\tau$  gennemløbe intervallet  $[1, 20]$  og som beskrevet i afsnit 4.2.1 udvalgt de værdier, der gav kildeserier med stærk årlig variation. I fase 1 vurderede jeg manuelt for hver  $\tau$ -værdi om en af de producerede kildeserier indeholdt et årligt signal, der virkede mindst lige så klart som hvad der ses i de oprindelige  $\delta^{18}\text{O}$  og  $\text{NO}_3^-$ -dataserier. I fase 2 udskrev jeg grafer med kildeserier for disse  $\tau$ -værdier (værdierne 3, 4, 5, 6, 8, 9, 10, 13, 14, 15, 18 og 21), og skrev  $\tau$  på bagsiden af graferne. Jeg blandede herefter siderne med grafer og udvalgte efter indbyrdes sammenligninger på kryds og tværs de 5 grafer, jeg synes havde det bedste årlige signal. Ved udvælgelsen fokuserede jeg især på kildeseriernes opførsel ved det specielle 1901-årlag (dybde 45,95 – 46,70 meter) og om de to årlag 1903–1904 (dybde 44,80 – 45,15 og 45,15 – 45,45 meter) blev klart adskilt. De 5 bedste værdier var 4, 6, 10, 14 og 15. For at afgøre hvilken af disse 5 resultater der var bedst, anvendte jeg den i afsnit 5.1 indførte tællemetode.

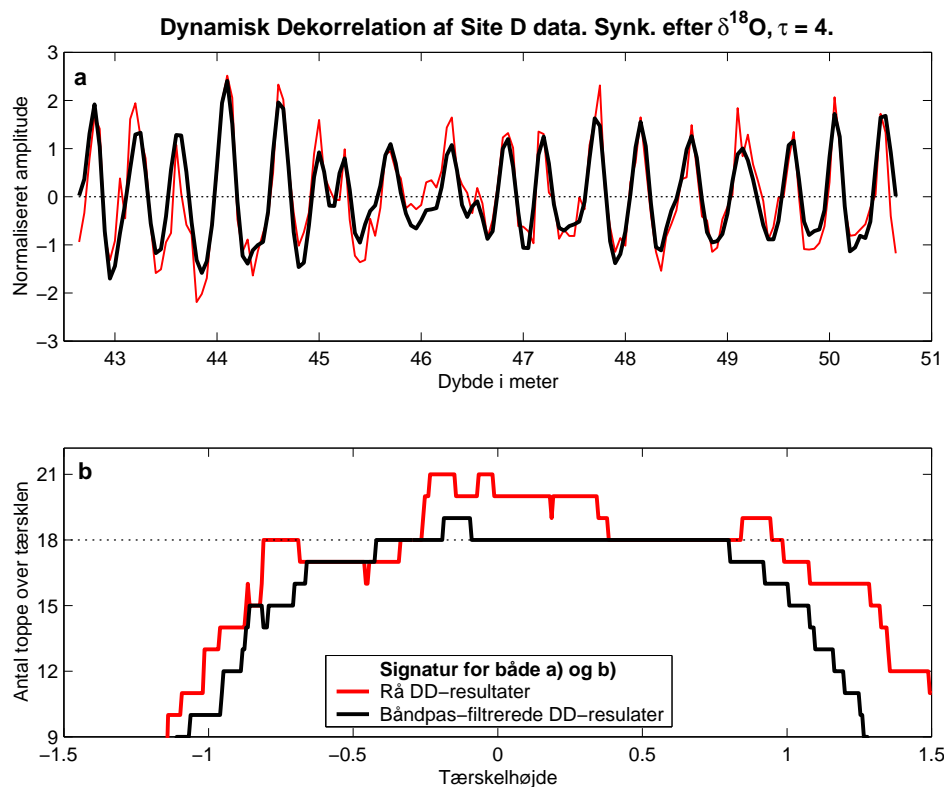
Grunden til at jeg ikke har anvendt tællemetoden gennem hele udvælgelsesproceduren er, at serierne skal beskæres i enderne for optællingsresultatet kan sammenlignes, både i forhold til hinanden og i forhold til det antal årlige toppe, serierne bør indeholde ifølge referencetidsskalaen. Dataserierne indeholder nemlig ikke kun akkurat perioden 1889–1912, men yderligere



Figur 5.1: Resultater fra Dynamisk Dekorrrelation af data fra Site D (IC1 - IC6), synkroniseret efter  $\delta^{18}\text{O}$  og med  $\tau = 4$ . Til sammenligning nederst det oprindelige datasæt. Den viste datering stammer fra [Steffensen, 1988].

et eller to år i hver ende. For at resultatet af optællingen af antallet af toppe skal kunne sammenlignes, er det derfor vigtigt at serierne beskæres så de dækker det samme antal årlag. Denne beskæring er ikke triviell, da serierne under synkroniseringsproceduren som beskrevet i afsnit 4.1 bliver forskudt op til  $3/4$  år indbyrdes, og beskæringen må derfor foregå manuelt, hvilket er meget tidskrævende for et stort antal serier. Desuden skal den relevante kilde serie under alle omstændigheder udvælges fra de 6 serier, der produceres for hvert  $\tau$ . Jeg synes sammenfattende at man opnår et rimeligt forhold mellem arbejdsmængde og objektivitet ved at foretage den beskrevne grovsortering manuelt og anvende en reproducerbar optællingsmetode til den mere detaljerede sammenligning.

Ingen af de 5 værdier gav  $(l, p)$ -grafer med et helt fladt plateau, men for  $\tau = 4$  opnås for tærskelværdier mellem  $-0,43$  og  $-0,19$  og mellem  $-0,08$  og  $0,80$  de 18-toppe, referencetidskalaen foreskriver. Resultatet ses på figur 5.2, hvor både kilde serien (IC1 fra figur 5.1), den filtrerede

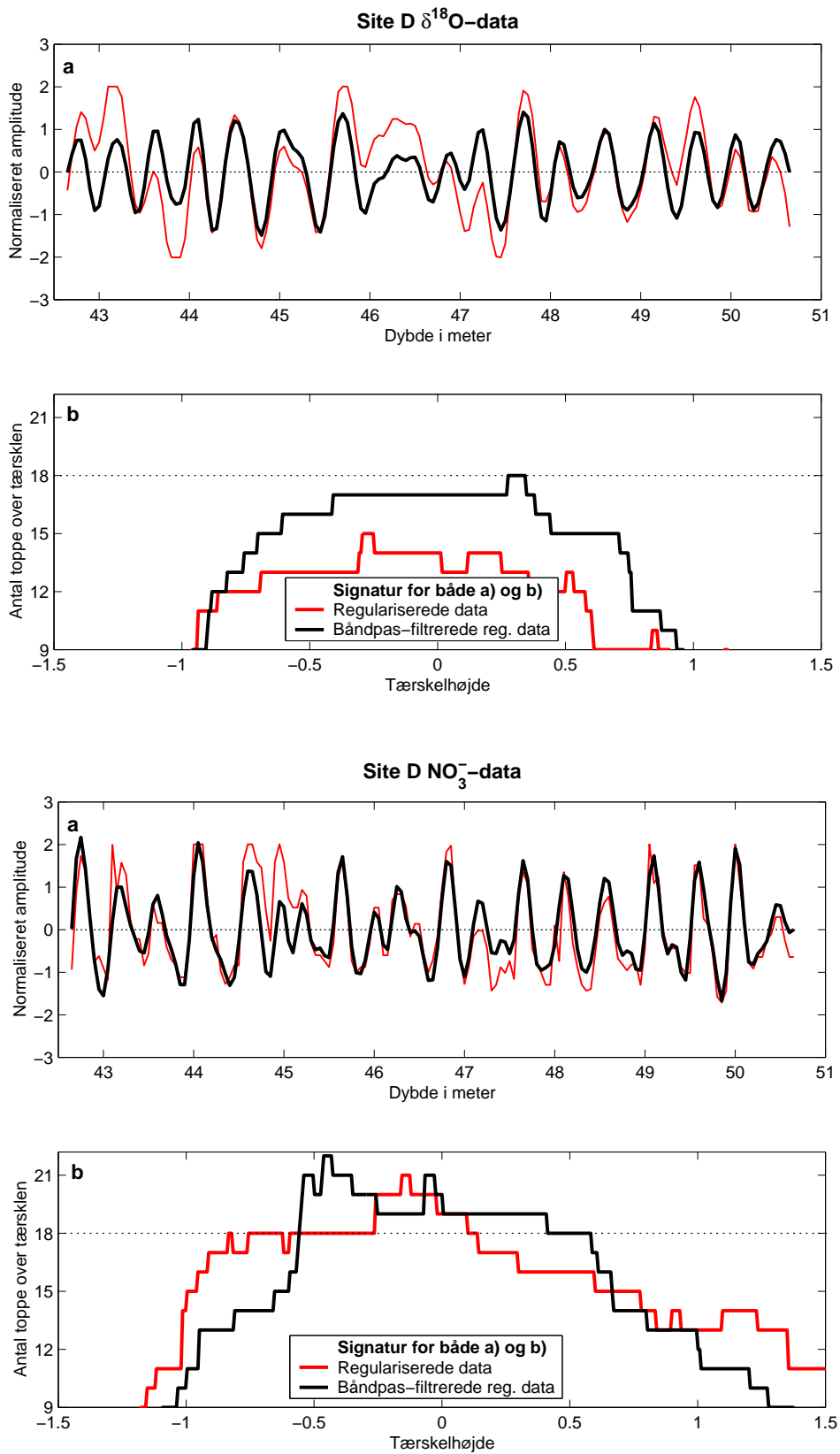


Figur 5.2: (a) Resultater fra Dynamisk Dekorrelation af data fra Site D, synkroniseret efter  $\delta^{18}\text{O}$  og med  $\tau = 4$ . (b) Antallet af toppe bestemt med metoden fra afsnit 5.1. Filtreeringen er foretaget som beskrevet i afsnit 5.1.1.

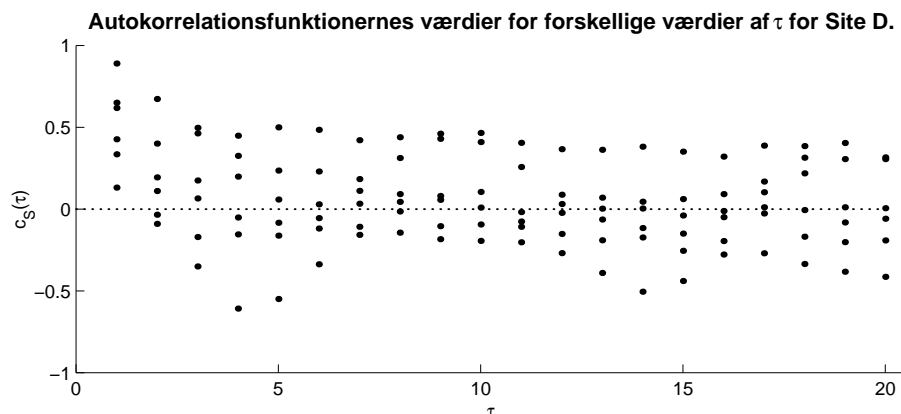
kildeserie og resultatet af optællingen af toppe i henholdsvis den rå og den filtrede kildeserie er samlet.

Det er det specielle 1901-årlag, der er kilde til den ekstra top, som giver 19 toppe for  $-0,19 < l < -0,08$ , og det ses da også af figur 5.1 at dette årlag i alle datasierne består af flere mindre toppe. H. B. Clausen, der har udarbejdet den anvendte referencetidsskala, har bekræftet at dette stykke af signalet med stor sikkerhed kun repræsenterer et årlag. Det ses at 1901-laget er mere veldefineret i kildeserien end i de fleste af de oprindelige datasier, hvilket vil sige at Dynamisk Dekorrelation i dette tilfælde har mindsket den tvetydighed, der er ved identifikation af årlag. Tilsvarende er årlagene 1903–1904 meget klarere adskilt i kildeserien end i  $\delta^{18}\text{O}$ , ECM og  $\text{SO}_4^{--}$ , hvilket indikerer at Dynamisk Dekorrelation faktisk er i stand til at samle et *fælles* årligt signal fra datasierne.

For at vurdere i hvor høj grad det årlige signal er blevet forbedret i forhold til det årlige signal i de oprindelige datasier har jeg talt antallet af toppe i datasierne med samme tællemetode og filtrering. Bedst tælbarhed opnås for  $\delta^{18}\text{O}$  og  $\text{NO}_3^-$ -serierne. Resultaterne ses på figur 5.3, hvoraf det klart ses at  $\delta^{18}\text{O}$ -signalet er kendetegnet ved god tælbarhed hvis det filtreres som beskrevet, men at årlagene 1903–1904 optræder som en top med en ”skulder”, og derfor kun tælles som en top. En top med skulder som denne forekommer ofte, når to



Figur 5.3: (a) De regulariserede  $\delta^{18}\text{O}$  (øverst) og  $\text{NO}_3^-$ -dataserier (nederst) fra Site D udsat for filtrering på samme måde som resultaterne fra Dynamisk Dekorrelation i figur 5.2. (b) Tælbarheden af årlag i de samme serier.



Figur 5.4: Værdier af kildesignalernes autokorrelationsfunktioner for forskellige værdier af  $\tau$  ved Dynamisk Dekorrelation af data fra Site D.

årlige toppe af den ene eller anden grund ikke er tilstrækkeligt adskilt eller når diffusion (se appendix B) har udvisket minimummet mellem de to toppe. Det ses desuden at  $\text{NO}_3^-$ -serien har dårlig tælbarhed uanset om den filtreres, idet antallet af toppe ikke er velbestemt for noget interval af tærskelværdier. Ved sammenligning af figur 5.2 og figur 5.3 ses det, at resultaterne fra Dynamisk Dekorrelation har væsentligt bedre tælbarhed end de oprindelige serier, og at Dynamisk Dekorrelation dermed forbedrer klarheden af det årlige signal betydeligt.

Hvad angår valget af  $\tau$  ud fra formodninger om kildeseriernes natur som beskrevet i afsnit 4.2.2, ses det at den forventede gode  $\tau$ -værdi  $\tau = \frac{f}{2} = \frac{8,8}{2} \approx 4$  i det ovenstående netop blev udvalgt som værdien der producerer de bedste kildeserier. For at gå lidt mere i detaljer betragtes figur 5.4, der viser værdierne af de 6 kildesignalers autokorrelationsfunktioner<sup>2</sup> for værdier af  $\tau$  mellem 1 og 20. Selvom punkterne (jævnfør fodnoten nederst) ikke kan forbindes med kurver, kan der anes en svingning med minimum for  $\tau = 4$ , maksimum for  $\tau = 9$  og minimum igen for  $\tau = 14$ .  $c_S(\tau)$  for et signal med perioden  $f = 8,8$  prøver pr. år vil netop have denne opførsel, selvom det andet minimum burde være placeret ved  $\tau = \frac{3}{2} \cdot 8,8 \approx 13$  istedet for  $\tau = 14$ . Dette kan tages som indtægt for at den årlige komponent er nogenlunde velrepræsenteret i mindst en kildeperiode for de fleste  $\tau$ . Jeg synes ikke at man med rimelig sikkerhed kan identificere andre tendenser i punkternes placering på figuren, selvom fragmenter af andre svingninger ser ud til at være til stede. Idet de to lave  $c_S(\tau)$ -værdier ved  $\tau = 4$  og  $\tau = 14$  efter alt at dømme stammer fra den kildeperiode, der indeholder den stærkeste komponent af det årlige signal, må man forvente god separation af det årlige signal netop hvor  $c_S(\tau)$ -værdierne for dette signal er bedst adskilt fra de øvrige værdier. Den antydede svingnings værdier er godt adskilt fra de resterende værdier for  $\tau$ -værdierne 3 – 6 og 13 – 15, hvilket

<sup>2</sup>Det er vigtigt at notere sig, at kildesignalerne ikke er de samme for de forskellige værdier af  $\tau$ , og at det derfor ikke giver mening at opfatte punkterne på figuren som værdier for de samme 6 autokorrelationsfunktioner. For hver værdi af  $\tau$  beregnes et sæt kildeserier, og figuren viser de 6  $c_S(\tau)$ -værdier for netop det valgte  $\tau$ . Kun i den optimale situation hvor separationen af kildeserierne er perfekt for alle  $\tau$ , er kildeserierne ens, og punkterne på figur 5.4 er da blot  $c_S(\tau)$  for disse kildeserier.

passer glimrende med de  $\tau$ -værdier (værdierne 3, 4, 5, 6, 8, 9, 10, 13, 14, 15, 18 og 21), der blev fundet ovenfor ved vurdering af kvaliteten af det årlige signal i kildeserierne. Ved sammenligning af de to metoders bud på gode værdier er kun  $\tau = 10$  lidt overraskende, idet det på figur 5.4 ser ud til at det årlige signals  $c_S(\tau)$ -værdi ligger relativt tæt på en anden kildeseries  $c_S(\tau)$ -værdi.

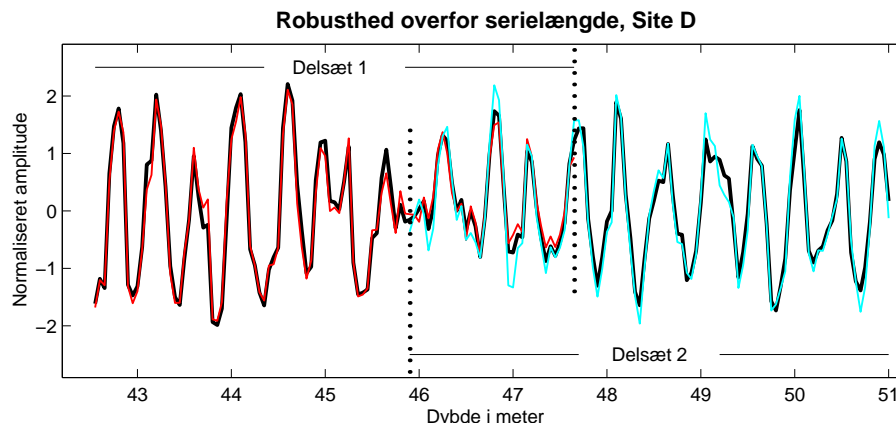
At vælge  $\tau$  ved at søge at sprede  $c_S(\tau)$ -værdierne godt indbyrdes, som det blev foreslået i afsnit 4.2.3, giver ikke  $\tau$ -værdier, der stemmer overens med de ovenfor opnåede resultater. Hvilke værdier der er bedst, afhænger naturligvis af hvilket af de i afsnit 2.4.4 opstillede kriterier, der anvendes, men  $\tau = 4$  er ikke blandt de 4 værdier, der separerer  $c_S(\tau)$ -værdierne bedst, uanset hvilket af de opstillede kriterier, der vælges. Ved minimering af f.eks.  $\mathcal{M}_2$  som defineret i kriterium 1 på side 55 bestemmes de fire bedste  $\tau$ -værdier til  $\tau = 18, 19, 15, 6$  hvis  $c_S(\tau)$ -værdierne først skaleres, så den største er 1 og den mindste er  $-1$ , og til  $\tau = 8, 16, 7, 12$  hvis skalering udelades. Heraf producerer Dynamisk Dekorrelation kildeserier med stærk - men ikke optimal - årlig variation for  $\tau = 6, 15$ , med nogenlunde klar årlig variation for  $\tau = 8, 18$ , mens  $\tau = 7, 12, 16, 19$  giver kildeserier med markant svagere årlig variation, end hvad der observeres i de oprindelige dataserier. Der kan således ikke konstateres noget generelt sammenfald mellem kildeserier med stærkt årligt signal og velseparerede  $c_S(\tau)$ -værdier med de anvendte kriterier for separation.

