

Sukker, fluorescens og SIMCA

Komplekse fluorescensspektre af en sukkerprøve opløst i vand kan fortælle, hvor prøven er produceret. Den kemometriske metode til at løse et sådant problem er Soft Independent Modeling of Class Analogy (SIMCA)

Af Lars Nørgaard, Søren Balling Engelsen og Rasmus Bro, Institut for Fødevidenskab, Det Biovidenskabelige Fakultet, Københavns Universitet

I 40'erne evaluerede man kvaliteten af sukker ved at belyse prøven med ultraviolet lys [1]. Jo mere blåligt skær prøven havde i det ultraviolette lys, jo mere uren var den. Uren er nok et stærkt ord at bruge om dagens sukkerproduktion. Sukker er et uhyre rent produkt: almindeligt roesukker består af mere end 99,9% sukrose og små mængder af andre komponenter som aminosyrer og phenoliske komponenter. Visse af disse komponenter er fluorescerende, og det er koncentrationen af disse, der ved UV-testen afgør, om prøven betragtes som ren.

I 1993 lancerede professor Lars Munck ideen om at bruge moderne fluorescensspektroskopi til kvalitetsanalyse af sukkerprøver, og han indledte derfor et samarbejde med Danisco Sugar. I stedet for kun at belyse sukkerprøven med en UV-lampe og visuelt vurdere resultatet, foreslog han at måle nøjagtige fluorescenslandskaber for hver prøve og at analysere disse kemometrisk for at finde ligheder, forskelle og grupperinger. Formålet var at udvikle nye online kvalitetsstyringssystemer til sukkerproduktion. Det er baggrunden for den analyse med SIMCA, vi nu demonstrerer.

