

## Eksplorativ evaluering af samspillet mellem dioxin og fede syrer i fiskemel

Hvordan spiller disse parametre sammen og kan vi opstille en model, der forudsiger dioxinindhold ud fra fedtsyrekoncentrationer? Disse to spørgsmål vil blive besvaret i to kemometriske klummer. I den første vil vi være eksplorative og se på samspillet mellem dioxin og fedtsyrerne

Af Thomas Skov, Karin Kjeldahl, Søren Balling Engelsen, Rasmus Bro og Lars Nørgaard, Københavns Universitet

Dioxin i fødevarer er et ømtåleligt emne og især grænseværdien for indholdet af dioxin har været til debat. At kunne måle dioxin er en forudsætning for at kunne fastslå, om indholdet er over en vis grænse. Derfor har man gennem længere tid forsøgt at finde og udvikle nye pålidelige tids- og omkostningseffektive screeningsmetoder, der kan erstatte den eksakte men dyre og langsomme kemiske måling [1].

Man ønsker at udnytte, at der i fiskemel er en systematisk sammenhæng mellem dioxinindholdet og sammensætningen af fedtsyrer. Da fedtsyresammensætningen er langt hurtigere og lettere at måle end dioxin, kan dette udnyttes til en hurtigmetode. Den flittige læser af den kemometriske klumme vil allerede nu tænke: "for pokker da, her kan multivariat kalibrering nok anvendes" – og ganske rigtigt. Men inden vi kommer så langt, er det altid godt at tage den eksplorative hat på og se på de primære variationer, der findes i data. Til det formål har vi før vist, at Principal Component Analysis (PCA) med fordel kan anvendes, og her vil vi gå mere i detaljer med, hvordan scoreplottet og især loadingplottet kan fortolkes.

### Data

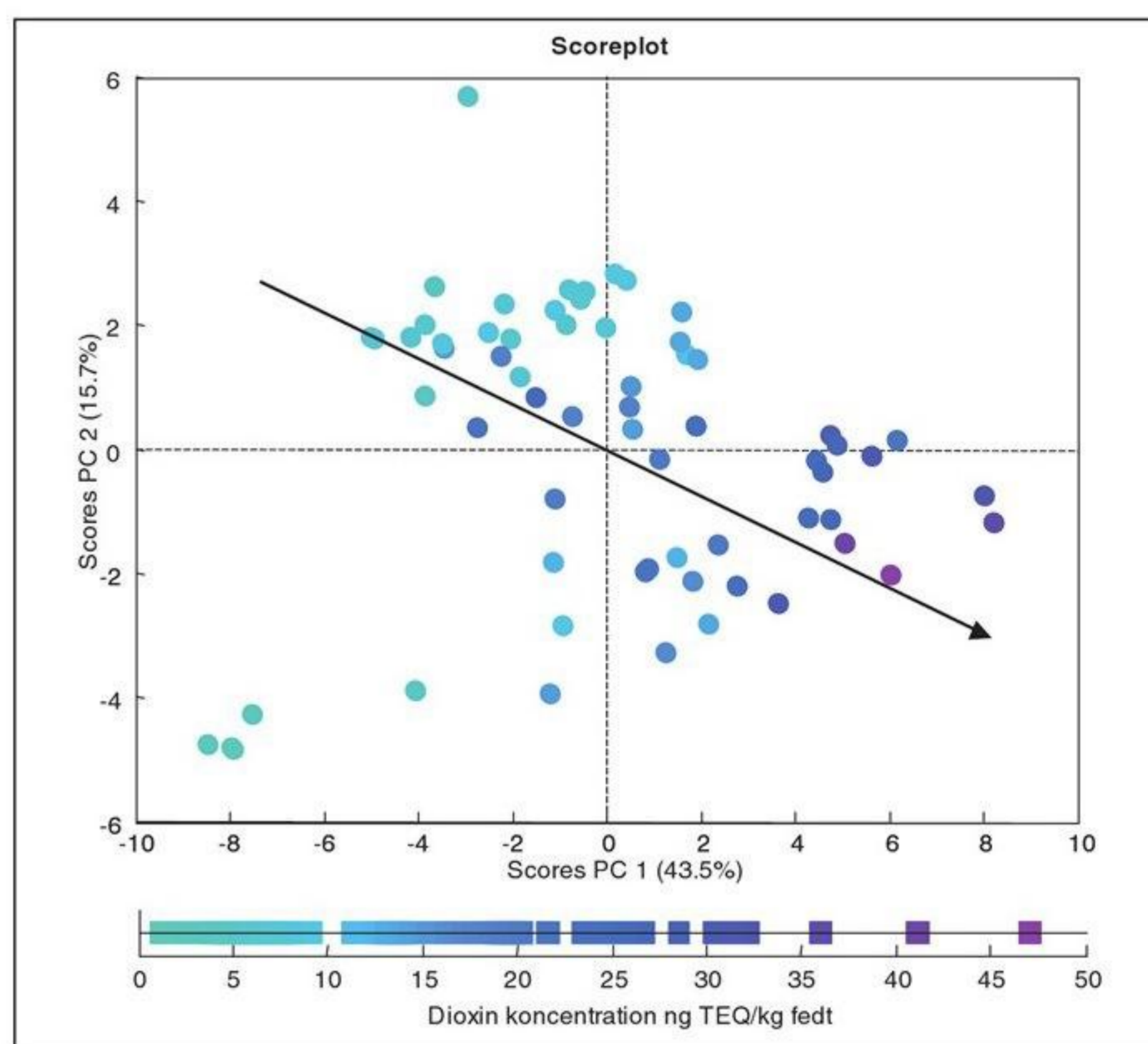
I 64 fiskemelsprøver er indholdet af dioxin bestemt til at ligge i intervallet 1,1 til 47,1 ng toksiske ækvivalenter pr. kilo fedt – (se [1] for information om dioxinbestemmelsen). For de samme prøver er sammensætningen af fedtsyrer bestemt gaskromatografisk med flammeioniseringsdetektion (GC-FID). Det giver en datablok med dimensionerne 64 prøver  $\times$  33 (32 fedtsyrer og dioxin). Da såvel dioxin som fedtsyrer har meget forskellige koncentrationsområder, er alle variablene autoskaleret, dvs. skaleret enkeltvis med én over standardafvigelsen for variabelen for at give dem samme mulighed for at påvirke modellen (Dansk Kemi nr. 1, 2008).