Afslutningsvis vil jeg præsentere resultatet af et følsomhedsstudium. Som beskrevet i afsnit 2.1 vil det af flere grunde formentlig være nødvendigt at gennemføre Dynamisk Dekorrelation på delmængder af datasæt, specielt hvis datasættet dækker et stort dybdeinterval. Det vil i dette tilfælde være uheldigt hvis måden, delmængderne vælges på, har afgørende indflydelse på resultaterne. Tilsvarende er afgrænsningen af datasættene i denne opgave dikteret af praktiske hensyn, og det vil være yderst uheldigt hvis f.eks. mindre variationer i længden af dataserierne ændrer markant på resultaterne. Jeg har undersøgt om afgrænsningen har betydning ved at dele Site D datasættet i to delsat og gennemføre Dynamisk Dekorrelation for såvel hele datasættet som de to delsat. Synkroniseringen er foretaget efter  $\delta^{18}\text{O}$  af hvert enkelt delsat for sig, og der anvendes  $\tau = 4$  i alle tilfældene. På figur 5.5 ses resultaterne af Dynamisk Dekorrelation af henholdsvis hele serien og af de første og sidste 60% af datasættet.

Det ses at der er næsten fuldstændig overensstemmelse mellem resultaterne fra analyse af hele datasættet og hver af de to delsat. Også ved sammenligning af de to delsat indbyrdes er overensstemmelsen mellem resultaterne bemærkelsesværdig, når man tager i betragtning, at kun de midterste 20% af det samlede antal datapunkter er inkluderet i begge delsat. Forskellene mellem resultaterne for de to delsat er begrænset til mindre uoverensstemmelser mellem tophøjder og minimumsdybder, og har ikke afgørende betydning for klarheden af det årlige signal.

Jeg har lavet tilsvarende analyser for NGRIP-data, og jeg har også forsøgt at opdele Site D datasættet i op til 10 overlappende stykker uden at kunne konstatere afgørende forskelle. Det kan således konkluderes, at Dynamisk Dekorrelation af iskernedata fungerer uafhængigt





Figur 5.5: Følsomhed overfor serielængde ved Dynamisk Dekorrelation. Den sorte kurve er resultatet af Dynamisk Dekorrelation af Site D data med  $\tau = 4$  (IC1 på figur 5.1). De røde hhv. cyan-blå kurver er resultaterne fra Dynamisk Dekorrelation af de første hhv. sidste 60% af datasættet.

af enkeltheder omkring afgrænsningen af datasættene, og at analysen derfor må forventes at kunne anvendes på lange datasæt ved sammenstykning af resultater fra analyse af kortere, eventuelt overlappende, sekvenser. Desuden indikerer resultatet, at blandingsmatricen  $\mathbf{A}$  er stationær med god tilnærmelse, hvilket i sig selv er et vigtigt resultat.

### 5.3 Resultater for NGRIP-datasættet

Arbejdet med at udvælge den bedste  $\tau$ -værdi ud fra vurdering af resultaterne af Dynamisk Dekorrelation af NGRIP-data var noget mere krævende end for Site D, idet det som beskrevet i afsnit 4.1 ikke entydigt var muligt at afgøre hvilken dataserie, det var bedst at fastholde ved synkronisering af det årlige signal. Jeg gennemprøvede derfor, med hver af de 12 dataserier fastholdt,  $\tau$ -værdier mellem 1 og 15. Hvilke af de 180 kombinationer, der gav kilde serier med klar årlig variation fremgår af tabel 5.1. De kombinationer, der ikke er markeret i tabellen, gav kilde serier med et årligt signal, som jeg vurderede var dårligere end de årlige signaler i de oprindelige dataserier. De kombinationer af fastholdt serie og  $\tau$ -værdi, der gav kilde serier med årlig svingning af sammenlignelig kvalitet i forhold til dataserierne, er markeret med  $\square$  i tabellen, mens en forbedring af det årlige signal er markeret med  $\boxtimes$ . Ud af de 13 gode kombinationer udvalgte jeg de 6 bedste, markeret med  $\blacksquare$  i tabellen, på samme måde som beskrevet for Site D, idet jeg uden at kende  $\tau$  og den fastholdte serie sammenlignede grafer på kryds og tværs.

I fase 3 talte jeg antallet af toppe i de 6 resterende serier med den beskrevne tællemetode for at afgøre hvilken kombination af  $\tau$  og fastholdt serie, der gav kilde serie med bedst årligt signal.

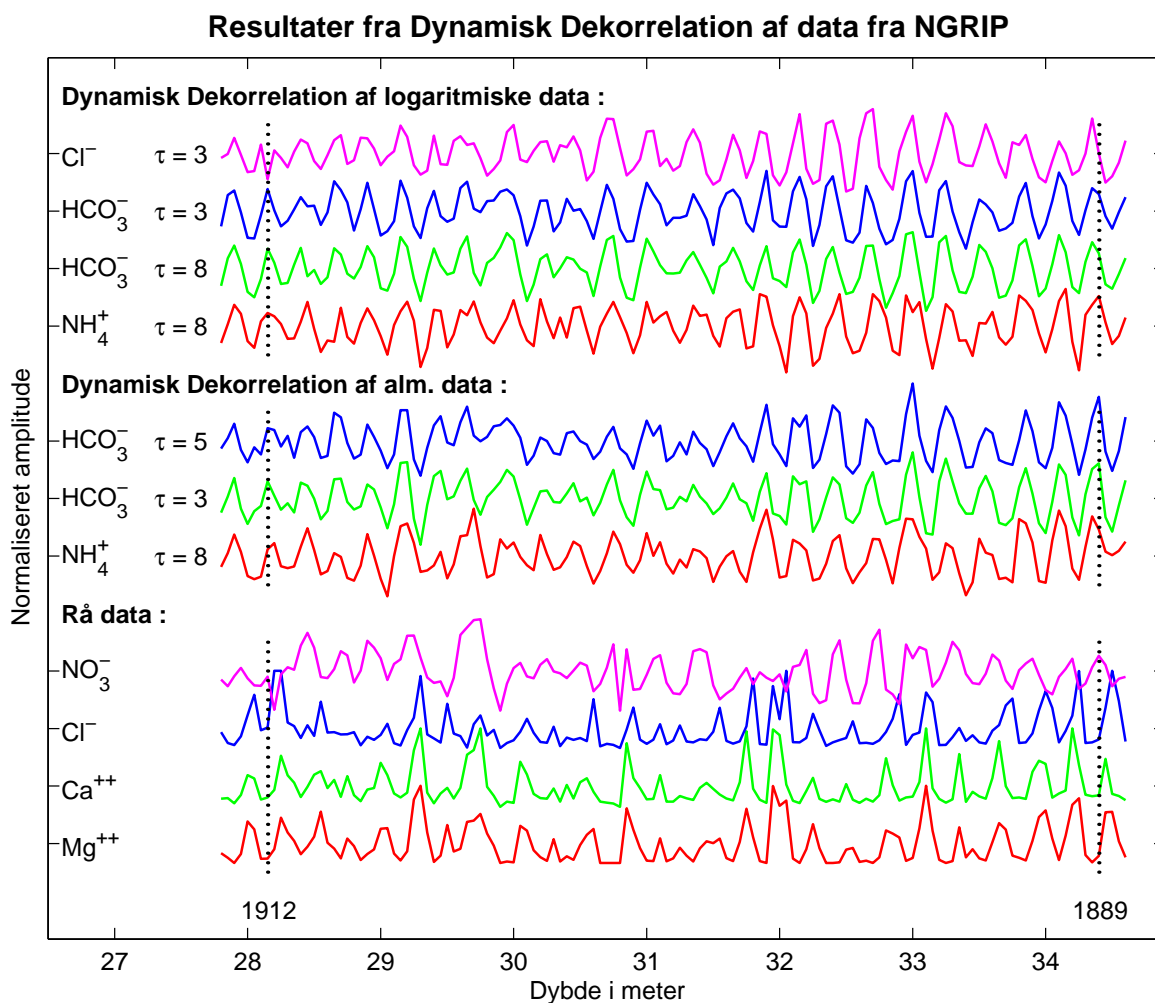
Jeg gentog hele ovenstående procedure for NGRIP datasættets logaritmisk transformerede variant. Resultaterne er helt parallelle, idet det dog i tabel 5.2 bemærkes at der er lidt flere

Fastholdt serie / $\tau$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Li <sup>+</sup>			<input type="checkbox"/>			<input type="checkbox"/>					<input type="checkbox"/>				
Na <sup>+</sup>		<input type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>						<input type="checkbox"/>	
NH <sub>4</sub> <sup>+</sup>			<input checked="" type="checkbox"/>		<input type="checkbox"/>			<input checked="" type="checkbox"/>			<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	
K <sup>+</sup>		<input type="checkbox"/>	<input type="checkbox"/>		<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input checked="" type="checkbox"/>				
Mg <sup>++</sup>			<input type="checkbox"/>								<input type="checkbox"/>				
Ca <sup>++</sup>		<input type="checkbox"/>	<input type="checkbox"/>					<input checked="" type="checkbox"/>			<input type="checkbox"/>				
F <sup>-</sup>		<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>			<input type="checkbox"/>						<input type="checkbox"/>	
MSA			<input checked="" type="checkbox"/>		<input type="checkbox"/>			<input type="checkbox"/>							
Cl <sup>-</sup>		<input type="checkbox"/>						<input type="checkbox"/>							
HCO <sub>3</sub> <sup>-</sup>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		<input type="checkbox"/>					
NO <sub>3</sub> <sup>-</sup>		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>						<input type="checkbox"/>	
SO <sub>4</sub> <sup>-</sup>			<input type="checkbox"/>					<input type="checkbox"/>					<input type="checkbox"/>	<input type="checkbox"/>	

Tabel 5.1: Oversigt over hvilke kombinationer af fastholdt serie og værdi af  $\tau$ , der giver gode resultater ved Dynamisk Dekorrelation af NGRIP data. Kombinationer, der er mærket , giver rimelige resultater, mens  angiver kombinationer, der giver gode resultater. De 6 bedste serier, der blev udvalgt til videre vurdering, er mærket . Se i øvrigt teksten for en uddybende forklaring.

Fastholdt serie / $\tau$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Li <sup>+</sup>		<input type="checkbox"/>													
Na <sup>+</sup>		<input type="checkbox"/>			<input checked="" type="checkbox"/>			<input type="checkbox"/>						<input type="checkbox"/>	
NH <sub>4</sub> <sup>+</sup>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>		<input type="checkbox"/>			<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	
K <sup>+</sup>		<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>			<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	
Mg <sup>++</sup>		<input type="checkbox"/>	<input type="checkbox"/>												
Ca <sup>++</sup>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>					<input type="checkbox"/>			<input type="checkbox"/>		<input type="checkbox"/>		
F <sup>-</sup>		<input type="checkbox"/>	<input checked="" type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>			<input type="checkbox"/>				
MSA		<input type="checkbox"/>	<input type="checkbox"/>		<input checked="" type="checkbox"/>						<input type="checkbox"/>			<input type="checkbox"/>	
Cl <sup>-</sup>			<input checked="" type="checkbox"/>					<input type="checkbox"/>			<input type="checkbox"/>				
HCO <sub>3</sub> <sup>-</sup>			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input checked="" type="checkbox"/>			<input type="checkbox"/>			<input type="checkbox"/>	
NO <sub>3</sub> <sup>-</sup>			<input type="checkbox"/>					<input type="checkbox"/>			<input type="checkbox"/>				<input type="checkbox"/>
SO <sub>4</sub> <sup>-</sup>		<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>			<input type="checkbox"/>							

Tabel 5.2: Som tabel 5.1, blot for logaritmisk transformerede data.



Figur 5.6: Resultater fra Dynamisk Dekorrelation af almindelige samt logaritmisk transformerede data fra NGRIP med forskellige valg af synkronisering og  $\tau$ -værdi. De angivne årlag markerer sommerens placering de pågældende år.

kombinationer af fastholdt serie og  $\tau$ -værdi, der giver pæne resultater. De 5 bedste kombinationer blev udvalgt på samme måde som beskrevet ovenfor for ikke-logaritmiske data og vil blive analyseret i det følgende. I figur 5.6 ses et udvalg af de bedste serier. Til sammenligning er 4 af de oprindelige dataserier medtaget. Det ses at ændringer i valget af fastholdt serie og  $\tau$ -værdi blandt de viste serier kun medfører mindre ændringer i de producerede kilde serier, og at de årlige signaler i de viste serier umiddelbart virker sammenlignelige. Hvilken af de viste serier der er bedst, må derfor afgøres ved optælling af toppe som beskrevet i afsnit 5.1.

Hvad angår synkroniseringen af serierne viser tabel 5.1 og 5.2 at det er de samme  $\tau$ -værdier, der giver gode kilde serier, uanset hvilken serie der fastholdes. Dette må siges at være forudsætningen for at synkroniseringen overhovedet kan foretages. Der er dog forskelle på hvor godt de enkelte serier egner sig til formålet. Umiddelbart ser  $\text{HCO}_3^-$  ud til at være en anelse bedre end de øvrige, men mens  $\text{Li}^+$  og  $\text{Mg}^{++}$  direkte er uegnede. Jeg har ikke nogen god forklaring på hvad der gør  $\text{HCO}_3^-$  mere velegnet end de andre serier, men ved

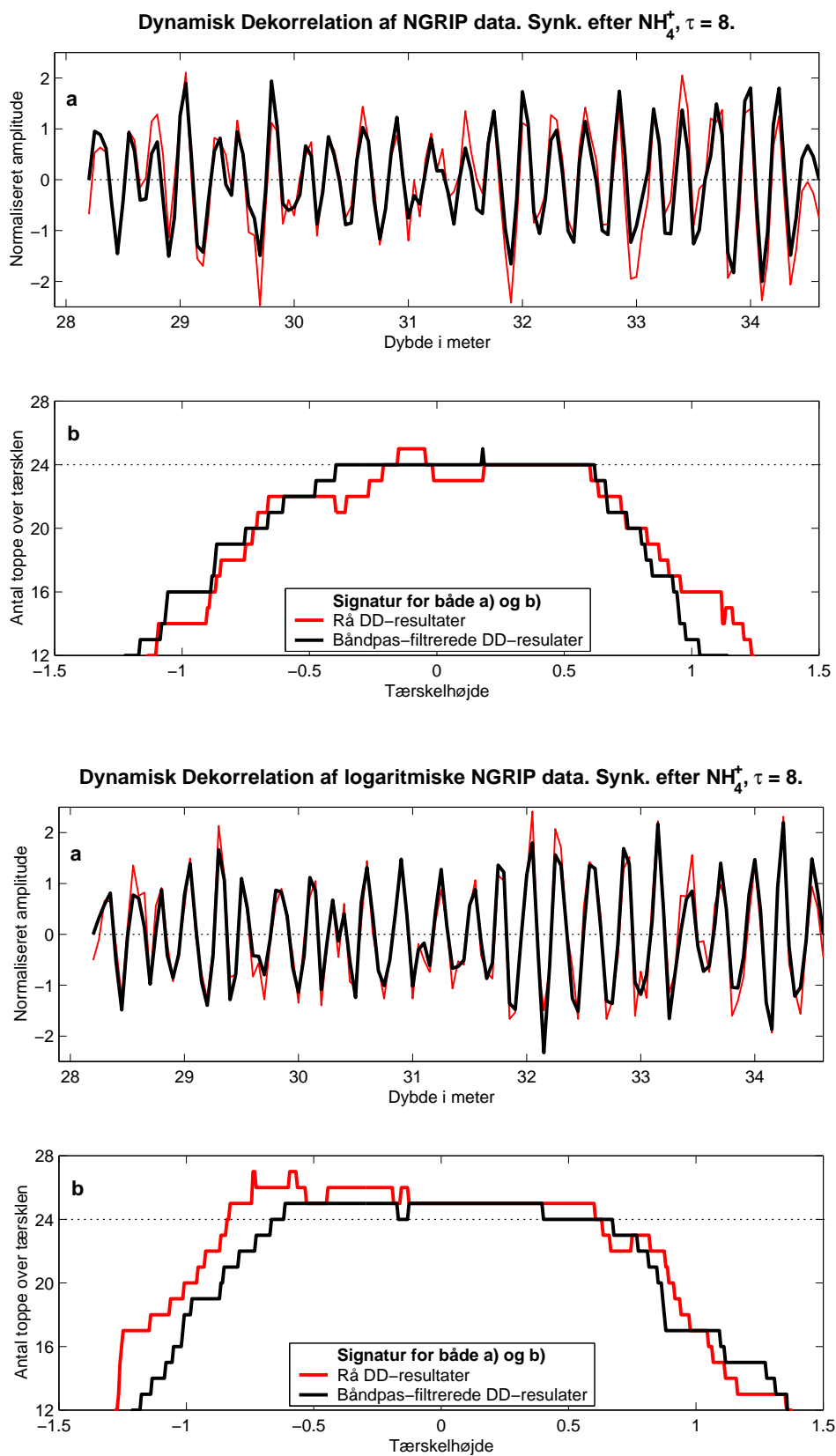
at betragte de regulariserede dataserier på figur 3.1 ses det at  $\text{HCO}_3^-$ -serien ikke udviser en særlig klar årlig svingning, men at den har to dybe dyk ved knap 32 meters dybde. Disse to hak kan tænkes at være dem der ”styrer” synkroniseringen fordi de giver store bidrag til krydskorrelationsfunktionerne.