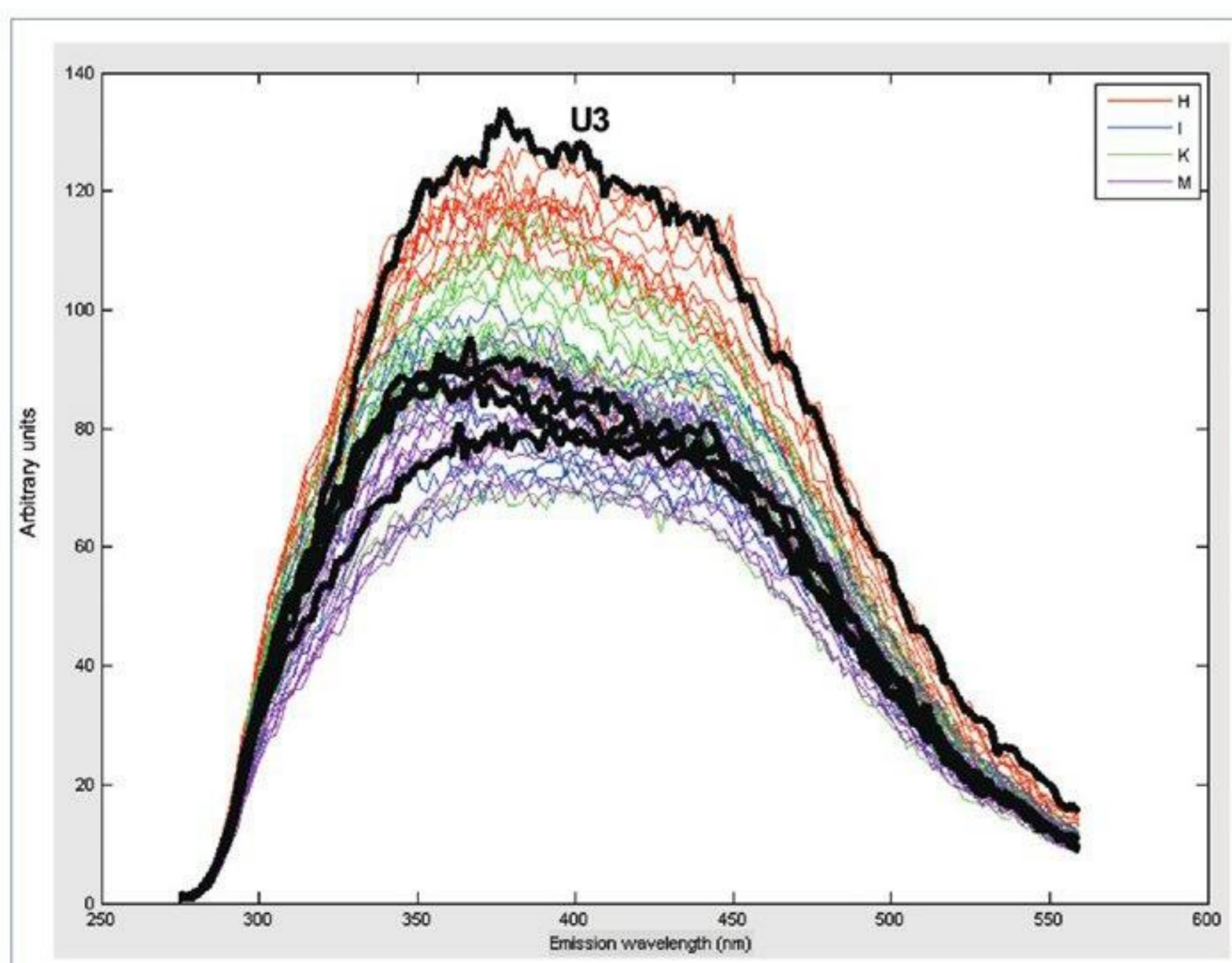
Data

Vi analyserer et prøvesæt med 57 prøver fra fire forskellige sukkerfabrikker (danske og tyske) leveret af Danisco. Sættet består af 14 prøver fra **fabrik H**, 12 fra **fabrik I**, 13 fra **fabrik K** og 13 fra **fabrik M**; derudover er der fem ukendte (U1-U5)