### Resultater

I figur 1 og 2 er henholdsvis scores og loadings for de første to principale komponenter plottet mod hinanden.

### Scoreplottet fortæller os hvordan

Af scoreplottet ses en klar sammenhæng mellem prøvernes indhold af dioxin og placeringen af prøverne. Man kan tolke dette, som at fiskemelsprøver, der ligger tæt på hinanden i scoreplottet, har mere eller mindre samme indhold af dioxin (se farveskala), men også samme fedtsyre sammensætning. Overordnet kan man sige, at scoreplottet fortæller os, *hvordan* prøverne opfører sig indbyrdes, men vi skal have fat i loadingplottet for at forklare, *hvorfor* prøverne ligger placeret, som de gør.



Figur 1. Scoreplot for PCA-model over dioxin- og fedtsyresammensætningen. Scoreplottet er farvet efter dioxinindholdet i niveauet 1,1-47,1 ng TEQ/kg fedt. Pilen indikerer den retning, der beskriver dioxin ud fra såvel farve som dioxins placering i loadingplottet – figur 2. Score- og loadingplottet beskriver ca. 59% af dioxin og fedtsyrevariationen (principal komponent 1: 43,5% og 2: 15,7%).

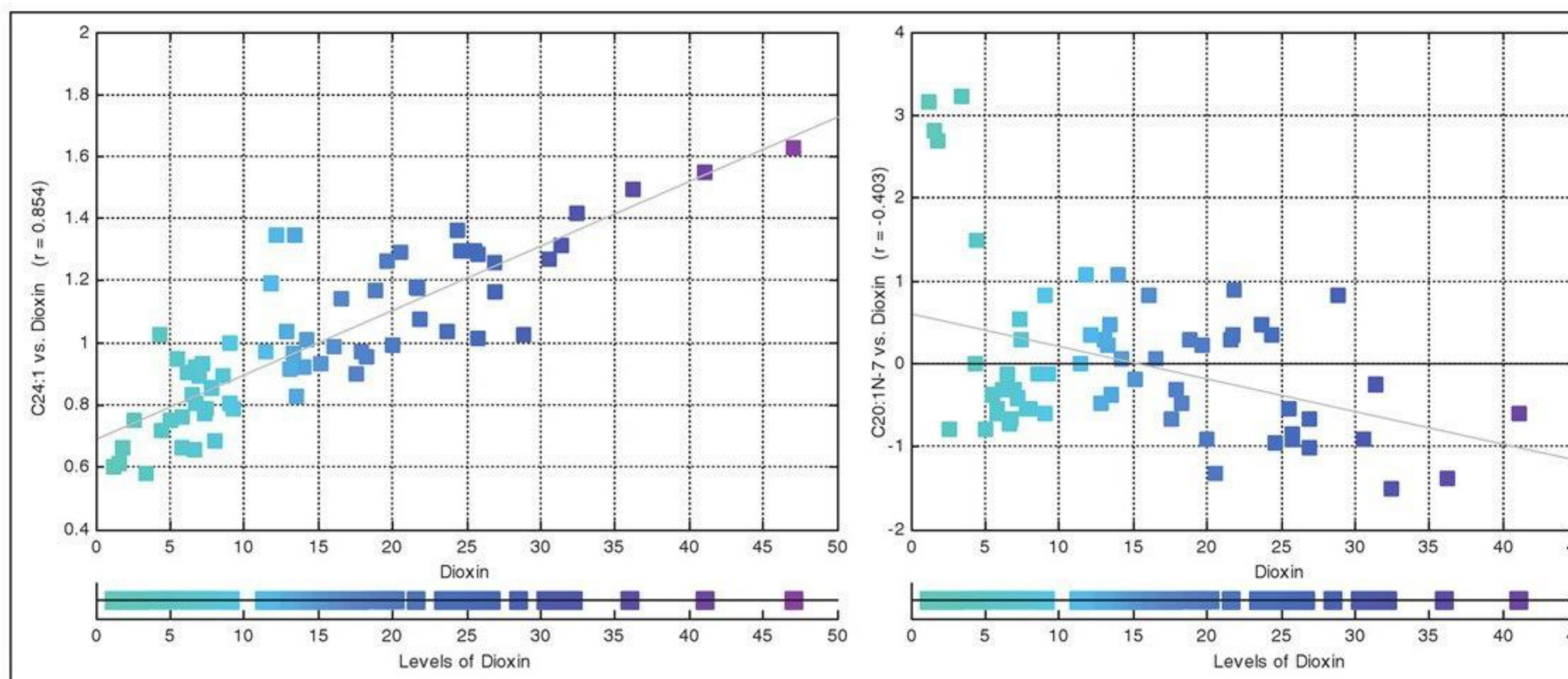
### Loadingplottet fortæller os hvorfor

Ud fra loadingplottet (figur 2), der indeholder både loadings for fedtsyrer og dioxin, kan man direkte se, at visse fedtsyrer hænger sammen med dioxin; nemlig de fedtsyrer, der ligger i samme retning som dioxin eller diagonalt modsat dioxin i plottet. Dioxin ligger tilmed