For almindelige data viste kombinationen af synkronisering efter  $\text{NH}_4^+$  og  $\tau = 8$  sig at give den bedste tælbarhed. Resultatet er vist øverst i figur 5.7. De tilsvarende resultater ved anvendelse af logaritmisk transformeret data er vist nederst til sammenligning. Tælbarheden af det årlige signal er god, idet  $(l, p)$ -grafene har et bredt plateau, hvor antallet af toppe er 24 som det må forventes ifølge referencetidsskalaen. Plateauet strækker sig fra  $-0,40$  til  $0,62$  kun afbrudt af en meget smal spids ved  $l = 0,17$ , der skyldes et hak i toppen ved 31,30 meters dybde. For de logaritmisk transformeret data ses også god tælbarhed, men antallet af toppe er for højt i forhold til referencetidsskalaen på grund af den tvedelte top i 30,35 meters dybde og den lille ekstratop i 31,15 meters dybde.

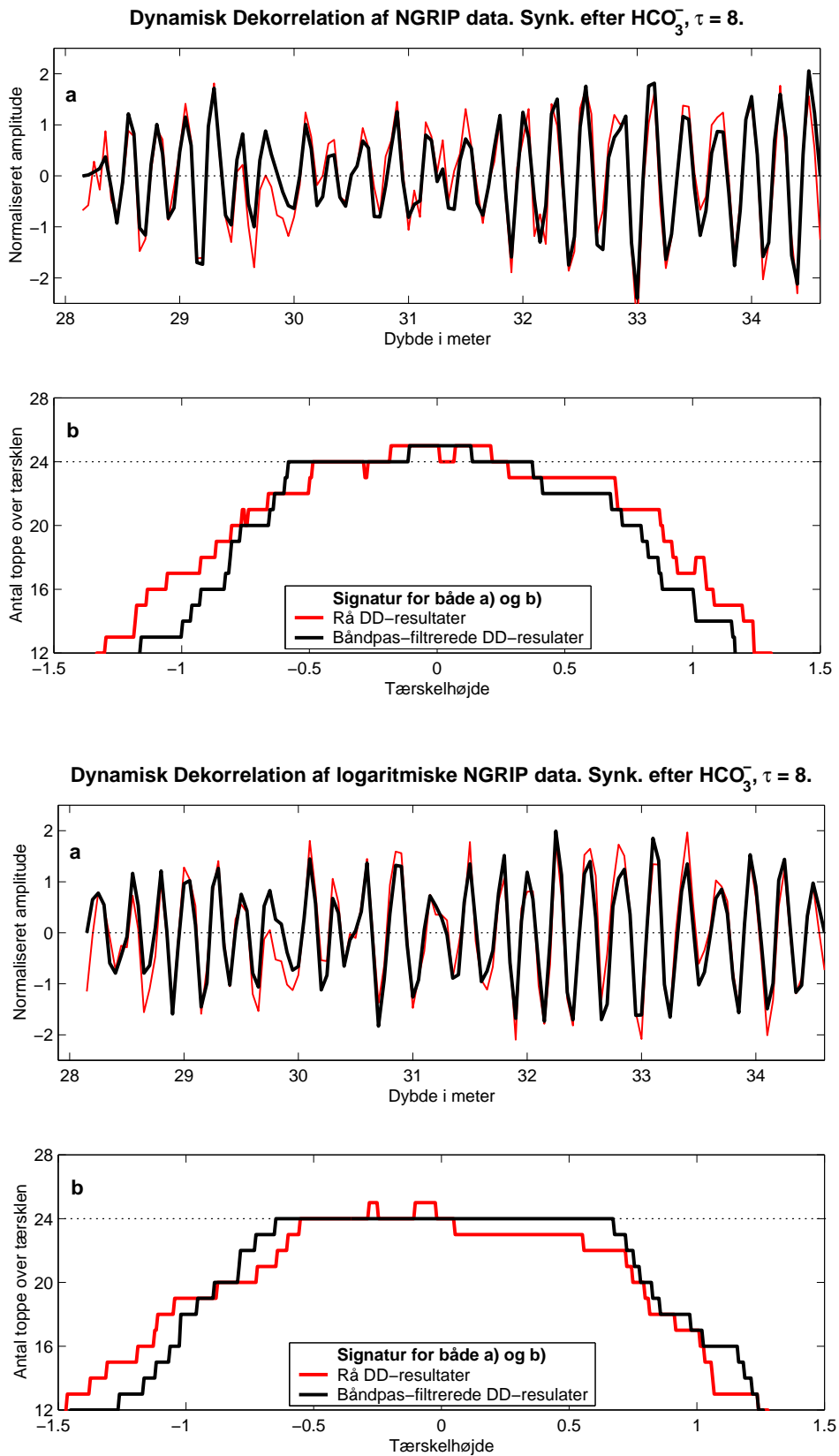
For logaritmisk transformeret data er synkronisering efter  $\text{HCO}_3^-$  kombineret med  $\tau = 8$  det bedste bud. Resultatet heraf ses nederst på figur 5.8. De tilsvarende resultater for almindelige data er vist øverst på grafen. For logaritmisk transformeret data er tælbarheden meget overbevisende med et plateau for  $-0,67 < l < 0,71$ , hvor antallet af toppe er 24 i overensstemmelse med referencetidsskalaen. Bemærk i øvrigt hvordan filtrering uden undtagelse forbedrer tælbarheden, endda meget markant i nogle tilfælde.

Til sammenligning ses i figur 5.9 antallet af toppe i de oprindelige samt logaritmisk transformeret  $\text{Mg}^{++}$  og  $\text{Ca}^{++}$ -dataserier med og uden filtrering.  $\text{Mg}^{++}$ -serien er valgt fordi denne serie har den bedste tælbarhed af de 12 dataserier, mens tælbarheden af  $\text{Ca}^{++}$ -serien er repræsentativ for mange af dataserierne fra NGRIP. Resultaterne ved optælling af toppe i  $\text{Mg}^{++}$ -serien viser, at filtrerede logaritmisk transformeret regulariserede data har en moderat god tælbarhed med en  $(l, p)$ -plateaubredde på ca. 0,75, mens de filtrerede regulariserede data udviser en plateaubredde på ca. 0,60. Antallet af toppe er dog ikke i overensstemmelse med hvad der forventes ifølge referencetidsskalaen, idet de to toppe ved 31,00 – 31,30 meters dybde ifølge referencetidsskalaen sammenlagt kun repræsenterer et årlag. For  $\text{Ca}^{++}$ -serien, der som nævnt er repræsentativ for mange af dataserierne, ses meget dårlig tælbarhed for både almindelige og logaritmisk transformeret data uanset filtrering. De eneste generelle træk man kan se ved at gennemgå  $(l, p)$ -graferne for alle 12 dataserier er, foruden den oftest dårlige tælbarhed, at der generelt tælles for mange toppe og at antallet af toppe er lidt mere veldefineret for de logaritmisk transformeret data, hvilket stemmer overens med resultatet for  $\text{Mg}^{++}$ -serien.

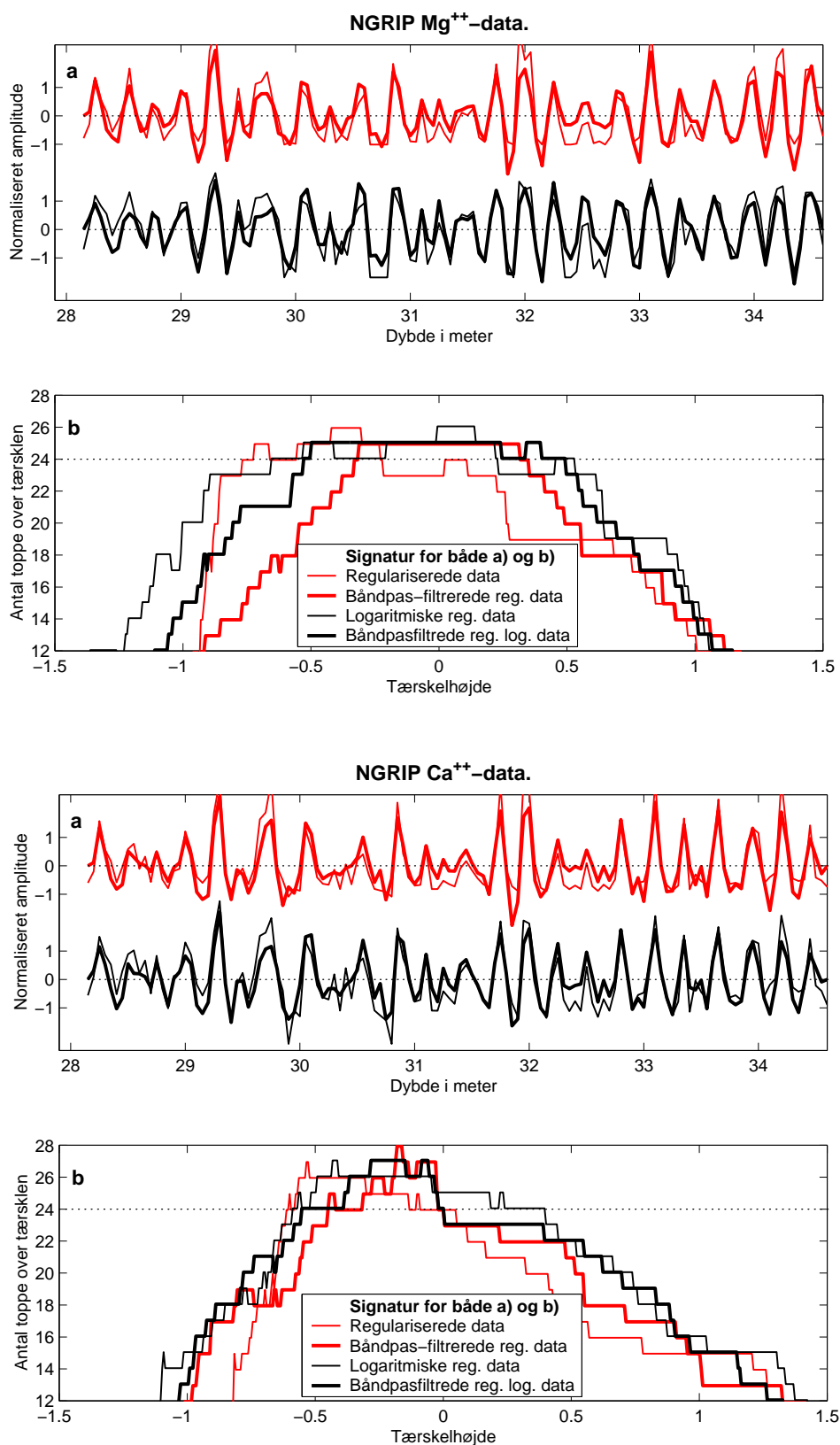
I oversigten i tabel 5.1 og 5.2 over værdier af  $\tau$ , der giver kilde serier med stærk årlig variation ses det at  $\tau = 3, 5, 8, 11, 14$  generelt virker godt, hvilket netop svarer til de første to maksima og de 3 første minima for  $c_S(\tau)$  for det årlige signal med periode  $f = 5,5$ . For det årlige signal vil  $c_S(\tau)$  have minima ved  $\tau = \frac{f}{2} = \frac{5,5}{2} \approx 3$ ,  $\tau = \frac{3f}{2} = \frac{3 \cdot 5,5}{2} \approx 8$  og  $\tau = \frac{5f}{2} = \frac{5 \cdot 5,5}{2} \approx 14$  og maksima ved  $\tau = 5,5$  og  $\tau = 2 \cdot 5,5 = 11$ . Vi kan dermed konstatere, at vores formodninger om kilde seriers egenskaber (jævnfør afsnit 4.2.2) gør det muligt at give et godt a priori bud



Figur 5.7: (a) Resultater fra Dynamisk Dekorrelation af data fra NGRIP, synkroniseret efter  $NH_4^+$  og med  $\tau = 8$ . Øverst med almindelige data, nederst med logaritmisk transformerede data. Filtringen er foretaget som beskrevet i afsnit 5.1.1. (b) Tælbarheden af samme serier.



Figur 5.8: (a) Resultater fra Dynamisk Dekorrelation af data fra NGRIP, synkroniseret efter  $\text{HCO}_3^-$  og med  $\tau = 8$ . Øverst med almindelige data, nederst med logaritmisk transformerede data. Filtringen er foretaget som beskrevet i afsnit 5.1.1. (b) Tælbarheden af samme serier.



Figur 5.9: (a) De oprindelige samt logaritmisk transformerede regulariserede Mg<sup>++</sup>-data (øverst) og Ca<sup>++</sup>-data (nederst) fra NGRIP udsat for filtrering på samme måde som resultaterne fra Dynamisk Dekorrelation i figur 5.8, og 5.7. (b) Tælbarheden af de nævnte dataserier med og uden filtrering.



Figur 5.10: Værdier af kildesignalernes autokorrelationsfunktioner for forskellige værdier af  $\tau$  ved Dynamisk Dekorrelation af data fra NGRIP

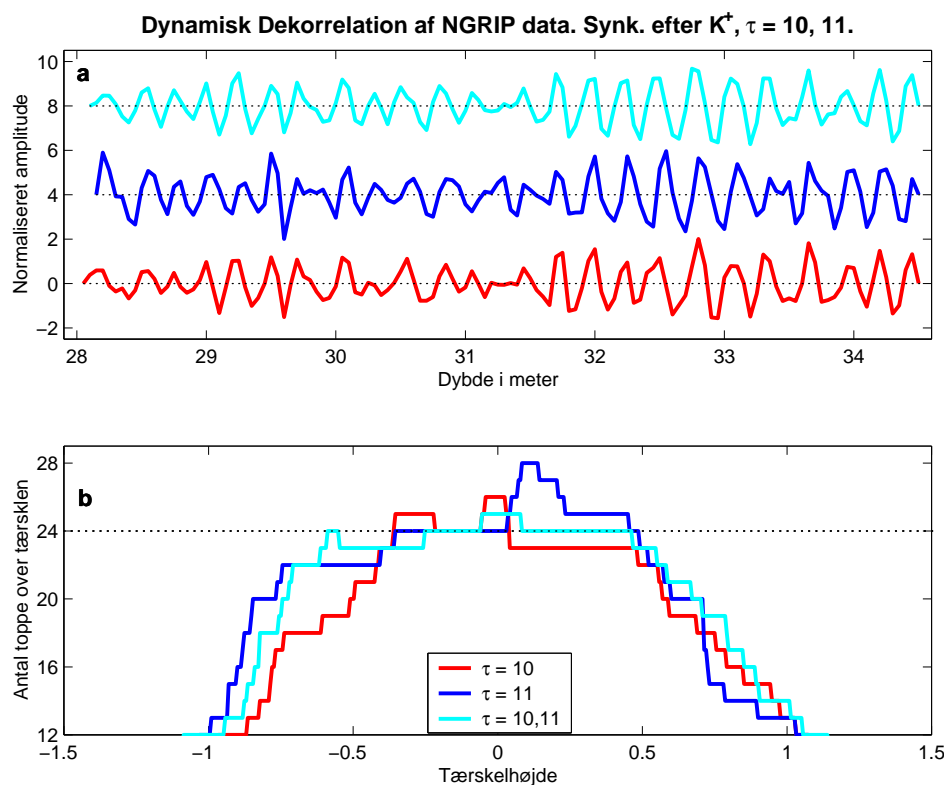
på  $\tau$ . Ved en nøjere sammenligning af kolonnerne i tabel 5.1 og 5.2 ses det at  $\tau = 3$  og  $\tau = 8$  er de bedste valg. En god tommelfingerregel ved Dynamisk Dekorrelation af NGRIP data er derfor at vælge  $\tau$  som  $\frac{f}{2}$  eller  $\frac{3f}{2}$ . Selvom  $f$  principielt ikke kendes før dateringen er foretaget, vil et godt estimat af antallet af prøver pr. år kunne tilvejebringes uafhængigt af resultaterne fra Dynamisk Dekorrelation, hvilket giver mulighed for et godt a priori bud på en velegnet værdi af  $\tau$ .

Figur 5.10 viser  $c_S(\tau)$ -værdierne for  $\tau$  mellem 1 og 15 ved Dynamisk Dekorrelation af NGRIP data. Der observeres en afgørende forskel på  $c_S(\tau)$ -værdierne i forhold til den tilsvarende figur 5.4 for Site D, idet ingen enkelte svingninger skiller sig ud. Istedet ser alle værdierne ud til følge en generelt trend mod numerisk lavere værdier mens de i nogen grad varierer i takt med autokorrelationsfunktionen for det forventede årlige signal. Dette tyder på at der er elementer af det årlige signal i mange af de producerede kildeserier, og at separationen derfor ikke er optimal.

De 2 laveste  $c_S(\tau)$ -værdier er adskilt fra resten af værdierne ved de forventede minima for det årlige signal ved  $\tau = 3$  og  $\tau = 8$ , og en enkelt værdi er pænt separeret fra resten af værdierne for  $\tau = 2, 4, 5, 12$  og  $15$ , men jeg synes ikke at man på baggrund af figuren alene kan slutte at de to minima skyldes en kildeserie med markant årligt signal som det var tilfældet for den tilsvarende figur 5.4 for Site D. Jeg må derfor konkludere at det for NGRIP-data ikke er muligt med rimelig sikkerhed at fastlægge en god værdi af  $\tau$  ud fra en figur som figur 5.10 alene. Da den tilsvarende figur for logaritmisk transformerede data i det store hele er identisk med figur 5.10, er den ikke vist her.

Jeg har afprøvet metoderne til sikring af velseparerede  $c_S(\tau)$ -værdier fra afsnit 4.2.3 på NGRIP-datasættet, men som det var tilfældet for Site D opnås ingen generel sammenhæng mellem god separation og stærk årlig variation i kildeserierne. Jeg må altså konkludere, at god separation af de 12  $c_S(\tau)$ -værdier i sig selv ikke er tilstrækkeligt for at opnå stærk årlig variation i kildeserierne.