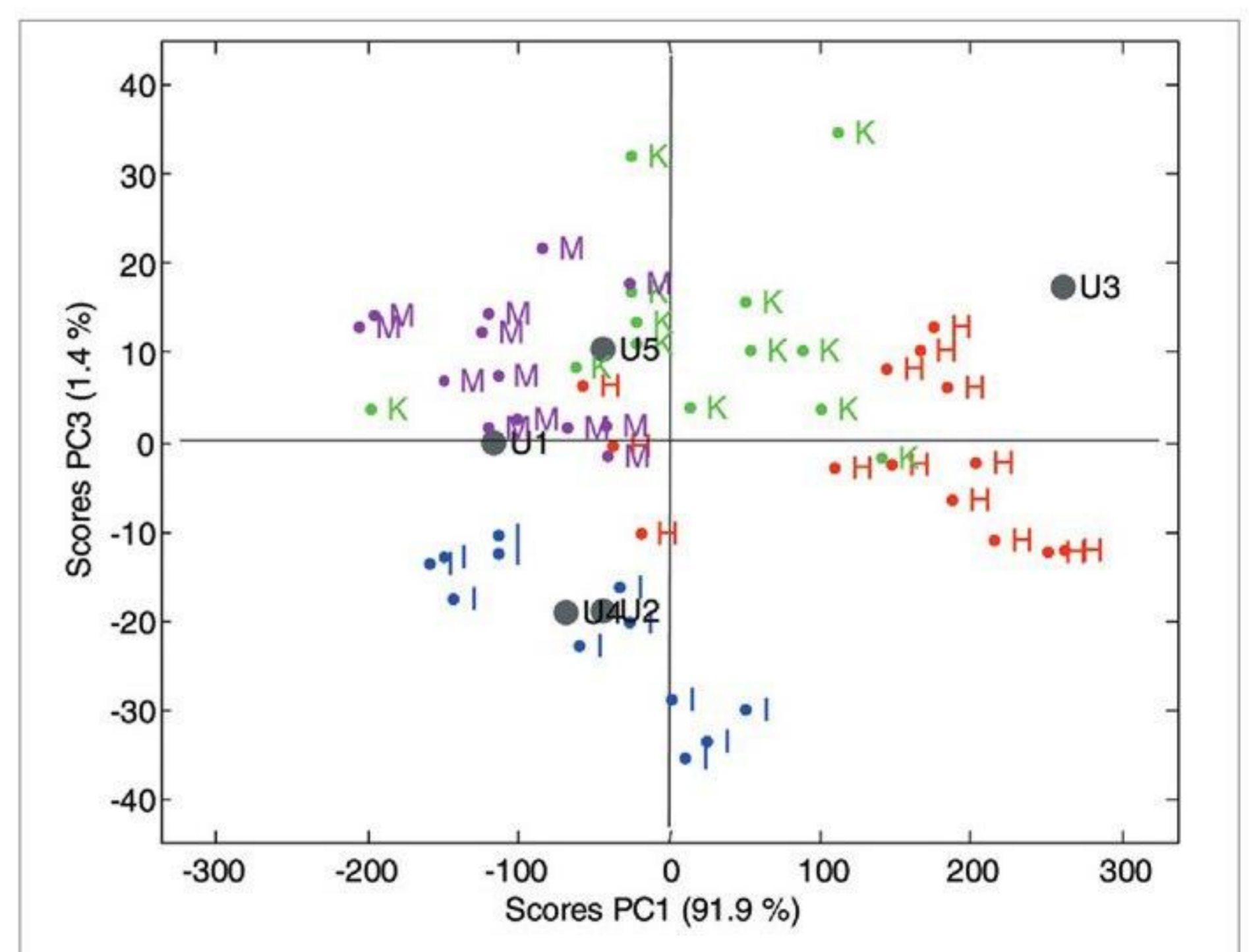


prøver i sættet, som klassificeres for at verificere metodens potentiale. Det er vigtigt, at de prøver, der udvælges til at beskrive de enkelte fabrikker, er repræsentative og dækkende. Ellers vil fremtidige klassifikationer fejlagtigt bedømme visse prøver til ikke at tilhøre grupperne.