langt ude af den første akse (PC1), hvilket betyder at den bidrager væsentligt til den variation, der er fundet og beskrevet i den første principale komponent (husk den første principale komponent er beskrivende for den retning i data som beskriver mest variation). Det ses ligeledes, at flere fedtsyrer ligger tæt på eller diametralt modsat dioxin.

Det er dog ikke nok bare at ligge i samme/modsat retning. Ligger en variabel langt fra (0,0) – f.eks. C24:1n-9 – har den stor lighed med dioxinindholdet, mens en variabel tættere på (0,0) – f.eks. C14:1n-5 – siger meget lidt om den variation, der beskrives i scoreplottet. Ligger en fedtsyre langt fra (0,0), men i en anden retning, bidrager den til den systematiske variation, men fortæller ikke meget om dioxin.

For blot at tage nogle eksempler så ses det, at når dioxinindholdet stiger, så stiger indholdet af fedtsyrerne C18:1n-9

Figur 3. Sammenhængen (korrelationen) mellem dioxin og henholdsvis C24:1n-9 (venstre) og C20:1n-7 (højre). Korrelationskoefficienten,  $r$  er ligeledes angivet –  $r = 1$ : perfekt korrelation,  $r = 0$ : ingen korrelation og  $r = -1$ : omvendt korrelation. Punkterne er farvet efter dioxinniveau (se figur 1).



og C24:1n-9, mens indholdet af C21:5n-3 og C16:4n-1 falder. Disse observationer kan let findes ved at tegne en linje gennem dioxins placering (●) og (0,0) og tilsvarende for fedtsyrernes placering (●) (figur 2).