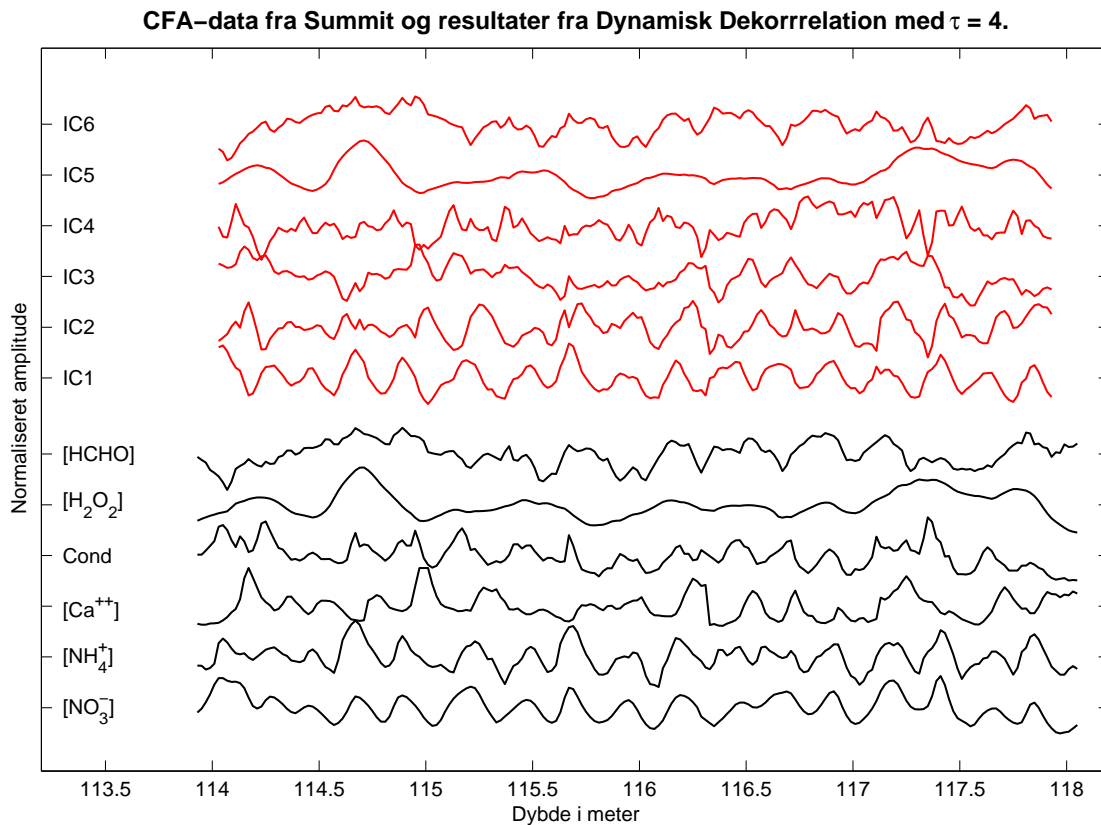


Figur 5.11: (a) Resultater fra Dynamisk Dekorrelation af logaritmisk transformerede data fra NGRIP synkroniseret efter  $K^+$ . Figuren viser et eksempel på anvendelse af to samtidige værdier af  $\tau$ . For overskuelighedens skyld er kun de filtrerede serier illustreret. (b) Tælbarheden af de 3 kildeserier.

Lad mig slutte af med at vise et eksempel på samtidig anvendelse af flere  $\tau$ -værdier, som det er beskrevet på side 34. Under databehandlingen afprøvede jeg en lang række par af  $\tau$ -værdier for at undersøge om resultaterne kunne forbedres af denne vej, men det er ikke lykket mig at opnå bedre resultater end hvad der er opnået med en enkelt  $\tau$ -værdi. Jeg vil dog alligevel vise et eksempel, hvor to værdier af  $\tau$  hver for sig ikke virker særligt godt, mens kombinationen af de to værdier giver udmærkede resultater. Resultaterne fremgår af figur 5.11 og viser resultaterne af Dynamisk Dekorrelation med  $\tau$ -værdierne 10 og 11 samt ved anvendelse af  $\tau_1 = 10, \tau_2 = 11, \alpha_1 = \alpha_2 = 0,5$ . Jeg vil ikke kommentere resultatet yderligere, men blot konstatere at det kan tænkes at det er muligt at forbedre de præsenterede resultater yderligere ved samtidig anvendelse af flere  $\tau$ -værdier.

## 5.4 Resultater for CFA-datasættet fra Summit

Det årlige signal i CFA-datasættet fra Summit er i de ubehandlede dataserier ganske klart repræsenteret i både  $\text{NO}_3^-$  og  $\text{NH}_4^+$ -serierne, og det er derfor begrænset hvad der kan vindes ved Dynamisk Dekorrelation. Dette ses tydeligt på figur 5.12, hvor de 6 ubehandlede dataserier fra 3. delset er vist nederst. Der vil kun blive præsenteret resultater fra det 3. delset, svarende



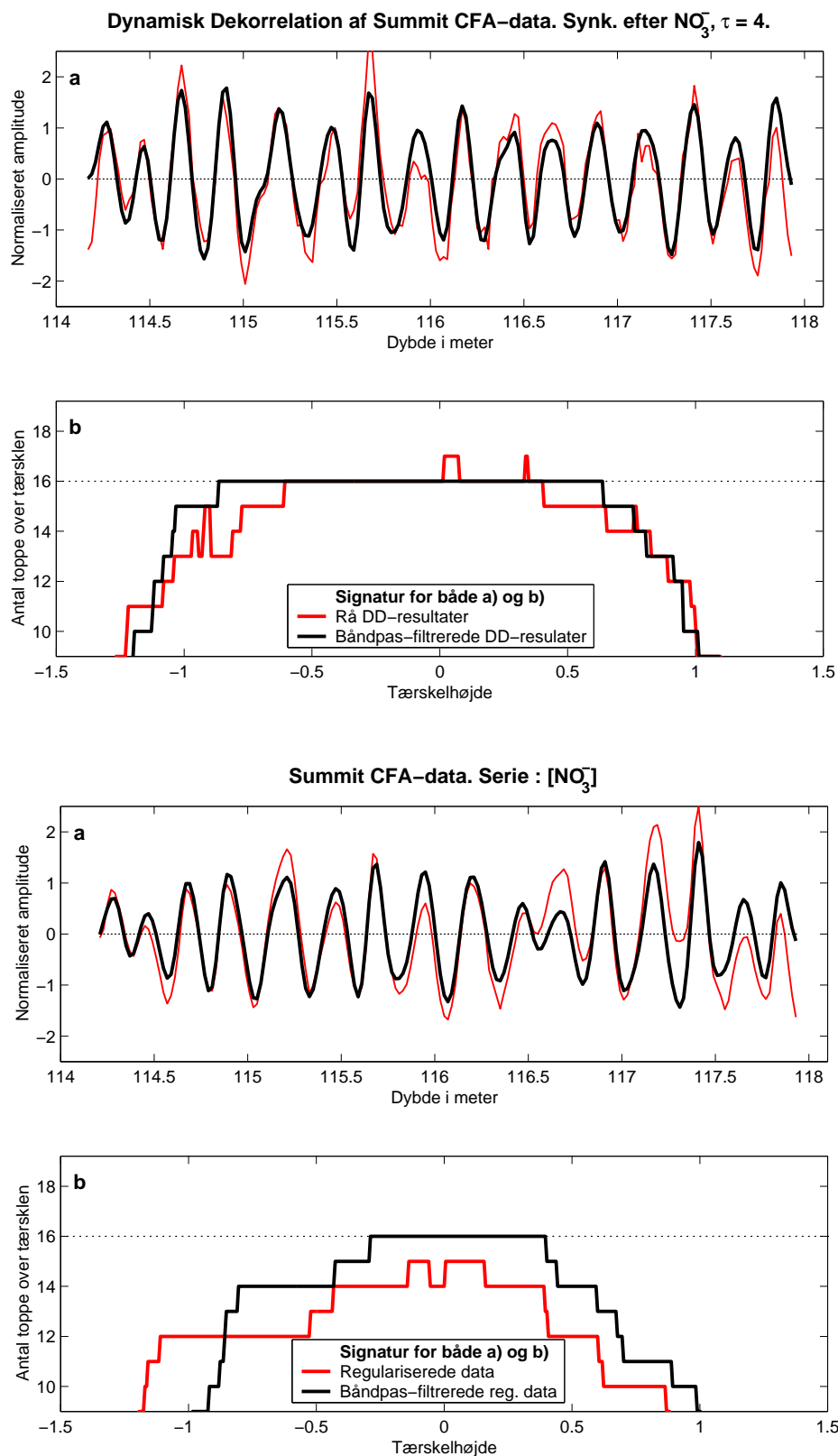
Figur 5.12: Resultater fra Dynamisk Dekorrelation af CFA-data fra Summit (IC1 - IC6), synkroniseret efter  $\text{NO}_3^-$  og med  $\tau = 4$ . Til sammenligning nederst det oprindelige datasæt.

til dybdeintervallet 113,93 – 118,05 meter, idet resultaterne fra analysen af de andre delset er næsten identiske.

Synkroniseringen af det årlige signal i serierne er som beskrevet i afsnit 4.1 foretaget med  $\text{NO}_3^-$  som fastholdt serie, da dette gav de bedste kildeserier. Dette er ikke overraskende, da figur 5.12 klart viser at  $\text{NO}_3^-$ -serien er den serie, der har den klareste årlige variation.

Udvælgelsen af kildeserier med stærk årlig variation er foretaget lidt anderledes for Summit-data end for data fra Site D og NGRIP, idet det årlige signal som nævnt er langt mere tydeligt i såvel dataserier som de producerede kildeserier. Sorteringen i fase 1 og 2 som beskrevet i afsnit 4.2.1 er derfor udeladt, og kvaliteten af det årlige signal vurderes udelukkende på baggrund af en optælling af antallet af toppe som beskrevet i afsnit 5.1. For de  $\tau$ -værdier hvor  $(l, p)$ -grafene har et plateau, er bredden af plateauet angivet nedenfor i tabel 5.3. Hvis de producerede kildeserier ikke gav anledning til et plateau på  $(l, p)$ -grafene, er  $\tau$ -værdien ikke medtaget i tabellen.

Det ses i tabellen, at  $\tau = 4$  giver kildeserien med bedst mulig tælbarhed, men at også  $\tau = 5$  og  $\tau = 6$  giver meget fine resultater. De producerede kildeserier for  $\tau = 4$  er vist i figur 5.12, og resultatet af optælling af toppe i kildeserien med stærk årlig variation (IC1 på figuren) er vist øverst i figur 5.13. Til sammenligning ses nederst i figur 5.13 tælbarheden af toppene i



Figur 5.13: Øverst : (a) Resultater fra Dynamisk Dekorrelation af CFA-data fra Summit, synkroniseret efter  $\text{NO}_3^-$  og med  $\tau = 4$ . Filtringen er foretaget som beskrevet i afsnit 5.1.1. (b) Tælbarhed af kildeserien. Nederst : (a) Den regulariserede  $\text{NO}_3^-$ -dataserie fra Summit udsat for filtrering på samme måde. (b) Tælbarhed for denne dataserie.

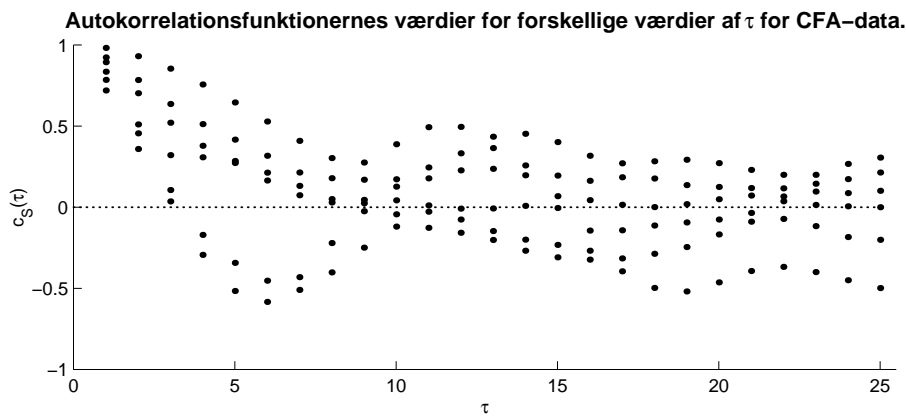
$\tau$	Plateaubredde	$\tau$	Plateaubredde
3	0,93	15	0,72
4	1,50	16	0,97
5	1,30	17	0,83
6	1,22	18	0,74
8	1,16	19	0,66
10	0,92	20	0,64
11	1,09	22	1,03
12	1,04	23	0,88

Tabel 5.3: Bredden af plateauet på  $(l, p)$ -grafene ved optælling af antallet af toppe i de producerede kildeserier ved Dynamisk Dekorrelation af CFA-data fra Summit.

$\text{NO}_3^-$ -serien, der er den serie blandt datasierne, der har den bedste tælbarhed.

For den ubehandlede  $\text{NO}_3^-$ -serie er tælbarheden rimelig god, karakteriseret ved en plateaubredde på 0,69. Ved sammenligning af denne plateaubredde med resultaterne i tabel 5.3 ses det at plateaubredden kan fordobles ved anvendelse af Dynamisk Dekorrelation, og at et stort antal  $\tau$ -værdier giver årlige signaler med forbedret tælbarhed i forhold til datasierne.

Jævnfør overvejelserne i afsnit 4.2.2 burde  $\tau = \frac{f}{2} = \frac{12}{2} = 6$  være et godt bud, og det ses da også af tabel 5.3 at  $\tau = 6$  er den 3. bedste værdi, ligesom  $\tau = f = 12$  giver pæn tælbarhed. Derimod giver  $\tau = \frac{3f}{2} = 18$  ikke kildeserier med væsentligt forbedret årligt signal i forhold til datasierne. Værdierne af  $c_S(\tau)$  for  $\tau$  mellem 1 og 25 ses på figur 5.14. Her ses ligesom for Site D klart adskilte svingende ”kurver”, der her blot er endnu klarere. 2 svingninger følger det forventede årlige signal, der med i snit 12 prøver pr. år har minimum ved  $\tau = 6$  og maksimum for  $\tau = 12$ , mens de 4 andre  $c_S(\tau)$ -værdier ligger vel adskilt herfra. For større værdier af  $\tau$  ses ikke noget klart system i værdiernes variation, bortset fra serien af meget lave  $c_S(\tau)$ -værdier for  $\tau \geq 18$ , der må skyldes en langsom svingning i komponenten IC5 eller IC6 i figur 5.12.



Figur 5.14: Værdier af kildesignalernes autokorrelationsfunktioner for forskellige værdier af  $\tau$  ved Dynamisk Dekorrelation af CFA-data fra Summit.

Heller ikke for CFA-datasættet er der overensstemmelse mellem gode kilde serier og god  $c_S(\tau)$ -separation ifølge kriterierne i afsnit 4.2.3. Faktisk giver  $\tau$ -værdierne 4, 5 og 6 de dårligst separerede  $c_S(\tau)$  værdier ifølge de anvendte kriterier, og at sikre god indbyrdes separation af alle  $c_S(\tau)$ -værdierne kan derfor heller ikke for CFA-datasættet anvendes som kriterium for valg af velegnede  $\tau$ -værdier.



## Del 6

# Konklusion

Det fremgår af resultaterne i del 5, at Dynamisk Dekorrelation er i stand til at producere en kildeSerie med et årligt signal, der er tydeligere end den årlige variation i de anvendte dataserier. Forbedringen skyldes primært to effekter :

1. Mens nogle dataserier indeholder flere eller færre end én top pr. år i gennemsnit, er der i de producerede kildeSerier netop én markant top pr. år. Dette betyder, at man med fordel kan udføre Dynamisk Dekorrelation på datasæt med flere parallelle dataserier og derefter foretage en egentlig datering på baggrund af det producerede årlige signal. En sådan procedure forener derved den sikre identifikation af årlag ved multiparameter-datering med fordelene ved at kunne identificere årlagene ud fra en enkelt dataserie.
2. I det producerede årlige signal er de årlige toppe mere ensartede hvad angår højde og adskillelse. Årlagene er dermed nemmere at optælle, hvilket må formodes især at ville have betydning ved anvendelse af en automatisk optællings-rutine.

For Site D data er begge effekter vigtige. I den producerede kildeSerie er de to 1903-1904 årlag opløst som to selvstændige toppe, mens de i  $\delta^{18}\text{O}$ -signalet optræder som en "top med skulder" (figur 5.1). At årlagene er adskilt i de 5 andre dataserier slår altså igennem i kildeSerien, der ellers i hovedtræk følger  $\delta^{18}\text{O}$ -signalets variation. På samme måde er 1901-årlaget nogenlunde veldefineret i kildeSerien, hvor de enkelte dataserier udviser en, to eller tre små toppe på samme sted. Desuden opnås en generelt mere jævn top-højde, der resulterer i en fordobling af plateaubredden på  $(l, p)$ -grafene fra 0,63 for  $\delta^{18}\text{O}$ -signalet til 1,23 for den producerede kildeSerie (figur 5.2 og 5.3), hvor en lille top dog forstyrrer plateauet for den producerede kildeSerie.

Den første effekt er meget markant for NGRIP-data, hvor det producerede signal udviser tydelige årlige toppe, mens der i dataserierne findes mange toppe, der ikke kan henføres til den årlige svingning. I dette tilfælde vil en multiparameter-datering kræve at man sideløbende anvender en håndfuld af dataserierne til identifikation af årlagene, mens det producerede årlige signal samler det årlige signal fra dataserierne i en grad, så identifikation af årlagene vil kunne

foretages med rimelig sikkerhed udelukkende på baggrund af den producerede kildeserie (figur 5.7 og 5.8). Hvad angår ensartetheden af de årlige toppe, ses det at man kan opnå plateau-bredden på 1-1,25, mod 0,75 for de oprindelige dataserier, der i øvrigt også havde et årlag for meget i forhold til referencetidsskalaen.

For CFA-data er de årlige toppe allerede ganske veldefinerede i de oprindelige dataserier, og effekten af Dynamisk Dekorrelation er derfor primært ensartning af de årlige toppe med fordobling af plateau-bredden som konsekvens.

De producerede serier udviser således et årligt signal, der er så klart, at man selv med den meget simple tællemetode, der anvendes til vurdering af resultaterne, i de fleste tilfælde er i stand til at fastlægge det samme antal år, som der ifølge referencetidsskalaerne burde være i de anvendte datasekvenser. Dette må tages til indtægt for, at antallet af årlag i de producerede serier med god præcision kan optælles med en automatisk metode.