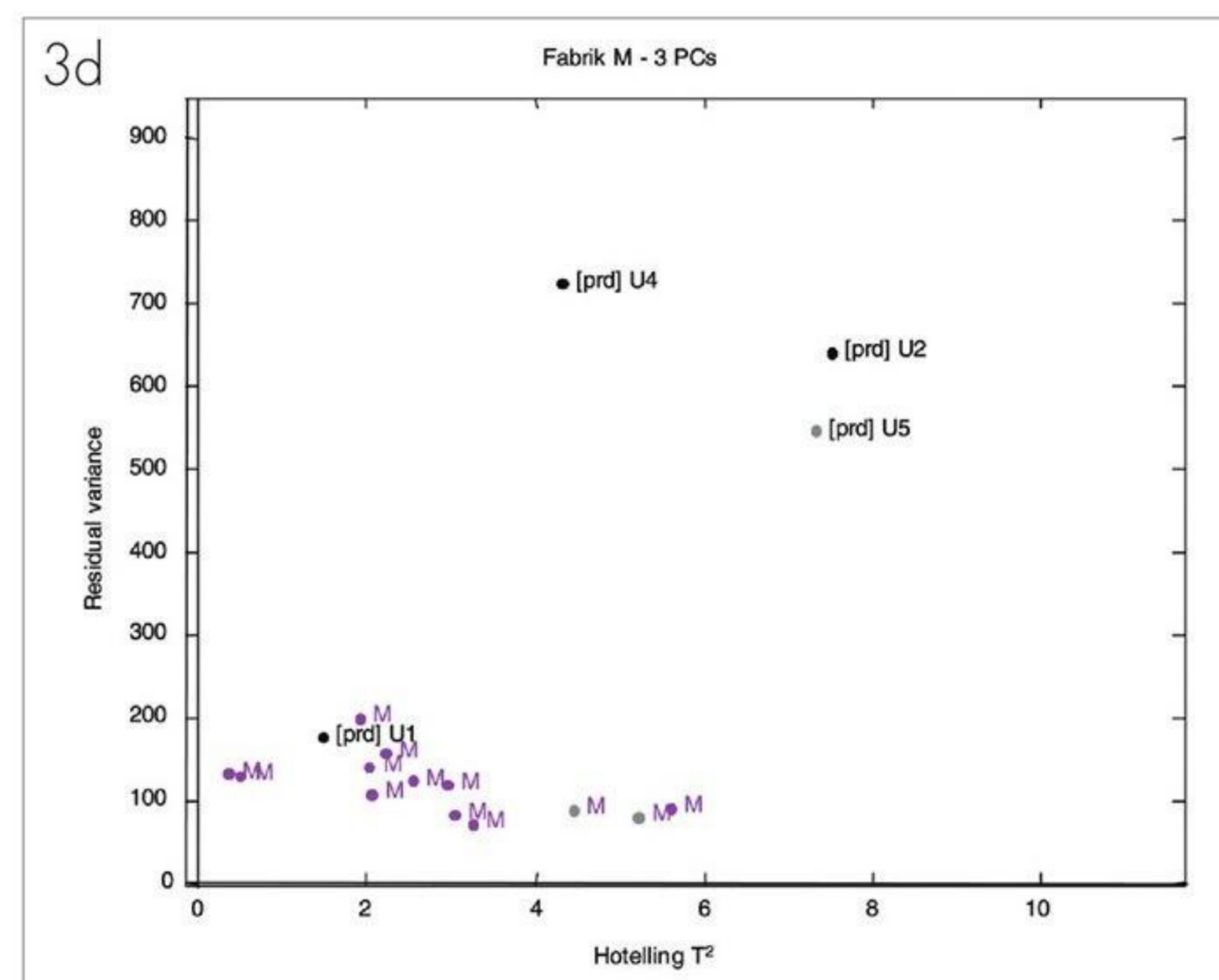
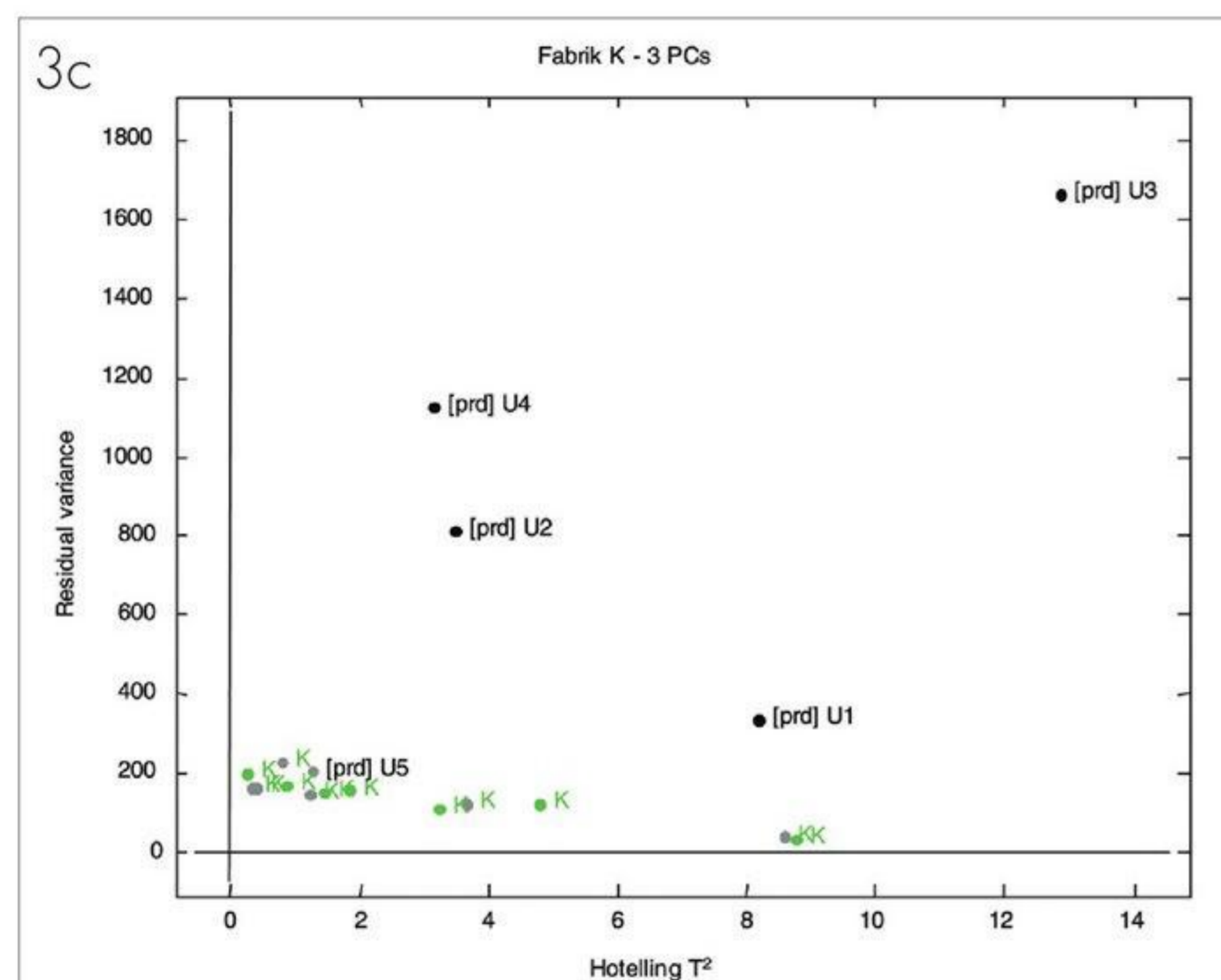