Vinklen mellem linjerne for dioxin og fedtsyrer viser, hvor meget variationen for fedtsyrerne ligner variationen for dioxin. F.eks. er vinklen mellem dioxin og C18:1n-9-linjerne tæt på  $0^\circ$ , mens den mellem dioxin og C21:5n-3-linjerne er tæt på  $180^\circ$ . Når vinklen er tæt på  $0^\circ$ , siger man, at de to variable er positivt korrelerede (når den ene stiger i koncentration, stiger den anden også), mens en vinkel på  $180^\circ$  fortæller, at variablene er omvendt korrelerede (når den ene stiger, falder den anden).

Ser man derimod på fedtsyren C20:1n-7, så ligger den med en vinkel tæt på  $90^\circ$  ift. dioxin. Dette betyder, at denne fedtsyre og dioxin ikke varierer systematisk sammen. Man anvender ofte terminologien, at de to variable er ortogonale – som betyder at de to variable ikke kan anvendes til at sige noget om hinanden.

Konklusionerne fra loadingplottet skal tages med det forbehold, at det viste plot kun fortæller en del af historien. Det gælder især, hvis den datavariation, der beskrives af PCA-modellen, er lille. De to første PCA-komponenter forklarer ca. 59% af variationen (figur 1 og 2) blandt fedtsyrerne og dioxin. For disse 59% er der en god korrelation mellem dioxin og C24:1n-9, men vi har i dette plot ikke mulighed for at se, hvad der sker i de resterende 41%. Vi kan plote andre PCA-komponenter og

blive endnu klogere eller plote rådata og få bekræftet, om der er en god sammenhæng mellem dioxin og fedtsyren C24:1n-9.

## Tjek loadingplottet i rådata

Som tommelfingerregel bør man også altid plote rådata for de interessante variable, for at tjekke hvor god sammenhængen (korrelationen) egentlig er. Korrelationen mellem dioxin og henholdsvis C24:1n-9 og C20:1n-7 er verificeret i figur 3.

I figur 3 ses at fortolkningen i loadingplottet er korrekt, nemlig at der er en god sammenhæng mellem dioxin og fedtsyren C24:1n-9 ( $r = 0,9$ ) og en ringe sammenhæng mellem dioxin og C20:1n-7 ( $r = -0,4$ ). I PCA-modeller, hvor variablene ligger i samme retning i loadingplottet, men hvor modellen ikke beskriver ret meget variation (f.eks.  $< 20\%$  i de første to PCA-komponenter), vil et plot af rådata (som i figur 3) sikre, at man ikke overfortolker de sammenhænge, man observerer i loadingplottet.

Har man den angivne ”forklaret varians” i baghovedet, er loadingplottet et utroligt nyttigt værktøj til at få overblik over samspillet mellem såvel fedtsyrer indbyrdes, men endnu vigtigere her også til dioxin.

## Outro

Vi har med denne klumme arbejdet eksplorativt med at finde sammenhæng mellem fiskemels fedtsyreprofiler og dioxinindhold. Sådanne empiriske sammenhænge fortæller desværre ikke noget om de kausale sammenhænge mellem fedtsyrer og dioxinindhold – vi kan blot få en idé til den hypotese, at fedtsyresammensætningen er afhængig af fiskens trofiske niveau og dermed indirekte korreleret til fiskens dioxinindhold.

I næste kemometriske klumme vil vi arbejde videre med disse data og se på, hvordan vi kan opstille en model, der kan forudsige dioxinindholdet ud fra sammensætningen af fedtsyrerne. Vi har dog allerede her fået et overblik over data og en kraftig antydning af, at dette er muligt og endog hvilke fedtsyrer, der højst sandsynligt vil kunne anvendes til formålet.