For dataserier, der antager ikke-negative værdier, viser analysen at værdierne med rimelighed kan opfattes som log-normalfordelte, selvom mindre afvigelser er blevet konstateret. Som konsekvens heraf kunne man forvente, at anvendelse af logaritmisk transformerende dataserier ville give et klarere årligt signal, idet den årlige variation vil ske mere symmetrisk omkring middelværdier. Resultater viser dog, at der kun er lidt at vinde ved at anvende logaritmisk transformerende data til Dynamisk Dekorrelation, selvom lidt flere  $\tau$ -værdier (jævnfør tabel 5.1 og 5.2) giver gode resultater ved anvendelse af logaritmisk transformerende data.

Valget af  $\tau$ -værdi har afgørende betydning for, om Dynamisk Dekorrelation producerer en kildeserie med stærk årlig variation. Ved en analyse af  $c_S(\tau)$  for en række mulige kildeserier (afsnit 4.2.2) nåede jeg frem til, at valgene  $\tau = \frac{f}{2}$  og  $\tau = \frac{3f}{2}$ , hvor  $f$  er det gennemsnitlige antal prøver pr. år, måtte forventes at give en god adskillelse af det årlige signal fra de resterende kildeserier. En systematisk sammenligning af de producerede kildeserier for alle  $\tau$ -værdier op til  $2 - 3f$  viste, at  $\tau = \frac{f}{2}$  i praksis var den værdi, der gav det bedste årlige signal for Site D, og at  $\tau = \frac{f}{2}$  og  $\tau = \frac{3f}{2}$  var de to bedste værdier for NGRIP-datasættet. For CFA-datasættet fra Summit var  $\tau = \frac{f}{2}$  det 3. bedste valg, men de tre bedste værdier gav ikke markant forskellige resultater, så også her er  $\tau = \frac{f}{2}$  et glimrende bud. Det er således muligt at give et godt bud på  $\tau$  ud fra kendskabet til det gennemsnitlige antal prøver pr. år.

At vælge  $\tau$  ved maksimering af adskillelsen af alle  $c_S(\tau)$ -værdierne har derimod ikke givet tilfredsstillende resultater. Hvilke  $\tau$ -værdier der bedst adskiller  $c_S(\tau)$ -værdierne afhænger af, hvilket af kriterierne i afsnit 4.2.3 man anvender, men for ingen af de 3 datasæt er der nogen sammenhæng mellem de værdier af  $\tau$  der gav adskilte  $c_S(\tau)$ -værdier, og de værdier af  $\tau$  der i afsnit 4.2.1 blev fundet ved at udvælge de kildeserier, der havde klareste årlig variation. Dette kan skyldes, at de anvendte kriterierne simpelthen ikke er gode nok. Der er dog god overensstemmelse mellem de  $\tau$ -værdier, jeg subjektivt vurderer giver den bedste adskillelse og



de  $\tau$ -værdier der minimerer f.eks.  $\mathcal{M}_2$  som defineret i kriterium 1 på side 55.

Det er således åbenbart af stor betydning at  $\tau$  vælges så  $c_S(\tau)$ -værdien for det årlige signal bliver optimalt adskilt fra de øvrige værdier. Dette er overraskende, idet det fra et teoretisk synspunkt kun er nødvendigt at  $c_S(\tau)$ -værdierne er så forskellige at egenværdiopløsningen af  $\widehat{\mathbf{Q}}$  kan gennemføres med tilstrækkelig god numerisk præcision. Dette tyder på, at der må være nogle faktorer der hindrer god separation, selvom  $c_S(\tau)$ -værdierne er tilstrækkeligt adskilt. Det kan tænkes, at fejlen ved at beregne  $\widehat{\mathbf{C}}_{\bar{X}}(\tau)$  ved cyklisk permuterede data har betydning, ligesom det kan skyldes at  $\widetilde{\mathbf{\Sigma}}$  i ligning 2.23 ikke er diagonal for endelige dataserier. En helt anden mulighed er at dybdeserie/tidsserie problematikken spiller ind, og at problemet er at vi anvender de tidslige egenskaber for signaler, hvis tidslige opløsning ikke i detalje er kendt.

Den anvendte analysemetode indeholder en lang række trin. For hvert trin har der været flere muligheder for at tilpasse metoden. Disse valg er foretaget dels på baggrund af teoretiske overvejelser, dels ved subjektive vurderinger og skøn, men ikke alle kombinationer af valgmulighederne er blevet undersøgt. Da de optimale valg i de enkelte trin er indbyrdes afhængige, er en empirisk test af alle kombinationer af valg meget omfattende, og det kan således ikke udelukkes at der findes en kombination af valg, der giver endnu mere sikker identifikation af årlagene end hvad der er opnået i dette arbejde.

Tilpasningen af metoden til de 3 datasæt har medført mindre forskelle i den optimale procedure for de 3 datasæt. Det er dog min opfattelse at disse forskelle primært kan henføres til at datasættene er meget forskellige. Forskellene skyldes blandt andet at antallene af prøver pr. år er afgørende forskellige, men også at de 3 datasæt indeholder dataserier, der beskriver meget forskellige fysiske og kemiske parametre.

Resultatet af dette arbejde tyder på at Dynamisk Dekorrelation kan anvendes på større datasæt uden yderligere tilpasning af metoden. Som udgangspunkt kan databehandlingen udføres som beskrevet i afsnit 4.3, hvor jeg vil anbefale at der anvendes logaritmisk transformeret data (for de serier, der er naturligt positive), at synkroniseringen foretages efter den serie, der umiddelbart ser ud til at have mest regelmæssig årlig variation, og at  $\tau$  vælges som  $\frac{f}{2}$  eller  $\frac{f}{4}$  eller at der anvendes en lige vægtet kombination af disse to værdier. Ved analyse af lange dataserier viser min analyse, at større datasæt kan analyseres ved opdeling i mindre (eventuelt overlappende) delsæt, der derefter sammenstykket.

Yderligere undersøgelser vil vise om Dynamisk Dekorrelation kan anvendes på andre iskerne-datasæt uden yderligere tilpasning. Viser dette sig at være tilfældet, kan Dynamisk Dekorrelation blive et værdifuldt bidrag til udviklingen af metoder til objektiv multiparameter-datering.



# Appendix A

## Definition af $\delta$ -notation

Isotopindholdet i en prøve angives i geofysikken oftest ved at angive den relative forskel i isotopforhold i forhold til en kendt standardprøve, den såkaldte  $\delta$ -værdi. Det var Harmon Craig, der i 1961 indførte delta-notationen og påpegede behovet for en global standardprøve som reference [Craig, 1961]. At anvende afvigelser fra en standard er motiveret af, at isotopforholdenes afvigelser fra standardværdien normalt er små. Desuden er det rent måleteknisk nemmere præcist at bestemme den relative afvigelse mellem to isotopforhold, end at bestemme absolutte forhold. Herunder vil jeg definere  $\delta^{18}\text{O}$ , men  $\delta$ -værdier defineres tilsvarende for andre isotoper.

I naturligt forekommende ilt er isotopforholdet  $[^{16}\text{O}] : [^{17}\text{O}] : [^{18}\text{O}]$  omtrent 99762 : 38 : 200. Når man er interesseret i fraktionering af iltisotoper er det derfor naturligt at vælge at måle forholdet mellem  $^{18}\text{O}$  og  $^{16}\text{O}$  fremfor det tilsvarende forhold mellem  $^{17}\text{O}$  og  $^{16}\text{O}$ , idet både den større forekomst og den større masseforskel gør måling af  $^{18}\text{O}$ -fraktionering nemmere end den tilsvarende måling af  $^{17}\text{O}$ -fraktionering. Forholdet mellem koncentrationerne af den tunge og den lette ilt-isotop betegnes  $R$ :

$$R = \frac{[^{18}\text{O}]}{[^{16}\text{O}]} \quad (\text{A.1})$$

Prøvens  $\delta^{18}\text{O}$ -værdi defineres ved sammenligning af  $R$  for prøven med  $R$  for en standardprøve:

$$\delta^{18}\text{O} = \frac{R_{\text{prøve}} - R_{\text{standard}}}{R_{\text{standard}}} \cdot 1000 \text{‰} \quad (\text{A.2})$$

Værdien er her angivet i promille, hvilket er standard.

Som standardprøve vil man ofte for ilt-isotoper vælge Standard Mean Ocean Water (oftest blot kaldet SMOW) eller en mere moderne variant heraf, Vienna-SMOW (VSMOW).

Nogle steder i litteraturen defineres isotopforholdet  $R$  som forholdet mellem koncentrationerne af den tunge isotop og den samlede koncentration af alle forekommende isotoper af det pågældende grundstof. På grund af at  $^{16}\text{O}$  udgør en så stor del af naturligt forekommende ilt er den resulterende forskellen mellem de to definitioner dog lille for iltets vedkommende.

I en del ældre skrifter (f.eks. Willi Dansgaards afhandling [*Dansgaard, 1961*]) anvendes den absolutte forskel mellem  $R_{prøve}$  og  $R_{standard}$  istedet for delta-værdier. En redegørelse for hvordan der omregnes mellem absolutte afvigelser og relative afvigelser findes i [*Dansgaard, 1961*], side 39.

## Appendix B

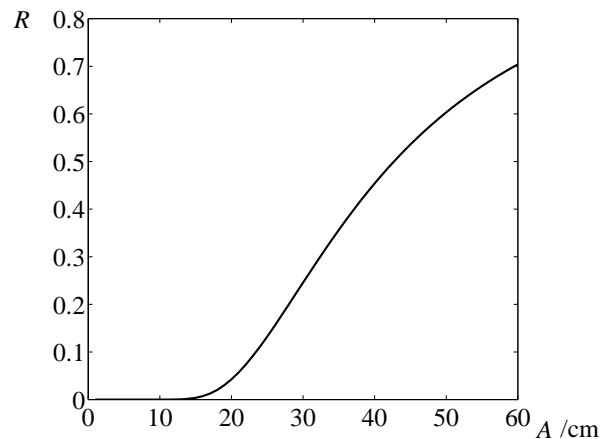
# Diffusion og tilbagediffusion

Den årlige svingning i  $\delta^{18}\text{O}$ -signalet har en amplitude på mindst 5‰ i nyfalden sne, men den nyfaldne sne undergår væsentlige forandringer, mens den bevæger sig ned gennem iskapens øverste hundrede meter. Sneens forvandling til is kaldes firnifikation, og kan modelleres med godt resultat [*Herron & Langway, Jr.*, 1980, *Anderson & Benson*, 1963], men skal ikke i detalje behandles her, hvor en overordnet beskrivelse er tilstrækkelig. Fra en densitet i omegnen af  $300 \text{ kg/m}^3$  pakkes snekrystallerne gradvist tættere sammen, efterhånden som den overliggende vægt sammenpresser sneen. Når densiteten når  $550 \text{ kg/m}^3$  er firnen ikke længere permeabel for luft og damp. Fra dette punkt, der kaldes pore-lukningsdybden, foregår den yderligere sammenpresning ved at den indelukkede luft komprimeres. Sammenpresningen fortsætter indtil densiteten når  $917 \text{ kg/m}^3$ , hvorefter isen er praktisk talt inkompressibel. I den første fase foregår densifikationen både ved at de enkelte krystaller pakkes tættere, men også ved rekrySTALLISERING, hvor vandmolekyler fordampes og kondenseres igen på andre krystaller. Herved flyttes de enkelt vandmolekyler en karakteristisk afstand  $L_{firn}$ , der er en samlet diffusionslængde for diffusionen via dampfasen i firnen indtil densiteten når  $550 \text{ kg/m}^3$ . Diffusionslængden afhænger af en lang række parametre, hvoraf skal nævnes temperaturen og den såkaldte *tortuosity*-faktor, der populært sagt angiver ”snørkelheden” af de mellemrum mellem krystallerne, vanddampen skal følge. I [*Johnsen*, 1977] beregnes  $L_{firn}$  til ca. 8 cm for de temperaturforhold, der hersker på indlandsisen. Amplituden af en svingning med bølgelængde  $A$  dæmpes ved denne diffusionsproces med en faktor  $R$  der er givet ved

$$R = \exp\left(\frac{-2\pi L_{firn}^2}{A^2}\right) \quad (\text{B.1})$$

Figur B.1 viser  $R$  som funktion af  $A$  for  $L_{firn} = 8 \text{ cm}$ .

For at kunne tælle de årlige svingninger må vi kræve at amplituden af svingningerne er mindst et par gange højere end den usikkerhed, der er forbundet med måling af isotopforholdet. Ved den målemetode der anvendes på Geofysisk Afdeling i København, er usikkerheden omkring 0,05‰, hvilket giver en ønsket minimumsamplitude på 0,15 - 0,25‰. Da amplituden af årstidssvingningen som sagt er omkring 5‰ før dæmpning, må vi kræve at 3 - 5% af den

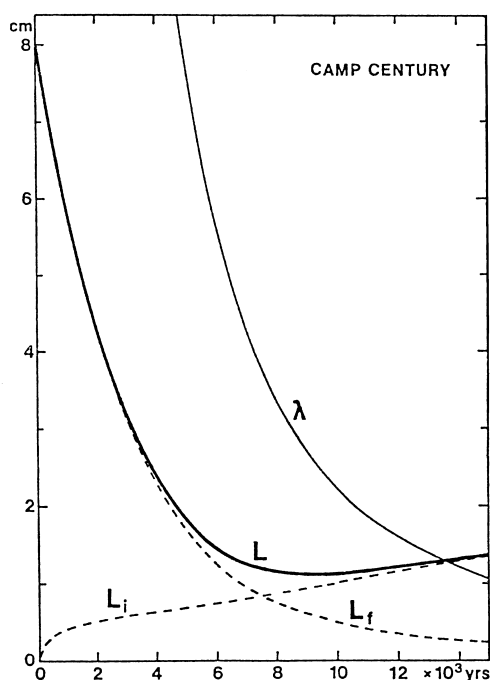


Figur B.1: Dæmpningsfaktor  $R$  som funktion af bølgelængde  $A$  ved diffusionslængde  $L_{firn} = 8$  cm. Bemærk at bølgelængder under 20 cm næsten diffunderer helt væk.

oprindelige amplitude er bevaret efter firnifikation. Som det ses på figur B.1 er dette opfyldt for svingninger med bølgelængder over 20 cm. Vi må derfor forvente at årstidssvingningen er velbevaret i Dye 3-iskernen, hvor den årlige akkumulation er over en halv meter isækvivalent, mens akkumulationen på Summit, hvor GRIP og GISP2-iskernerne blev boret, er på ca. 25 cm og dermed i den nedre ende af det tilstrækkelige. Ved NGRIP er akkumulationen 19,5 cm is om året [Dahl-Jensen *et al.*, 2001], hvilket betyder at vi kun kan forvente at årstidssvingningens amplitude efter firnifikation er en anelse større end måleusikkerheden.

Også længere nede i iskappen end pore-lukningsdybden foregår der diffusion, men her er diffusionshastigheden størrelsesordener mindre. Diffusionslængden  $L_{firn}$  er som sagt ca. 8 cm efter firnifikationen er overstået, men idet årlagene bliver udtyndet med dybden på grund af isens flydning, bliver den effektive firn-diffusionslængde mindre med dybden, mens diffusionslængden som følge af diffusion i is vokser. For at kunne beregne den samlede diffusionslængde i de nedre dele af iskappen er det derfor af essentiel betydning at have bestemt udtyndingsraten som funktion af dybden ved hjælp af en realistisk flydemodel. På figur B.2 ses de to bidrag til den samlede diffusionslængde for Camp Century-iskernen, hvor en DJ-model er anvendt til at bestemme udtyndingsraten.

Når den samlede diffusionslængde (med udtyndingseffekten) er kendt, er det muligt at genskabe de "tabte" svingninger ved hjælp af en teknik, der kaldes tilbagediffusion [Johnsen, 1977, Johnsen *et al.*, 1999, Johnsen *et al.*, 2000]. Diffusionens effekt på et signal svarer til at signalet foldes med et gaussisk filter :  $F_1 = \frac{1}{L\sqrt{2\pi}} \exp\left(\frac{-(z-z')^2}{2L^2}\right)$ , hvilket i frekvens-domænet svarer til responset givet ved ligning B.1. De bortdiffunderede svingninger kan genskabes ved at man konstruerer et filter, der forstærker de enkelte frekvensbånd i signalet for at kompensere for dæmpningen (jævnfør ligning B.1). Her må man vælge den maksimale frekvens, man mener at kunne rekonstruere, idet man ellers også forstærker den højfrekvente støj i signalet voldsomt. På denne måde rekonstrueres dynamikken i frekvensområdet, der ligger under en valgt afskæringsfrekvens, mens alt over denne frekvens lades tilbage uforstærket eller eventuelt helt



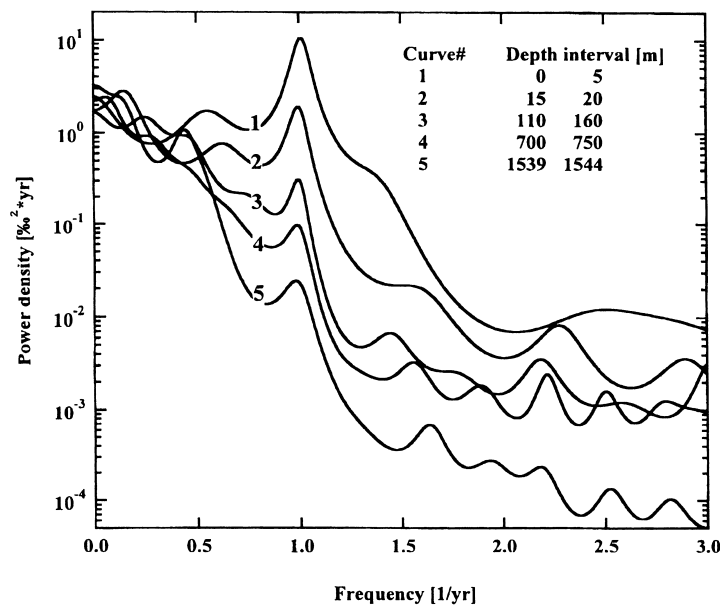
Figur B.2: Diffusionslængden  $L$  som funktion af dybden for Camp Century-iskernen.  $L$  består af 2 bidrag:  $L_f$  fra diffusion i firnen og  $L_i$  fra diffusion i isen. Kurven  $\lambda$  angiver den modellerede årlagstykkelse i den anvendte DJ-model. Fra [Hammer *et al.*, 1978]

fjernes. På figur B.3 ses spektraltætheder<sup>1</sup> for  $\delta^{18}\text{O}$ -signalet fra 5 segmenter fra GRIP-iskernens Holocæn-del. Det årlige signal ses som en tydelig top, der skiller sig klart ud fra den næsten flade støj-niveau, der starter ved frekvenser på  $1,5 - 2 \text{ år}^{-1}$ . Så længe det årlige signal klart kan adskilles fra støjen, som det er tilfældet på figur B.3, kan tilbagediffusion foretages uden risiko for at det rekonstruerede signal er en matematisk konstruktion uden fysisk indhold.