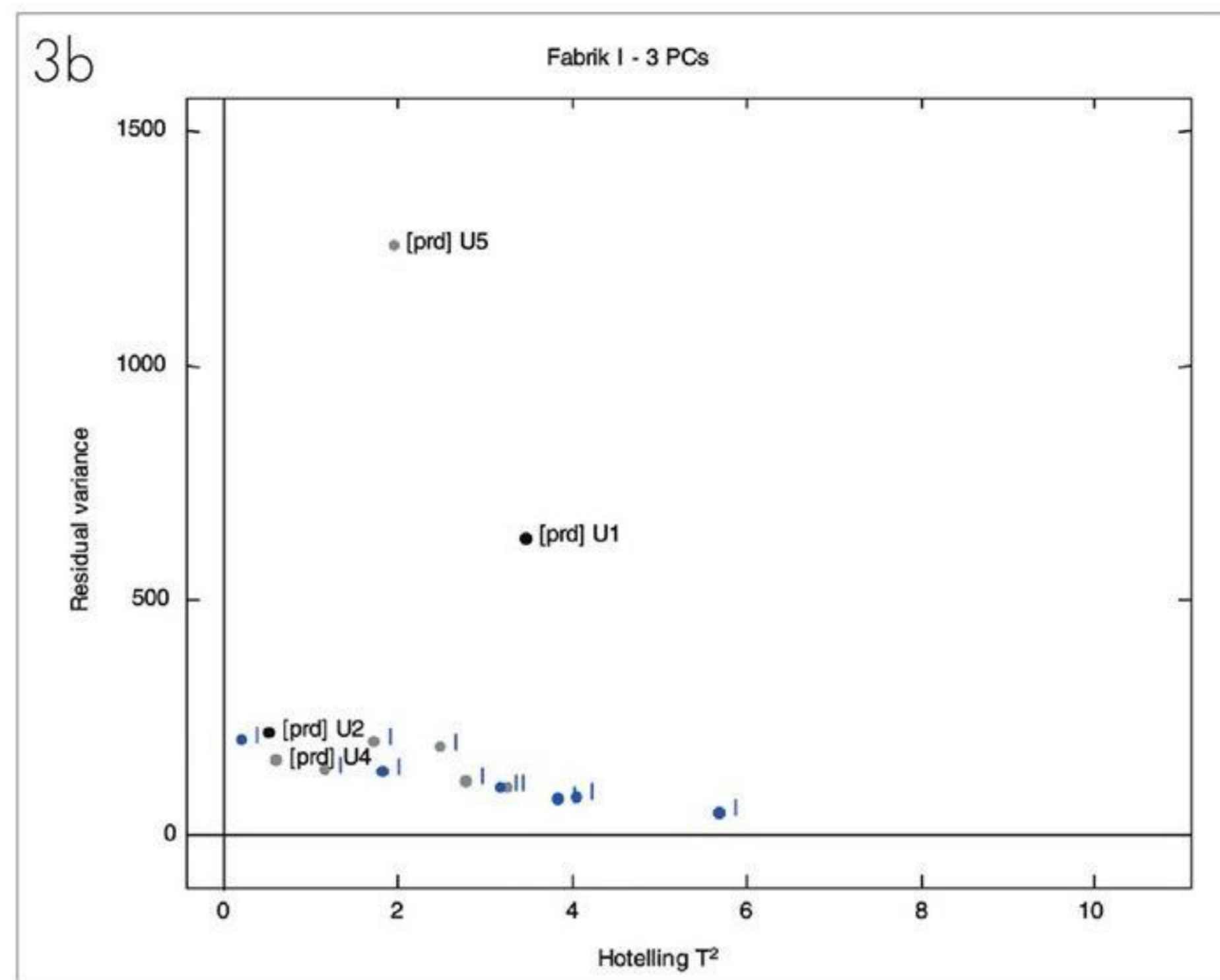
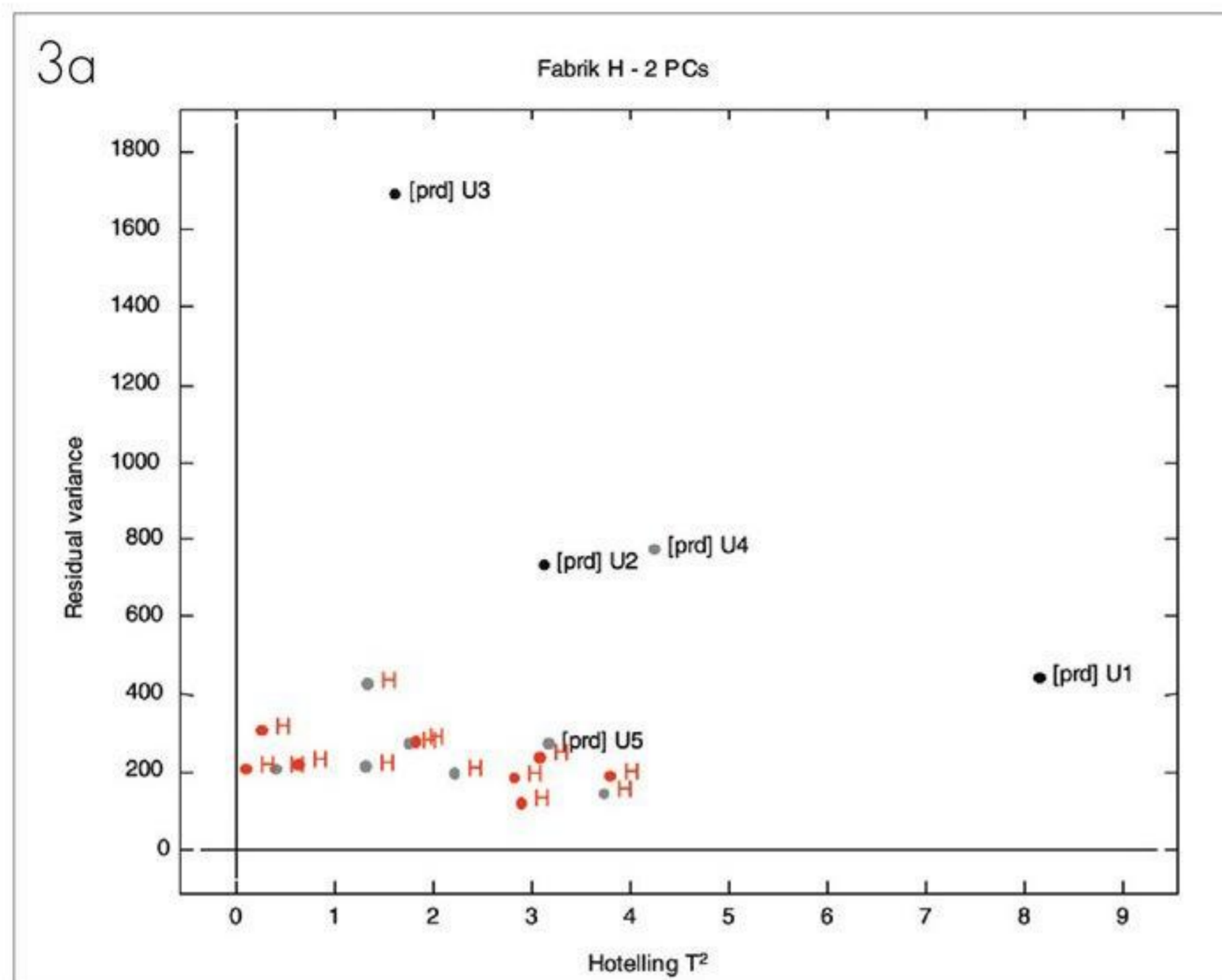
Prøveforberedelsen består i at opløse 2,25 gram sukker i ►