E-mail-adresser:

Thomas Skov: thsk@life.ku.dk

Karin Kjeldahl: kkd@life.ku.dk

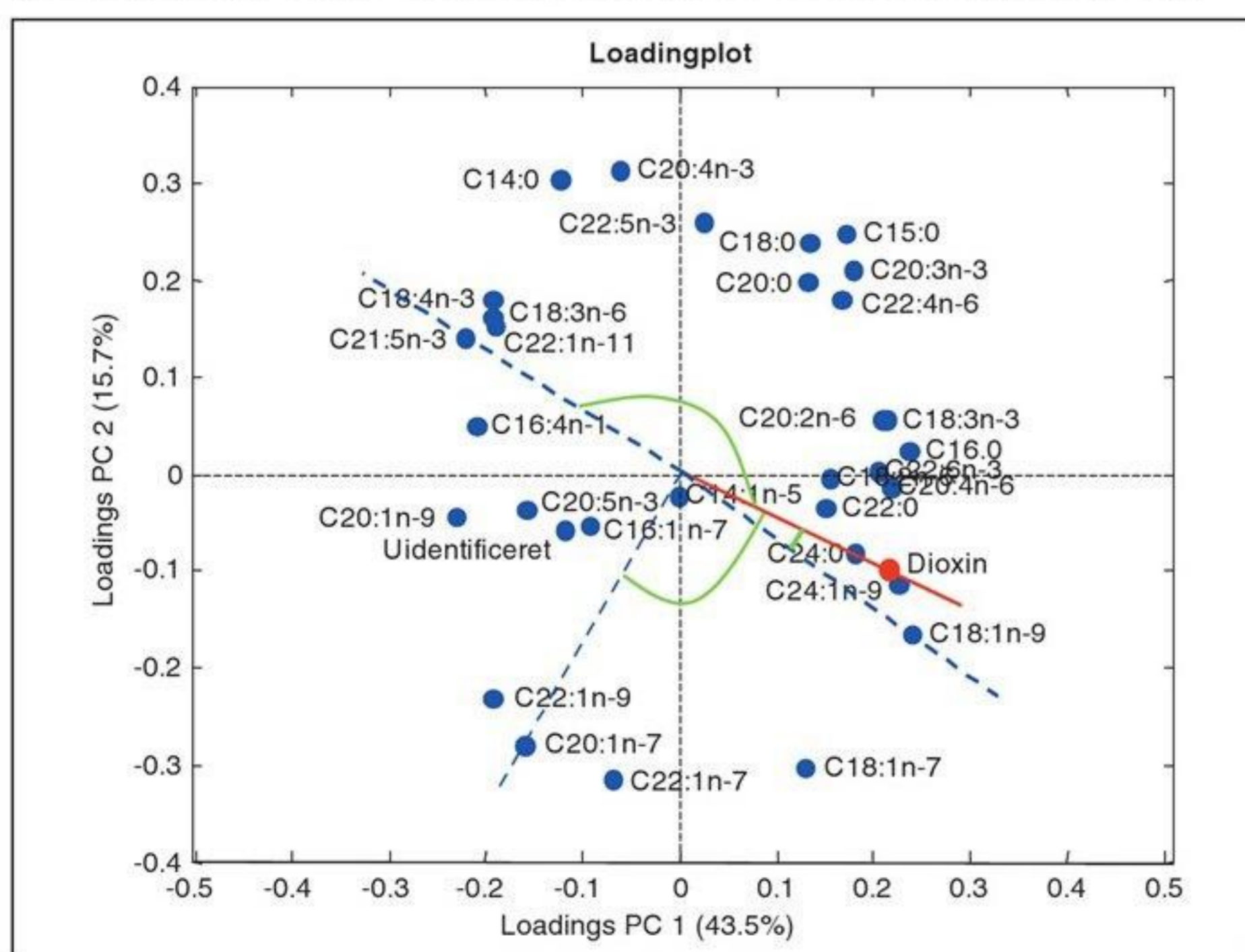
Rasmus Bro: rb@life.ku.dk

Søren Balling Engelsen: se@life.ku.dk

Lars Nørgaard: lan@life.ku.dk

Reference

1. Rapid dioxin assessment in fish products by fatty acid pattern recognition, Marc Bassompierre, Lars Munck, Rasmus Bro and Søren Balling Engelsen, Analyst, 129, 553-558, 2004.



Figur 2. Loadingplot med streger indtegnet for den retning som dioxin samt fedtsyrerne C21:5n-3, C18:1n-9 og C20:1n-7 beskriver. De grønne linjer indikerer vinklen mellem to variable – f.eks. ses en vinkel mellem dioxin og C21:5n-3 på tæt ved  $180^\circ$ .