På figur B.4 ses en sekvens fra ca. 60 meters dybde i Crête-iskernen, hvor  $\delta^{18}\text{O}$ -signalet (C) er blevet tilbagediffunderet (B). Det ses, at dynamikken i det rekonstruerede signal minder meget om det signal, der ses helt i toppen af iskernen, hvor diffusion ikke har dæmpet de høje svingninger endnu.

Valget af afskæringsfrekvens er kritisk, ligesom bestemmelse af diffusionslængden som sagt kræver viden om (eller modellering af) udtyndingsraten som funktion af dybden. I [Johnsen, 1977] anvendes tilbagediffusions-teknikken på et stykke af Camp Century-iskernen fra 1156 meters dybde hvor  $L$  beregnes til 1,8 cm. Resultatet er, at det ca. 22 cm lange kerne-stykke viser sig at indeholde 6 bortdiffunderede årlag, men i et korrektiv til artiklen korrigeres  $L$  til 1,2 cm, hvilket betyder at der rekonstrueres 10 istedet for 6 årlag. Eksemplet viser klart, at svingninger med perioder under 2-3 gange diffusionslængden ikke med sikkerhed kan siges at stamme fra det originale signal, men lige så godt kan være en utilsigtet "bivirkning" af rekonstruktionen. Dette faktum har nok været medvirkende til, at tilbagediffusions-teknikken

<sup>1</sup>De viste spektraldensiteter er estimater beregnet ved hjælp af den såkaldte Maximum-entropi metode (MEM), der er velegnet til især korte dataserier. Se [Johnsen & Andersen, 1978] for en diskussion af metoden.

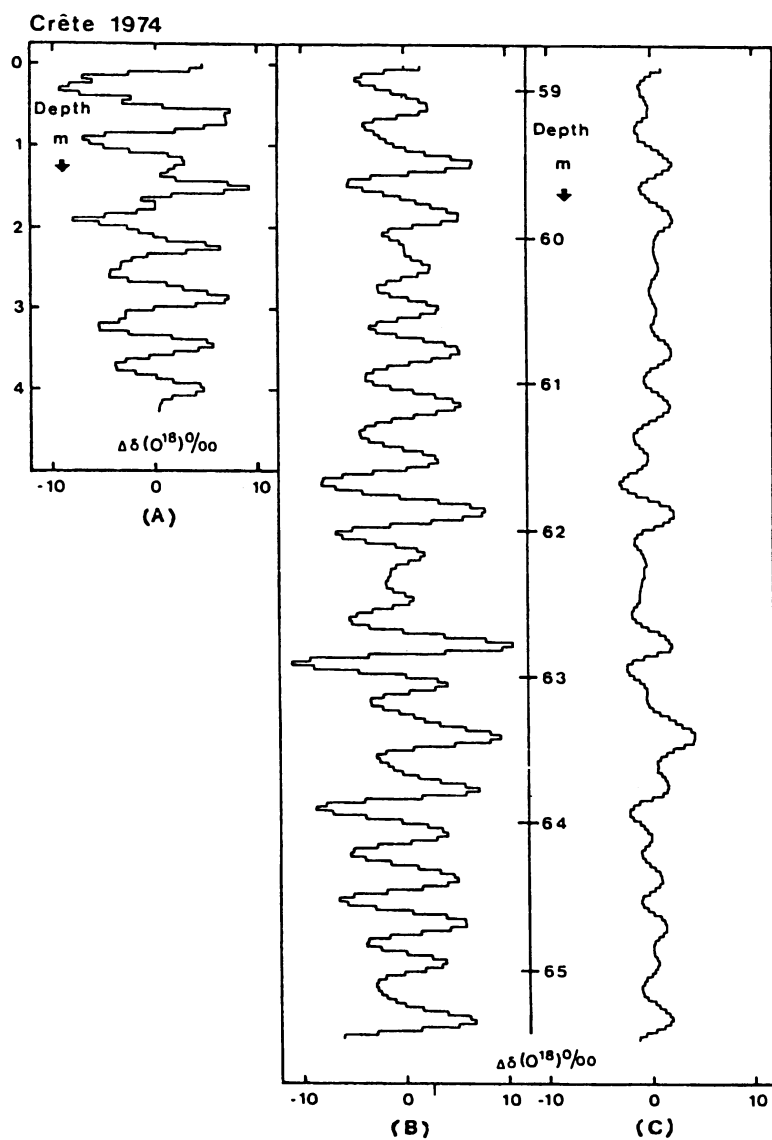


Figur B.3: MEM Spektaltætheder (se fodnoten side 87) for 5 5-meters intervaller af GRIP-iskernen. Det årlige signal er tydeligt, men forskellen på kurverne 3-5 viser, at der sker en kraftig svækkelse af signalet selv efter firnifikationens ophør. Denne svækkelse er meget kraftigere end hvad man måtte forvente som resultat af almindelig diffusion i is. [Johnsen *et al.*, 2000] foreslår derfor at diffusionen også sker gennem vandfasen, der på trods af temperaturen kan eksistere langs krystalgrænser [Rempel *et al.*, 2001]. Dette meget væsentlige bidrag til den samlede diffusion kaldes *excess diffusion*. Figur fra [Johnsen *et al.*, 2000].

ikke har vundet udbredelse i vide kredse. Det er denne usikkerhed - sammen med det faktum at jeg primært bruger  $\delta^{18}\text{O}$ -data fra små dybder, hvor diffusionen endnu ikke har dæmpet amplituden helt så meget - der har gjort, at jeg har valgt ikke at anvende tilbagediffunderede  $\delta^{18}\text{O}$ -data i denne opgave. Der er dog ingen tvivl om at tilbagediffusion er nødvendig, hvis  $\delta^{18}\text{O}$ -data skal kunne bruges til stratigrafisk datering af is fra alt andet end de øverste dele NGRIP-iskernen, selvom årlagstykkelsen her er lige på grænsen til at rekonstruktionen kan foretages.

Til sidst vil jeg nævne en anvendelse af teorien bag tilbagediffusion, der kan blive vigtig ved behandling af data fra Continuous Flow Analysis (CFA). CFA er en målemetode hvor et stykke iskerne smeltes og koncentrationen af udvalgte ioner samt evt. andre parametre måles i smeltevand. Smeltevand pumpes fra et smeltehoved gennem en række tynde slanger før selve målingen foretages, og i mellemtiden sker der en vis opblanding i væskekolonnen. Konsekvensen heraf er, at en pludselig ændring i koncentrationen af en bestemt ion vil blive målt som en noget langsommere gradvis ændring, hvilket matematisk svarer til foldning med et ikke-symmetrisk filter. Det er håbet, at dette filters karakteristika kan bestemmes ved måling af opstillingens respons på en pludselig koncentrationsændring, og at man derefter kan korrigere for opblandingen ved at folde det målte signal med et "omvendt" filter, der genskaber de oprindelige koncentrationsændringer.





Figur B.4:  $\delta^{18}\text{O}$ -profil fra de øverste 5 meter af Crête-iskernen (A), og fra 59 til 65,5m (C). (B) viser sidstnævnte stykke efter tilbagediffusion. Fra [Johnsen, 1977].



## Appendix C

# Artikel i Annals of Glaciology

I samarbejde med Katrine Krogh Andersen udarbejdede jeg i foråret 2001 en artikel til Annals of Glaciology om resultaterne af dynamisk dekorrelation af iskernedata. Marie-Louise Siggaard-Andersen var ansvarlig for målingen af ionkromatografidata fra NGRIP og Henrik B. Clausen stod for dateringen af de anvendte dataserier. Artiklen følger nedenfor i den endelige version, som er blevet godkendt til publikation i Annals of Glaciology, vol. 35.

Der er et par mindre forskelle i resultaterne i forhold til de resultater, jeg har præsenteret i opgaven :

- Der er lidt forskel på de anvendte datasæt, idet ECM-data i artiklen ikke anvendes for Site D og  $\text{HCO}_3^-$  ikke er medtaget i NGRIP-datasættet.
- Regulariseringen af data er foretaget på en lidt anden måde. I datasættene, der blev anvendt i artiklen er unormalt høje toppe manuelt identificeret og beskåret til et niveau, der svarer til højden af "almindelige" årlige toppe. Jeg besluttede mig sidenhen for at anvende regulariseringsmetoden, der er beskrevet i kapitel 3.3, idet den identificerer den sporadiske komponent ud fra objektive kriterier og dermed giver reproducerbare resultater.
- Kombinationen af  $\text{Cl}^-$  som fikseret serie og  $\tau = 2$  angives som blandt de bedste kombinationer. Jeg havde ikke, da jeg udvalgte resultater til artiklen, foretaget en så grundig sammenligning af resultaterne som den sammenligning der beskrives i afsnit 5.3. Den anvendte kombination er ikke blandt de allerbedste, selvom resultaterne er rimelig gode.
- Filtreringen foretages i artiklen med et lavpasfilter og ikke med et båndpasfilter som beskrevet i afsnit 5.1.1. Forskellen er næsten uden betydning for de i artiklen præsenterede serier, men kan være vigtig for kildeserier, der foruden det årlige signal indeholder en trend eller svingninger med perioder på flere år.

# Extracting the annual signal from Greenland ice core chemistry and isotopic records

Sune Olander Rasmussen,<sup>1</sup>  
Katrine Krogh Andersen,  
Marie-Louise Siggaard-Andersen,  
and Henrik B. Clausen

Department of Geophysics, NBIfAFG, University of Copenhagen,  
Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

Stratigraphic dating of ice cores by identification and counting of annual cycles in e.g. chemical measurements requires skill and experience. The work presented here investigates a method of data enhancement which is a first step towards an automated and more objective method of annual layer counting. The method of dynamical decorrelation is briefly introduced and is applied to data from Site D and NorthGRIP in central Greenland. With this method the measured data series are decomposed into a number of independent source series, one of which exhibits a more pronounced annual variation than the input data themselves. The annual variation is more regular in that (1) some double and triple peaks in the measured series are replaced by single peaks in the extracted signal and (2) the resulting annual peaks have a much more uniform height. A simple method of determining the number of annual peaks in a series is set up. Using this method, it is shown that it is easier to determine the number of annual peaks in the series produced by dynamical decorrelation, than in the original data series. Dynamical decorrelation may thus be used to improve data series prior to dating.

## Introduction

In this paper we investigate the possibility of extracting a common annual signal from combined chemistry and isotopic records from central Greenland ice cores. Records from Site D (Steffensen, 1988) and NorthGRIP (Dahl-Jensen and others, 2001) are analysed. The objective is to extract a signal that exhibits a more pronounced annual variation than what can be found directly in the measured records. The extracted signal can be used as the basis for e.g. automated counting of annual layers in ice core records.

The method introduced here is based upon the assumption that the measured records can be considered to be linear combinations of a set of underlying source signals. The method allows us to determine the source signals from the measured records if their autocorrelation functions are different. The extracted source signals may reflect variations in physical parameters such as temperature, but they may also be purely mathematical constructions. From an ice core dating point of view, even the latter could be useful if the obtained signal shows pronounced annual variation.

---

<sup>1</sup>Corresponding author, e-mail [olander@gfy.ku.dk](mailto:olander@gfy.ku.dk).

## Treating ice core records as time series

Most ice core records are sampled at constant depth intervals. If precipitation is unevenly distributed in time, ice core records are equivalent to time series sampled at varying (and unknown) time intervals. Firnification processes, thinning of annual layers due to ice flow, and changes in the amount of precipitation during different climatic periods will lead to further changes in the sampling rate with depth.

In this work, we use data series each spanning only around 20 years, corresponding to less than 10 meters. Within each period, we may assume the overall climatic situation to be unchanged, and thus that the annual precipitation does not undergo large changes. Firnification processes and thinning of annual layers may also be disregarded when using short series at these depths. It will be assumed that the precipitation is relatively evenly distributed throughout the year, which is supported by investigations by e.g. Shuman and others (1995). Adding up these assumptions, we will consider the measured records to be sampled at approximately constant time intervals, thus treating the "depth series" as ordinary time series.

When analysing longer data series, one should account for firnification in the upper layers and the thinning of annual layers due to ice flow. Also, when analysing data from periods with changing precipitation patterns and rates special care should be taken.

## Dynamical Decorrelation

We assume that the measured data series (the rows of a  $N \times T$  matrix  $\mathbf{X}$ ) are formed by a linear combination of an unknown set of underlying source signals  $\mathbf{S}$  (also  $N \times T$ ) as  $\mathbf{X} = \mathbf{A}\mathbf{S}$ , where the  $N \times N$  matrix  $\mathbf{A}$  is called the mixing matrix. A common approach is to assume that the source signals form an orthogonal set and to arrange the series in the set with decreasing variance. This approach is known as Principal Component Analysis (PCA) and is widely used in time series analysis.

However, the assumption of orthogonality of the source signals is not necessarily viable, and some other assumption about the source signals must be chosen. Methods that assume statistically independent source signals are known as Independent Component Analysis (ICA), but the assumption of independence of the source signals does not alone contain enough information to determine  $\mathbf{S}$  uniquely. Many choices of additional assumptions can be made (for a short summary, see e.g. Hyvärinen (1999)), but here the method of Molgedey and Schuster (1994), further developed by Hansen and others (2000), will be applied. Only a brief overview will be given here, and only the formalistic simple version of infinitely long series is considered. When performing the analysis on real finite data series, a number of mathematical technicalities must be considered, but the basic idea remains the same.

When performing PCA and most versions of ICA, no temporal information from the data is used; the data could be sampled at random times or be without temporal structure at all. In our case with ice core records, the source signals are expected to be e.g. the annual signal, other (quasi-)periodic signals and noise. These signals have very different periods and thereby autocorrelation functions, and the idea of Molgedey and Schuster is to make use of this fact. Given a linear mixture of the source signals, the method is able to separate the source signals by only assuming that one can choose a time lag  $\tau$ , for which the values of the autocorrelation functions of the different source signals are clearly distinct.

## Methodology

The correlation function between two infinitely long normalized discrete signals  $x_{i,t}$  and  $x_{j,t}$ ,  $t \in \mathbb{Z}$ , can be defined as  $c_{ij}(\tau) = \lim_{T \rightarrow \infty} \left[ \frac{1}{2T+1} \sum_{t=-T}^T x_{i,t} x_{j,t+\tau} \right]$ . The correlation functions are all collected in the correlation matrix  $\mathbf{C}_X(\tau)_{ij} = c_{ij}(\tau)$ , which for  $\tau = 0$  is the crosscorrelation matrix. We now define the quotient matrix  $\mathbf{Q} = \mathbf{C}_X(\tau) \mathbf{C}_X(0)^{-1}$ . Using the basic assumption that  $\mathbf{X} = \mathbf{A}\mathbf{S}$ , we get that  $\mathbf{C}_X(\tau) = \mathbf{X}\mathbf{X}^\top = \mathbf{A}\mathbf{C}_S(\tau)\mathbf{A}^\top$ , where  $\mathbf{C}_S$  is the correlation matrix for the source signals, and conclude that  $\mathbf{Q} = \mathbf{A}\mathbf{C}_S(\tau)\mathbf{C}_S(0)\mathbf{A}^{-1}$ . The fact that the source signals were assumed to be independent results in  $\mathbf{C}_S(\tau)$  being diagonal, such that the quotient matrix is given as

$$\mathbf{Q} = \mathbf{A} \begin{bmatrix} \frac{c_{11}(\tau)}{c_{11}(0)} & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & 0 & \frac{c_{NN}(\tau)}{c_{NN}(0)} \end{bmatrix} \mathbf{A}^{-1} \quad (\text{C.1})$$

This closely resembles an eigenvalue decomposition of  $\mathbf{Q}$ , and thus  $\mathbf{A}$  (and thereby  $\mathbf{S}$ ) can be computed by diagonalizing  $\mathbf{Q}$ , which itself can be readily calculated from the measured data.  $\mathbf{A}$  is uniquely determined as long as the diagonal in the above equation consists of significantly different values. This means that if a choice of  $\tau$  exists where the autocorrelation functions of the source signals are significantly different, the source signals can be uniquely determined. There is, however, no way of choosing  $\tau$  a priori, and one has to find a suitable value by trial-and-error. A number of technical and numerical details must be taken into consideration. For a thorough discussion of these, and a proof of the fact that  $\mathbf{Q}$  always has a real-valued eigenvalue decomposition, see Hansen and others (2000).

## Performing Dynamical Decorrelation

A suitable set of data series from the same depth interval must be available for the analysis. Only series that are anticipated to contain the annual signal should be included. For evaluation purposes, we have chosen two different data sets which have been dated by annual layer counting. This dating has been verified with the help of known historical markers (Hammer and others, 1978).

For Site D the concentrations of  $\text{Cl}^-$ ,  $\text{NO}_3^-$  and  $\text{SO}_4^{2-}$  and dust mass have been measured at 5 cm resolution, whereas  $\delta^{18}\text{O}$  samples were cut in 8 samples/year according to a cutting scale based on an ice flow model (Steffensen, 1988). For this application the latter thus had to be resampled in order to be used together with the chemistry data. The resampling was done using statistical cubic interpolation. The sequence covers the years 1890-1910 in the depth interval 42.2 - 51.15 m. The relatively high annual accumulation rate of  $0.37 \text{ m yr}^{-1}$  (Clausen and others, 1988) at site D combined with the fact that the selected series are taken from a shallow depth give a sampling frequency of approx. 9 samples per year.

Also new data from NorthGRIP have been analysed. The NorthGRIP(1) ice core deep drilling started during the field season of 1996 9.85 m below surface and terminated at a depth of 1371 m during the 1997 field season (Dahl-Jensen and others, 2001). From the top down to a depth of 349.8 m a continuous series of 5 cm resolution was sampled for ion chromatographic analysis. This series corresponds to a time interval of approximately 2000 years. The chemistry samples were manually decontaminated and cut in the field using a microtome knife. They were stored frozen in Coulter accuvettes until measurement. Samples representing core depths from 116.05 to 221.65 m were measured by M. Hansson at the University of Stockholm, Sweden. The rest of the series was measured at the Dept. of Geophysics, University of Copenhagen.