Figur 1. Fluorescens-emissionsspektre af 57 sukkerprøver. Prøverne er fra fire forskellige fabrikker og fem ukendte er fremhævet.



Figur 2. PCA-scoreplot af alle 57 prøver inkl. de fem ukendte (U1-U5). Prøver fra fire fabrikker er angivet med betegnelserne H, I, K og M.



Figur 3a-d. Residualvarians vs. T^2 for de fire PCA-modeller inkl. placering af de ukendte prøver i modellen. Se tekst for uddybning.

15,0 mL ionbyttet vand. Fluorescens-emissionsspektre optages på opløsningen ved fire forskellige excitationbølgelængder (230, 240, 290, 340 nm) med et LS50B-instrument fra Perkin-Elmer [1]. I dette eksempel anvendes kun emissionsspektrene ved excitation på 240 nm. De ses i figur 1 farvet efter fabriktilhørsforhold. I den videnskabelige publikation [2] udnyttes hele landskaber i dataanalysen, men det gemmer vi til en senere klumme.

Ved inspektion af figur 1 ses visse tendenser i rådata: prøver fra fabrik H har høje intensiteter, prøver fra fabrik K lidt lavere, og prøver fra fabrik I og M har overlappende spektrale signaturer. De ukendte prøver placerer sig med både høje og lave intensiteter med U3 som prøven med højeste intensitet. Det er ikke umiddelbart let at klassificere prøverne ift. de fire grupper, men U3 kunne man tentativt fristes til at foreslå som en prøve fra fabrik H.

Global PCA-model

For at få et bedre overblik over data laves en PCA på alle prøver inkl. de ukendte (centrerede data). Scoreplottet i figur 2 viser komponent 1 vs. komponent 3, da denne kombination grupperer

fabrikkerne bedst. Ud fra dette plot ses, at **fabrik I** skiller sig ud fra de tre andre fabrikker ved en placering med høje negative komponent 3 scoreværdier, mens **fabrikkerne M, K og H** adskilles langs komponent 1.

Prøverne U2 og U4 placerer sig fint i I-gruppen, og det er en stærk indikation af, at disse prøver er fra fabrik I. U1 og U5 er sværere at placere entydigt, mens U3 er placeret tæt på H-gruppen uden dog at være helt tæt på. Figur 2 viser kun en lille del af variationen, så endelige konklusioner er svære at drage uden en mere tilbundsående analyse.

SIMCA

Første trin i SIMCA er at lave lokale PCA-modeller for hver gruppe (centrerede data). For disse grupper når vi frem til følgende modeller primært ved hjælp af fuld krydsvalidering og inspektion af struktur/støj i loadingplots:

Fabrik H: 2 komponenter

Fabrik I: 3 komponenter, én outlier fjernet

Fabrik K: 3 komponenter

Fabrik M: 3 komponenter

Tre komponenter kan være lidt vel rigeligt for et datasæt med kun 11 prøver, men i dette tilfælde er det tydeligt at den tredje komponent er reel bl.a. vurderet ud fra den spektrale form af loadingvektoren. Det ville dog have været godt at have et større datasæt.

Vi vil nu klassificere de fem ukendte prøver ift. de fire PCA-modeller. Til det formål projiceres de fem ukendte prøver ind i hver af de fire PCA-modeller ved det optimale antal komponenter, og modellerne inspiceres vha. Residualvarians vs. Hotelings T^2 -plottet (Dansk Kemi, nr. 3, 2008). I figur 3a-d ses de fire relevante plots.

Fabrik H: Prøve U5 ser ud til at placere sig i denne gruppe; ingen andre prøver tilhører gruppen.

Fabrik I: Prøve U2 og U4 må klart anses for at være prøver fra fabrik I.

Fabrik K: Prøve U5 placerer sig som den eneste prøve i gruppen.

Fabrik M: Prøve U1 er en prøve fra fabrik M; alle andre prøver ligger langt fra M-modellen (U3 ligger langt uden for den viste del af plottet).

Ovenstående betyder, at U3-prøven ikke kan klassificeres som tilhørende nogle af de fire fabrikker. Derfor må den betegnes som en outlier, der måske stammer fra en femte fabrik. Det står i modsætning til rådata og delvist til det globale PCA-scoreplot, men ved at fokusere med SIMCA ses det, at prøven ikke er fra

fabrik H grundet høj residualvarians. Det betyder, at prøvens mønster afviger fra mønsteret for de kendte prøver fra **fabrik H**.

Vi har således klassificeret alle fem prøver ift. deres tilhørsforhold til de fire PCA-modeller, og den efterfølgende kontrol hos Danisco bekræftede resultaterne.

Outro

Baseret på dette indledende studie blev der efterfølgende lavet en række studier inden for området bl.a. med henblik på at karakterisere de fluorescerende komponenter og yderligere raffinering af fluorescensmetode til karakterisering og proceskontrol.

E-mail-adresser:

Lars Nørgaard: lan@life.ku.dk

Søren Balling Engelsen: se@life.ku.dk

Rasmus Bro: rb@life.ku.dk

Referencer

1. Nørgaard L, *Classification and prediction of quality and process parameters of thick juice and sugar by fluorescence spectroscopy and chemometrics*, Zuckerindustrie 120(11), 970-981, 1995.
2. Bro R, *Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis*, Chemometrics and Intelligent Laboratory Systems, 46(2), 133-147, 1999.

Ren gas er god kemi

strandmollen.dk

Hjemmesiden med den bedste rådgivning om gas til alle formål!

STRANDMØLLEN



70102107

Strandmøllen A/S har 90 års erfaring med produktion og salg af gas. Det forpligter. Derfor går vi aldrig på kompromis med kvalitet, sikkerhed og rådgivning.