The samples measured in Copenhagen were melted in the Coulter accuvettes and under a clean bench they were decanted into Dionex 5 ml polyvials with filter caps. The polyvials and filter caps were precleaned in  $> 18 \text{ M}\Omega$  water from an Elgastat UHP from Elga Ltd.

The measurements were done using a Dionex 500 micro bore ion chromatograph with two channels. A single standard solution was used for quantification of the contents of  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{NH}_4^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{F}^-$ ,  $\text{MSA}^-$ ,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ , and  $\text{SO}_4^{2-}$  in the samples. In the standard solution the concentrations of  $\text{Li}^+$  and  $\text{MSA}^-$  were  $0.2$  and  $0.3 \mu\text{equiv.kg}^{-1}$ . The concentrations of the other ions in the standard solution were ranging from  $1$  to  $7 \mu\text{equiv.kg}^{-1}$ . The chromatograms were integrated using Peaknet 4.30 software and the peak heights were used in the evaluation except when the peak area seemed to be more reproducible.

A sequence from the depth interval 28.1-34.65 m (corresponding to 24 years) was selected because of excellent data quality and because the firnification and thinning of layers have not proceeded far at this depth. The measured concentrations of 11 ions were used.  $\delta^{18}\text{O}$  was left out of the analysis (1) because the signal is lost at greater depths due to diffusion in the relatively thin annual layers and (2) in order to make the data set significantly different from that of Site D.

With a sample size of 5 cm and with an annual accumulation rate of  $0.195 \text{ m yr}^{-1}$  (Dahl-Jensen and others, 2001), the sampling frequency is approx. 5.5 samples per year.

## Data preprocessing

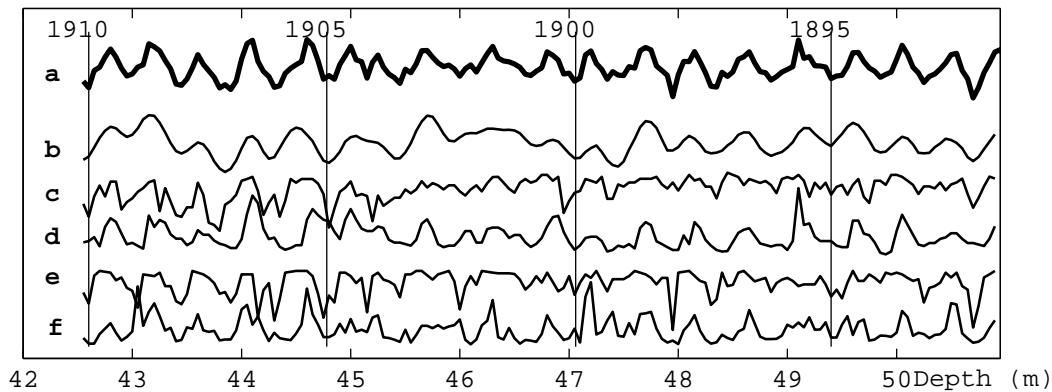
Missing data cannot be handled in this analysis, and must therefore be replaced by realistic values. The series used here only had isolated holes and a simple mean-of-neighbours scheme was applied to replace missing data and measurements that seemed improbable. The top of abnormal high peaks ascribed to volcanic eruptions was cut off such that the height of these peaks was reduced to approx. twice the average peak height. Prior to further analysis, the series were normalized to zero mean and unit variance in order to weigh the series equally. As a consequence of the normalization, *all* series on the plots are presented without dimensions.

The concentration of different chemical components peak at different times of the year, but as the objective of this work is to extract the annual signal from all the data series and to collect them into only one component, we decided to adjust the series such that they all peak simultaneously.

First one series is chosen relative to which the others are shifted. To synchronize another series with the fixed series, the time lag which maximizes the *absolute* value of the time lagged crosscorrelation function is computed and the other series is shifted accordingly. The crosscorrelation and the time lag computed here only relate to the synchronization of the series, and they have no relation to the dynamical decorrelation described in the previous sections. The maximum allowed time lag should be less than the number of samples per year to prevent the series from being shifted a full annual cycle. It should however still be large enough to allow synchronization of winter and summer peaks. After trying different values, we have settled on a maximum time lag of  $3/4$  of an average year. If the obtained crosscorrelation is negative, the sign of the adjusted series is changed such that the peaks are synchronized and have the same sign. The latter is done only to facilitate comparison between series.

For Site D, we have chosen to use the  $\delta^{18}\text{O}$  profile as the fixed signal and synchronized and/or inverted the others relative to this.

For the NorthGRIP data, we have systematically tried each of the series as the fixed series, and we found that using  $\text{Cl}^-$ ,  $\text{K}^+$  or  $\text{Na}^+$  worked best. This is probably due to the fact that they all exhibit



Figur C.1: Results of dynamical decorrelation of Site D data (a) with  $\tau = 5$  compared to the raw but synchronized data series of (b)  $\delta^{18}\text{O}$ , (c)  $\text{SO}_4^{2-}$ , (d)  $\text{NO}_3^-$ , (e)  $\text{Cl}^-$ , and (f) dust mass.

only one major annual peak which is rather broad. In this work, the importance of the synchronization has been investigated by performing the analysis both with and without the synchronizing of peaks. Although the difference is moderate, synchronizing the peaks gives somewhat better results.

## Results

The dynamical decorrelation method was applied with values of the time shift  $\tau$  ranging from 1 to 30, corresponding to a few annual cycles. This interval should be sufficiently large to reveal differences in the autocorrelation functions of the source series. The value of  $\tau$  that produced the source series with the most distinct annual component was individually chosen.

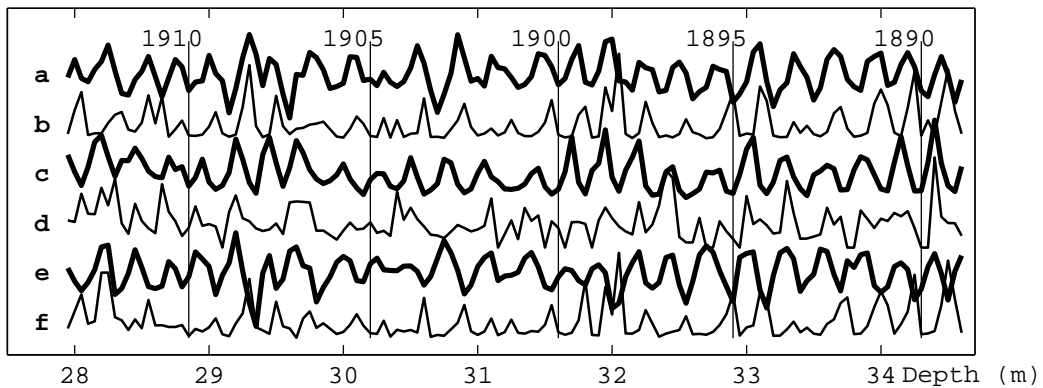
For Site D, the method produced the most pronounced annual signal for  $\tau = 5$ . The source signal displaying the annual variation is in figure 1 displayed together with the 5 input data series. The vertical lines indicate the dating used by Steffensen (1988). It is seen that the produced annual signal somewhat resembles the  $\text{NO}_3^-$ -series, but that it is significantly improved especially at year 1901, where the triple peak of the  $\text{NO}_3^-$ -series is replaced by a broad single maximum and at year 1908, where a double peak is replaced by a single peak.

As stated in the previous section, the synchronization of the NorthGRIP data series was performed prior to the analysis with different choices of the fixed series. Depending on the choice of fixed series, different values of  $\tau$  performed best. A total number of approx. 100 tests were performed with different combinations of fixed series and time lags  $\tau$ . The 3 combinations that performed best are displayed in figure 2 together with the series that was held fixed. In general our results show that the annual signal produced by dynamical decorrelation shows the closest resemblance to the one of the input series that was held fixed.

It is seen that the produced signals exhibit clear annual variation with peaks of comparable size and that the 3 source series agree well on where the peaks should be situated, although the sizes of the peaks vary considerably across the series.

To test the robustness of the results, the analysis was performed on two overlapping series containing either the first 60% or the last 60% of the series from Site D respectively, using the same value of  $\tau$ . The results agreed almost perfectly in the interval where the two parts overlapped. Also 10 samples were removed from each end of the series and the result was compared with the result from the analysis





Figur C.2: Examples of results from dynamical decorrelation (DD) of NorthGRIP data compared to the raw but synchronized data series. (a) DD with  $\text{Na}^+$  fixed and  $\tau = 3$ . (b) Raw  $\text{Na}^+$ . (c) DD with  $\text{K}^+$  fixed and  $\tau = 5$ . (d) Raw  $\text{K}^+$ . (e) DD with  $\text{Cl}^-$  fixed and  $\tau = 2$ . (f) Raw  $\text{Cl}^-$ .

with the full data set. Again almost perfect agreement was achieved, and from these two tests we thus conclude that dynamical decorrelation produces results that are largely indifferent to even considerable changes in input data selection.

The same method has also been applied to a different sequence of NorthGRIP data from a depth of approx 110 m. The results were not convincing, indicating that the method demands data with a resolution of at least 4-5 samples per year.

## A measure of performance

In order to evaluate the performance of the method, a simple "dating" scheme has been set up. For a given series, the number of peaks rising above a certain threshold is counted. The number of peaks should in the optimal case be equal to the number of years for a wide variety of threshold levels, such that the exact value of the threshold level is not critical to the result. Applying the scheme to count the number of peaks of e.g.  $\sin(2\pi t)$ ,  $t \in [0, 10]$ , the number of peaks will be 10 for threshold levels  $l \in ]-1, 1[$ , and zero for all other  $l$ . If the number of peaks  $p$  is plotted against  $l$ , it is from this graph alone easily verified that the number of peaks must be 10, and that the result does not depend on the exact choice of  $l$ .

In general the series' "countability" will be evaluated using this counting method. A series will be considered easy to count if the  $(l, p)$ -graph exhibits a plateau of (almost) constant  $p$  over a wide range of  $l$ , while the absence of such a plateau indicates that the series cannot be dated using this simple counting method.

This method of counting annual peaks is crude, and should therefore not necessarily yield perfect results. On the other hand it is our assumption that if the method of dynamical decorrelation improves the countability of the series when using such a simple counting method, it is most likely that the produced source series exhibit a stronger annual signal than the original data series and will be easier to date using almost *any* counting method.

The method counts peaks of any size as long as they cross the threshold level. A peak consisting of only one measurement that just barely exceeds the threshold level is thus counted as a peak. It is thus anticipated that noisy data series will have too many peaks when counted with this counting method.

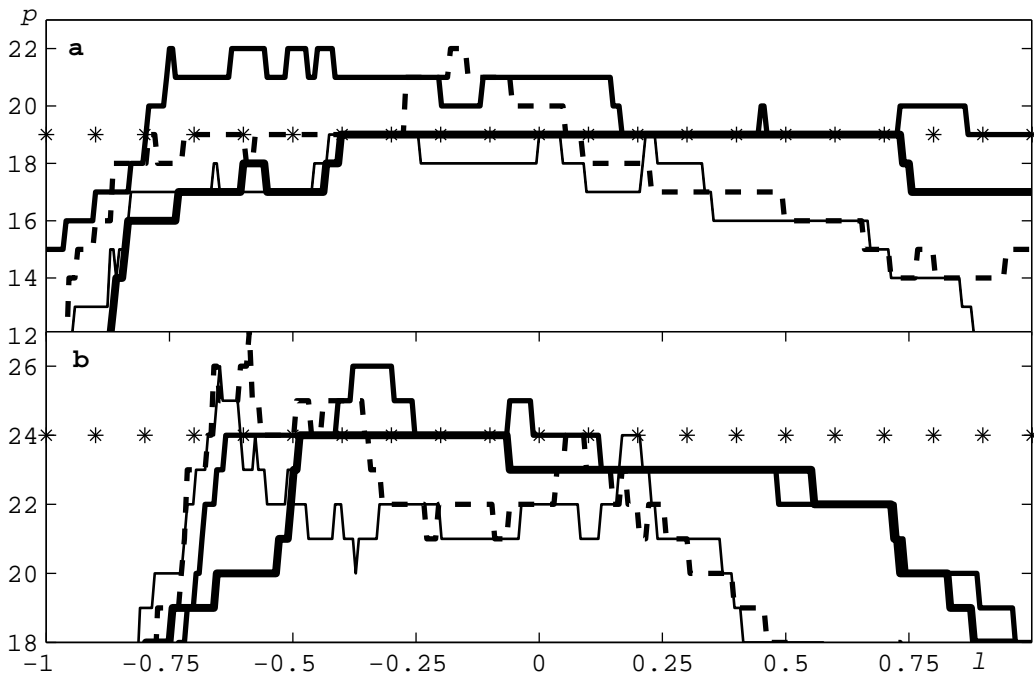


Figure C.3: The number of peaks  $p$  versus threshold level  $l$  for (a) Site D and (b) NorthGRIP. Results from the dynamical decorrelation filtered with a low-pass filter (bold line), DD unfiltered (medium line), raw filtered (thin line) and raw unfiltered (dashed line). The number of years obtained by manual counting (19 for Site D and 24 for NorthGRIP) is marked with asterisks.

To overcome this problem, we have tried to run the data through a low-pass filter before counting the number of peaks. A 5th order Butterworth digital low-pass filter with a cutoff frequency of  $4.5 \text{ yr}^{-1}$  provided good results.

The counting method has been used on both filtered and unfiltered results from the dynamical decorrelation and on raw data for comparison. The produced  $(l, p)$ -plots show considerable improvement in the countability of the number of peaks (figure 3).

Especially for Site D (figure 3a) the results are promising. When the counting method is applied on the filtered annual component from the dynamical decorrelation, the same number of years as obtained by manual layer counting is found for a wide interval of threshold levels  $l \in [-0.42, 0.71]$ . For the unfiltered signal, the number of peaks is less well-defined. The number of peaks in the raw data is counted in the same way but provides no usable results for dating, nor does the filtered but otherwise raw data.

For NorthGRIP, only the  $(l, p)$ -plot for the case where  $\text{Na}^+$  is fixed and  $\tau = 3$  is shown (figure 3b). Much like the Site D results, only the filtered dynamical decorrelation series provides good countability. With  $l \in [-0.50, -0.07]$  the result agrees with the 24 years determined by manual counting, while 23 years are found when  $l \in [-0.07, 0.55]$ . This arises from the fact that the peak representing the year 1904 (see figure 2) is less pronounced than the other annual peaks and thus only is counted if the threshold level is lower than  $-0.07$ . Features like the "two-level-plateau" in figure 3b can thus be used to identify doubtful annual layers.

## Discussion

When comparing the series in figure 1 and 2, it is seen that the method of dynamical decorrelation is able to extract a signal that exhibits a more pronounced annual variation than the input series themselves. It is demonstrated that the method cures the series of some problems of double or triple maxima and that the peaks in the produced signal have a relatively well-defined height. This facilitates automatic counting of the peaks by a simple method as demonstrated in figure 3.

The results produced are sensible to the choice of  $\tau$  and to some extent to which series is held fixed, but is otherwise fairly robust. By comparing the 3 series of figure 2 that are the results of dynamical decorrelation, one sees that although the produced series vary somewhat for different choices of fixed series and  $\tau$ , the overall appearance is similar.

Dynamical decorrelation is thus able to improve data series significantly prior to e.g. dating. Nevertheless the benefit in practical applications is at this moment limited by the need of choosing how to synchronize the data, choosing a suitable value for  $\tau$ , and selecting an appropriate cut-off frequency for the low-pass filter. Future investigations will show whether it is possible to make choices that will work well for longer data series from the same ice core. It is hoped that especially values of  $\tau$  can be found that will work on longer records from the same climatic period.

## Acknowledgments

The authors wish to thank Professor L.K. Hansen, Technical University of Denmark, for introducing us to the method of Molgedey and Schuster and for good discussions. The authors are indebted to J.P. Steffensen for fruitful discussions and suggestions. KKA is supported by the Carlsberg Foundation.

## References

- Clausen and others, 1988:** Clausen H.B., N.S. Gundestrup, and S.J. Johnsen 1988. Glaciological investigations in the Crête area, central Greenland: a search for a new deep-drilling site. *Ann. Glaciol.* 10, 10-15.
- Dahl-Jensen and others, 2001:** Dahl-Jensen D, N.S. Gundestrup, and others 2001. The NorthGRIP deep drilling program. This volume : 35A151.
- Hammer and others, 1978:** Hammer, C.U., H.B. Clausen, W. Dansgaard, N.S. Gundestrup, S.J. Johnsen, and N. Reeh 1978. Dating of Greenland ice cores by flow models, isotopes, volcanic debris, and continental dust. *J. Glaciol.* 20, 3-26.
- Hansen and others, 2000:** Hansen, L.K., J. Larsen, and T. Kolenda 2000. On Independent Component Analysis for Multimedia Signals. In Guan, L., S.Y. Kung and J. Larsen (eds.) *Multimedia Image and Video Processing*. Boca Raton, FL, CRC Press, 175-199.
- Hyvärinen, 1999:** Hyvärinen, A. 1999. Survey on Independent Component Analysis. *Neural Computing Surveys*, 2, 94-128.
- Molgedey and Schuster, 1994:** Molgedey, L., and H. Schuster 1994. Separation of Independent Signals using Time-Delayed Correlations. *Phys. Rev. Let.* 72(23), 3634-3637.
- Shuman and others, 1995:** Shuman, C.A., R.B. Alley, S. Anandakrishnan, J.W.C. White, P.M. Grootes, and C.R. Stearns 1995. Temperature and Accumulation at the Greenland Summit: Comparison of high-resolution isotope profiles and passive microwave brightness temperature trends. *J. Geophys. Res.* 100(D5), 9165-9177.
- Steffensen, 1988:** Steffensen, J.P. 1988. Analysis of the seasonal variation in Dust,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ , and  $\text{SO}_4^{2-}$  in two central Greenland firn cores. *Ann. Glaciol.* 10, 171-177.

# Litteratur

- [Alley *et al.*, 1997] Alley, R. B., C. A. Shuman, D. A. Meese, A. J. Gow, K. C. Taylor, K. M. Cuffey, J. J. Fitzpatrick, P. M. Grootes, G. A. Zielinski, M. Ram, G. Spinelli, and B. Elder, 1997, Visual-stratigraphic dating of the GISP2 ice core: Basic, reproducibility, and application, *Journal of Geophysical Research*, 102, (C12), 26367–26381.
- [Anderson & Benson, 1963] Anderson, D. L., and C. S. Benson, The densification and diagenesis of snow, In W. Kingery (editor), *Ice and Snow; properties, processes, and applications*, Proc. conf. M.I.T., Feb. 12-16, 1962, chapter 30, 391–403, M.I.T. Press, Cambridge, Massachusetts, 1963.
- [Azuma, 1994] Azuma, N., 1994, A flow law for anisotropic ice and its application to ice sheets, *Earth and Planetary Science Letters*, 128, 601–614.
- [Azuma & Higashi, 1985] Azuma, N., and A. Higashi, 1985, Formation processes of ice fabric pattern in ice sheets, *Annals of Glaciology*, 6, 130–134.
- [Biscaye *et al.*, 1997] Biscaye, P. E., F. E. Grousset, M. Revel, S. Van der Gaast, G. A. Zielinski, A. Vaars, and G. Kukla, 1997, Asian provenance of glacial dust (stage 2) in the Greenland Ice Sheet Project 2 ice core, Summit, Greenland, *Journal of Geophysical Research*, 102, (C12), 26765–26781.
- [Bishop, 1995] Bishop, C., *Neural networks for pattern recognition*, Oxford University Press, 1st edition, 1995.
- [Clausen & Hammer, 1988] Clausen, H. B., and C. U. Hammer, 1988, The Laki and Tambora eruptions as revealed in Greenland ice cores from 11 locations, *Annals of Glaciology*, 10, 16–22.
- [Clausen & Langway, Jr., 1989] Clausen, H. B., and C. C. Langway, Jr., The ionic deposits in polar ice cores, In H. Oeschger and C. C. Langway, Jr. (editors), *Dahlem Konferenz: The Environmental Record in Glaciers and Ice Sheets*, Physical, Chemical, and Earth Sciences Research Report 8, 225–248, John Wiley, New York, 1989.
- [Craig, 1961] Craig, H., 1961, Standard for reporting concentrations of deuterium and oxygen-18 in natural waters, *Science*, 133, (3467), 1833–4.
- [Dahl-Jensen, 1989] Dahl-Jensen, D., 1989, Steady thermomechanical flow along two-dimensional flow lines in large grounded ice sheets, *Journal of Geophysical Research*, 94, (B8), 10355–10362.
- [Dahl-Jensen *et al.*, 1993] Dahl-Jensen, D., S. J. Johnsen, C. U. Hammer, H. B. Clausen, and J. Jouzel, Past accumulation rates derived from observed annual layers in the GRIP ice core from Summit, Central Greenland, In R. W. Peltier (editor), *Ice in the Climate System*, vol. 12 of *NATO ASI Ser. I*, 517–532, Springer-Verlag, New York, 1993.

- [Dahl-Jensen *et al.*, 1998] Dahl-Jensen, D., K. Mosegaard, N. Gundestrup, G. D. Clow, S. J. Johnsen, A. W. Hansen, and N. Balling, 1998, Past temperatures directly from the Greenland ice sheet, *Science*, *282*, 268–271.
- [Dahl-Jensen *et al.*, 2001] Dahl-Jensen, D., N. Gundestrup, H. Miller, O. Watanabe, S. J. Johnsen, J. P. Steffensen, H. B. Clausen, A. Svensson, and L. B. Larsen, 2001, The NorthGRIP deep drilling program, *Annals of Glaciology*, *35*, (in press).
- [Dansgaard, 1953] Dansgaard, W., 1953, The abundance of  $^{18}\text{O}$  in atmospheric water and water vapour, *Tellus*, *5*, 461–469.
- [Dansgaard, 1954] Dansgaard, W., 1954, The  $\text{O}^{18}$ -abundance in fresh water, *Geochimica et Cosmochimica Acta*, *6*, (nos. 5/6), 241–260.
- [Dansgaard, 1961] Dansgaard, W., 1961, The isotopic composition of natural waters, *Meddelelser om Grønland*, *165*, (2), Særtryk, s. 1–120.
- [Dansgaard, 1964] Dansgaard, W., 1964, Stable isotopes in precipitation, *Tellus*, *16*, 436–468.
- [Dansgaard, 2000] Dansgaard, W., *Grønland i istid og nutid*, Rhodos, Copenhagen, 2000.
- [Dansgaard *et al.*, 1969] Dansgaard, W., S. J. Johnsen, J. Møller, and C. C. Langway, Jr., 1969, One thousand centuries of climatic record from Camp Century on the Greenland ice sheet, *Science*, *166*, 377–381.
- [Dansgaard & Johnsen, 1969] Dansgaard, W., and S. J. Johnsen, 1969, A flow model and a time scale for the ice core from Camp Century, Greenland, *Journal of Glaciology*, *8*, (53), 215–223.
- [Ditlevsen & Marsh, 1998] Ditlevsen, P. D., and N. D. Marsh, 1998, New method for identification of sources for chemical time series and its application to the Greenland Ice Sheet Project ice core record, *Journal of Geophysical Research*, *103*, (D5), 5649–5659.
- [Fuhrer *et al.*, 1999] Fuhrer, K., E. W. Wolff, and S. J. Johnsen, 1999, Timescales for dust variability in the Greenland Ice Core Project (GRIP) ice core in the last 100,000 years, *Journal of Geophysical Research*, *104*, (D24), 31043–31052.
- [Glen, 1955] Glen, J. W., 1955, The creep of polycrystalline ice, *Proceedings of the Royal Society, A*, *228*, (1175), 519–538.
- [Grootes & Stuiver, 1997] Grootes, P. M., and M. Stuiver, 1997, Oxygen 18/16 variability in Greenland snow and ice with  $10^{-3}$ - to  $10^5$ -year time resolution, *Journal of Geophysical Research*, *102*, (C12), 26455–26470.
- [Hammer, 1980] Hammer, C. U., 1980, Acidity of polar ice cores in relation to absolute dating, past volcanism, and radio-echoes, *Journal of Glaciology*, *25*, (93), 359–372.
- [Hammer *et al.*, 1978] Hammer, C. U., H. B. Clausen, W. Dansgaard, N. Gundestrup, S. J. Johnsen, and N. Reeh, 1978, Dating of Greenland ice cores by flow models, isotopes, volcanic debris, and continental dust, *Journal of Glaciology*, *20*, (82), 3–26.
- [Hammer *et al.*, 1994] Hammer, C. U., H. B. Clausen, and C. C. Langway, Jr., 1994, Electrical conductivity method (ECM) stratigraphic dating of the Byrd Station ice core, Antarctica, *Annals of Glaciology*, *20*, 115–120.

- [*Hammer et al.*, 1999] Hammer, C. U., K. K. Andersen, H. B. Clausen, D. Dahl-Jensen, C. Schött Hvidberg, and P. Iversen, The stratigraphic dating of the GRIP ice core, Special report, Department of Geophysics, Niels Bohr Institute of Astronomy, Physics and Geophysics, University of Copenhagen, 1999.
- [*Hansen et al.*, 2000] Hansen, L. K., J. Larsen, and T. Kolenda, On independent component analysis for multimedia signals, In L. Guan, S. Y. Kung, and J. Larsen (editors), *Multimedia Image and Video Processing*, chapter 7, 175–199, CRC Press, 2000.
- [*Herron & Langway, Jr.*, 1980] Herron, M. M., and C. C. Langway, Jr., 1980, Firn densification: An empirical model, *Journal of Glaciology*, 25, (93), 373–385.
- [*Hvidberg*, 1996] Hvidberg, C. S., 1996, Steady-state thermomechanical modelling of ice flow near the centre of large ice sheets with the finite-element technique, *Annals of Glaciology*, 23, 116–123.
- [*Hyvärinen*, 1999] Hyvärinen, A., 1999, Survey on independent component analysis, *Neural Computing Surveys*, 2, 94–128.
- [*Johnsen*, 1977] Johnsen, S. J., Stable isotope homogenization of polar firn and ice, In *Proc. of Symp. on Isotopes and Impurities in Snow and Ice, I.U.G.G. XVI, General Assembly, Grenoble Aug./Sept., 1975*, 210–219, IAHS-AISH Publ. 118, Washington D.C., 1977.
- [*Johnsen & Andersen*, 1978] Johnsen, S. J., and N. Andersen, 1978, On power estimation in maximum entropy spectral analysis, *Geophysics*, 43, 681–690.
- [*Johnsen et al.*, 1995] Johnsen, S., D. Dahl-Jensen, W. Dansgaard, and N. Gundestrup, 1995, Greenland palaeotemperatures derived from GRIP bore hole temperature and ice core isotope profiles, *Tellus*, 47B, 624–629.
- [*Johnsen et al.*, 1997] Johnsen, S. J., H. B. Clausen, W. Dansgaard, N. S. Gundestrup, C. U. Hammer, U. Andersen, K. K. Andersen, C. S. Hvidberg, D. Dahl-Jensen, J. P. Steffensen, H. Shoji, Á. E. Sveinbjörnsdóttir, J. White, J. Jouzel, and D. Fisher, 1997, The  $\delta^{18}\text{O}$  record along the Greenland Ice Core Project deep ice core and the problem of possible Eemian climatic instability, *Journal of Geophysical Research*, 102, (C12), 26397–26410.
- [*Johnsen et al.*, 1999] Johnsen, S. J., H. B. Clausen, J. Jouzel, J. Schwander, Á. E. Sveinbjörnsdóttir, and J. White, Stable isotope records from Greenland deep ice cores: The climate signal and the role of diffusion, In J. S. Wettlaufer, J. Dash, and N. Untersterner (editors), *Ice Physics and the Natural Environment*, vol. 56 of *NATO ASI Ser. I*, 89–107, Springer-Verlag, Berlin, etc., 1999.
- [*Johnsen et al.*, 2000] Johnsen, S. J., H. B. Clausen, K. M. Cuffey, G. Hoffmann, J. Schwander, and T. Creyts, Diffusion of stable isotopes in polar firn and ice: The isotope effect in firn diffusion, In T. Hondoh (editor), *Physics of Ice Core Records*, 121–140, Hokkaido University Press, Sapporo, 2000.
- [*Langway, Jr.*, 1967] Langway, Jr., C. C., 1967, Stratigraphic analysis of a deep ice core from Greenland, *CRREL Research Report 77*, 1–130.
- [*Meeker et al.*, 1995] Meeker, L. D., P. A. Mayewski, and P. Bloomfield, A new approach to glaciochemical time series analysis, In R. Delmas (editor), *Ice Core Studies of Global Biogeochemical Cycles*, vol. 30 of *NATO ASI Ser. I*, 383–400, Springer-Verlag, Berlin Heidelberg, 1995.

- [Meese *et al.*, 1997] Meese, D. A., A. J. Gow, R. B. Alley, G. A. Zielinski, P. M. Grootes, M. Ram, K. C. Taylor, P. A. Mayewski, and J. F. Bolzan, 1997, The Greenland Ice Sheet Project 2 depth-age scale: Methods and results, *Journal of Geophysical Research*, 102, (C12), 26411–26423.
- [Molgedey & Schuster, 1994] Molgedey, L., and H. Schuster, 1994, Separation of independent signals using time-delayed correlations, *Physical Review Letters*, 72, (23), 3634–3637.
- [Moore *et al.*, 1994] Moore, J. C., E. W. Wolff, H. B. Clausen, C. U. Hammer, M. R. Legrand, and K. Fuhrer, 1994, Electrical response of the Summit-Greenland ice core to ammonium, sulphuric acid, and hydrochloric acid, *Geophysical Research Letters*, 21, (7), 565–568.
- [Nye, 1951] Nye, J. F., 1951, The flow of glaciers and ice-sheets as a problem in plasticity, *Proceedings of the Royal Society, Ser A*, 207, (1091), 554–572.
- [Nye, 1957] Nye, J. F., 1957, The distribution of stress and velocity in glaciers and ice-sheets, *Proceedings of the Royal Society, Ser. A*, 239, (1216), 113–133.
- [Nye, 1963] Nye, J. F., 1963, Correction factor for accumulation measured by the thickness of the annual layers in an ice sheet, *Journal of Glaciology*, 4, (36), 785–788.
- [Petit *et al.*, 1999] Petit, J. R., J. Jouzel, D. Raynaud, N. I. Barkov, J. M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delaygue, M. Delmotte, V. M. Katlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. Pépin, C. Ritz, E. Saltzman, and M. Stievenard, 1999, Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, 399, 429–436.
- [Reeh, 1988] Reeh, N., 1988, A flow-line model for calculating the surface profile and the velocity, strain-rate, and stress fields in an ice sheet, *Journal of Glaciology*, 34, (116), 46–54.
- [Reeh, 1989] Reeh, N., Dating by ice flow modeling: A useful tool or an exercise in applied mathematics?, In H. Oeschger and C. C. Langway, Jr. (editors), *Dahlem Konferenz: The Environmental Record in Glaciers and Ice Sheets*, Physical, Chemical, and Earth Sciences Research Report 8, John Wiley, New York, 1989.
- [Rempel *et al.*, 2001] Rempel, A. W., E. D. Waddington, J. S. Wettlaufer, and M. G. Worster, 2001, Possible displacement of the climate signal in ancient ice by premelting and anomalous diffusion, *Nature*, 411, 568–571.
- [Schwander *et al.*, 2001] Schwander, J., J. Jouzel, C. U. Hammer, J. R. Petit, R. Udisti, and E. Wolff, 2001, A tentative chronology for the EPICA Dome Concordia ice core, *Geophysical Research Letters*, 28, (22), 4243 – 4246.
- [Shuman *et al.*, 1995] Shuman, C. A., R. B. Alley, S. Anandakrishnan, J. W. C. White, P. M. Grootes, and C. R. Stearns, 1995, Temperature and accumulation at the Greenland Summit: Comparison of high-resolution isotope profiles and satellite passive microwave brightness temperature trends, *Journal of Geophysical Research*, 100, (D5), 9165–9177.
- [Steffensen, 1988] Steffensen, J. P., 1988, Analysis of the seasonal variation in dust,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ , and  $\text{SO}_4^{2-}$  in two central greenland firn cores, *Annals of Glaciology*, 10, 171–177.
- [Steffensen *et al.*, 1996] Steffensen, J. P., H. B. Clausen, and J. M. Christensen, On the spatial variability of impurity content and stable isotope composition in recent summit snow, In E. W. Wolff and R. C. Bales (editors), *Chemical Exchange Between the Atmosphere and Polar Snow*, vol. 43 of *NATO ASI Ser. I*, 607–616, Springer-Verlag, New York, 1996.



- [Wagner *et al.*, 2001] Wagner, G., C. Laj, J. Beer, C. Kissel, R. Muscheler, J. Masarik, and H. A. Synal, 2001, Reconstruction of the paleoaccumulation rate of central Greenland during the last 75 kyr using the cosmogenic radionuclides  $^{36}\text{Cl}$  and  $^{10}\text{Be}$  and geomagnetic field intensity data, *Earth and Planetary Science Letters*, 193, 515–521.
- [Wilhelms *et al.*, 1998] Wilhelms, F., J. Kipfstuhl, H. Miller, K. Heinloth, and J. Firestone, 1998, Precise dielectric profiling of ice cores: A new device with improved guarding and its theory, *Journal of Glaciology*, 44, (146), 171–174.



# Figurfortegnelse

1.1	Nye-modellens ligevægts-antagelser. . . . .	11
1.2	Horisontal flydehastighed beregnet fra Glens lov og DJ-modellens forsimplede profil. . .	12
1.3	ECM-signal og $\delta^{18}\text{O}$ -sekvens for et kort stykke af Crête-iskernen. . . . .	17
1.4	Eksempel på resultater fra linescanning af NGRIP-iskernen . . . . .	18
1.5	Datering af GISP2-iskernen : Bestemmelse af årlag. . . . .	20
3.1	NGRIP datasættet med/uden regularisering. . . . .	43
3.2	Kumulerede fordelingsfunktioner for NGRIP-data. . . . .	45
3.3	Kumulerede fordelingsfunktioner for Site D data. . . . .	47
3.4	Kumuleret fordelingsfunktion for ECM-data fra Site D . . . . .	48
3.5	Kumuleret fordelingsfunktion for ECM-data fra Site D, udsnit. . . . .	48
5.1	Kildeserier produceret ved DD af Site D data og det oprindelige datasæt. . . . .	60
5.2	Resultater fra DD af Site D data og tælbarheden heraf. . . . .	61
5.3	Tælbarhed for de oprindelige Site D data. . . . .	62
5.4	$c_S(\tau)$ for Site D. . . . .	63
5.5	Følsomhed overfor serielængde ved DD af Site D data. . . . .	65
5.6	Forskellige resultater af DD af NGRIP data, oversigt. . . . .	67
5.7	Resultater fra DD af NGRIP data og tælbarheden heraf. Fastholdt $\text{NH}_4^+$ , $\tau = 8$ . . . . .	69
5.8	Resultater fra DD af NGRIP data og tælbarheden heraf. Fastholdt $\text{HCO}_3^-$ , $\tau = 8$ . . . . .	70
5.9	Tælbarheden af de oprindelige $\text{Mg}^{++}$ og $\text{Ca}^{++}$ -dataserier fra NGRIP. . . . .	71
5.10	$c_S(\tau)$ for NGRIP. . . . .	72
5.11	Resultater fra DD af NGRIP data med to samtidige $\tau$ -værdier . . . . .	73
5.12	Kildeserie produceret ved DD af Summit CFA-data og det oprindelige datasæt. . . . .	74
5.13	Resultater fra DD af Summit-data samt den rå $\text{NO}_3^-$ -serie og tælbarheden heraf. . . . .	75
5.14	$c_S(\tau)$ for CFA-data fra Summit. . . . .	76
B.1	Dæmpningsfaktor som funktion af bølgelængden. . . . .	86
B.2	Diffusionslængden $L$ som funktion af dybden for Camp Century-iskernen. . . . .	87
B.3	MEM Spektaltætheder for 5 5-meters intervaller af GRIP-iskernen. . . . .	88
B.4	$\delta^{18}\text{O}$ -profiler fra Crête-iskernen. . . . .	89
C.1	Artikel i Annals of Glaciology, figur 1. . . . .	96
C.2	Artikel i Annals of Glaciology, figur 2. . . . .	97
C.3	Artikel i Annals of Glaciology, figur 3. . . . .	98