

PCA som eksplorativt værktøj – karakterisering af en dynamisk proces

Eksplorativ data analyse er helt central i den kemometriske tilgang til problemløsning. Her vil vi demonstrere anvendelsen af PCA til at give overblik over en friteringsproces monitoreret med to forskellige typer af analyser.

Af Søren Balling Engelsen, Lars Norgaard og Rasmus Bro, Københavns Universitet

Ved intern kvalitetskontrol i industrien er der tradition for at analysere produkter og procesprøver med en række laboratoriemetoder som tørstof, protein, % fedt, iodtal, anisidantal, energindhold, vandaktivitet og lignende uden at udnytte kovariationen imellem disse. Samtidigt begynder man at benytte moderne spektroskopiske metoder som nærinfrarød (NIR) spektroskopi til fx at bestemme % fedt uden egentlig at udnytte det fulde potentiale, som disse spektroskopiske data indeholder.

Friteringsproces

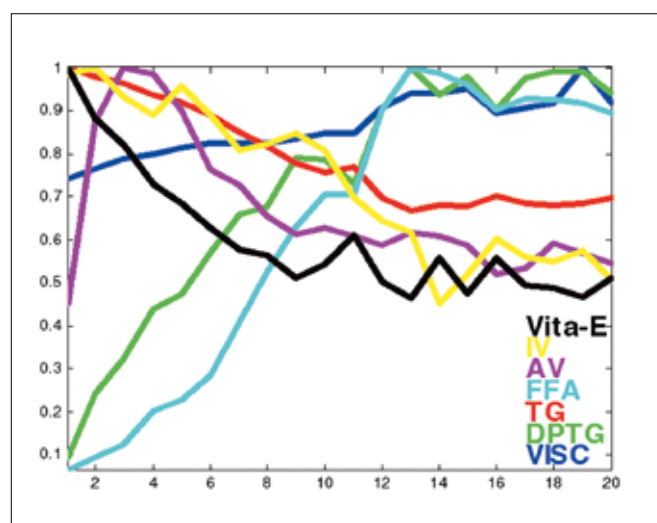
Friteringsprocessen er et glimrende eksempel på en kompleks fødevarerindustriell proces. Det tilsyneladende simple spørgsmål: "Hvornår skal fritureolien udskiftes?" viser sig at være et komplekst optimeringsproblem, der først og fremmest adresseres ud fra en sensorisk problemstilling, men der vil også være en stribe afledte problemstillinger, såsom lovmæssige krav, krav til holdbarheden af de friterede produkter, økonomisk rentabilitet, toksikologiske og ernæringsmæssige konsekvenser osv. Derudover vil svaret på spørgsmålet være afhængig af en stribe forskellige procesparametre som fx kvaliteten af den fritureolie, der benyttes, friteringsobjektet, friteringstemperaturen, mængden af dagligt tilført ny olie (turnover), ilt-tilførsel, brugen og karakteren af antioxidanter, filtrering, vedligeholdelse af friteringsudstyret, vedligeholdelse af fritureolien osv. For at kunne opnå en tilfredsstillende kontrol med alle disse parametre må processen monitoreres detaljeret.

I denne klumme vil vi se på hvordan PCA kan hjælpe med til at give et overblik over en forårsrulle friteringsproces [1]. Ved denne proces friteres forårsruller i en blanding af raps- og palmeolie i en semikontinuerlig proces over 4 uger. Det totale volumen for friteringsprocessen er 3000 liter og der tilføres dagligt 40 liter ny olie for at kompensere for tabet af olie i forårsrullerne. Hver nat overføres olien til en hviletank. Om morgenen overføres olien tilbage til friteringskarret og opvarmes til friteringstemperaturen, hvorefter der udtages en prøve. I alt 20 prøver (der friteres ikke i weekenderne) udtages fra

fritureolie-processen. Fem prøver i uge 1, weekend, fem prøver i uge to, weekend, osv.

For at kontrollere processen benytter fabrikanten en række standard kvalitetsanalyser: viskositet (VISC), % di- og polymere triglycerider (DPTG), % triglycerider (TG), % frie fedtsyrer (FFA), anisidantal (AV), iodtal (IV) og total vitamin E (vita-E). For at komplementere disse kvalitetsparametre har vi desuden analyseret de samme prøver med visuel og nærinfrarød spektroskopi i en 10 mm transmissionskuvette. Til dette formål har vi benyttet NIRSystems 6500, der skanner prøverne i området 400 nm til 2498 nm med 2 nm's intervaller; absorbanser ved i alt 1050 spektrale variable registreres for hver prøve.

I figur 1 og 2 er de kemiske og spektroskopiske data vist. De kemiske målinger er max-normeret til værdien én og viser forlø-



Figur 1. Kemiske målinger som funktion af antal måledage (1 til 20). Betydningen af forkortelser er angivet i teksten.

bet som funktion af antal dage for hver måling. Fælles træk kan aflæses af kurven direkte, men forløbet er helt klassisk. Mængden af FFA og viskositeten stiger, mens TG og IV falder. Kun anisidintallet som repræsenterer sekundære oxidationsprodukter har et lidt mere komplekst forløb. For at få et samlet overblik over kvalitetsmålingerne laves en PCA på disse standard analyse data. Derved er det muligt at se hvor mange underliggende fænomener, der er afspejlet i data og samtidig forstå dynamikken og sammenhængene mellem de forskellige variable.

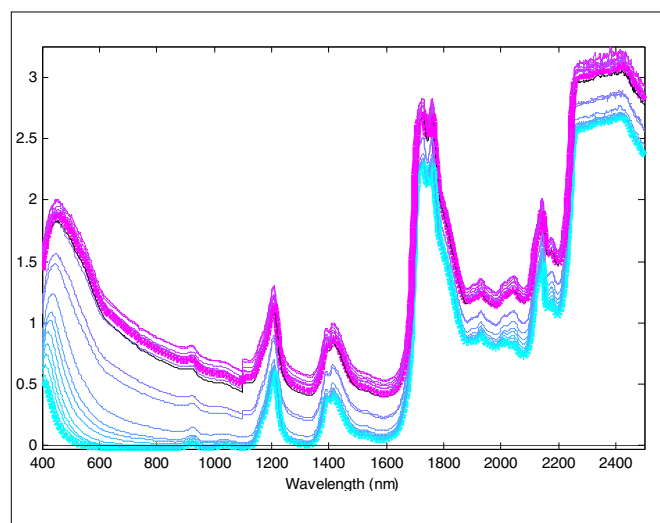
PCA på kemiske data

PCA modellen for de kemiske data kan skrives

$$X_{auto} = TP' + E$$

hvor X og E har dimensionen 20×8 . Data autoskaleres; det vil sige at fra hver variabel fjernes gennemsnittet og variablen skaleres, så alle variable har samme variation. Dermed vil modellen forsøge at afspejle alle variable frem for at fokusere på de, der er målt i størst numeriske værdier. Da dette er en eksplorativ analyse, fokuseres ikke på en eksakt bestemmelse af det korrekte antal komponenter. I stedet medtages rigeligt med komponenter, og den efterfølgende visualisering fokuserer så i første omgang på de første komponenter, som per definition er de, der beskriver mest af data.

I figur 3 ses bi-plottet for komponent et (PC1) mod komponent to (PC2). Af bi-plottet kan udledes at variationen mellem olieprøver bliver mindre hen over processen, og at de efter de første to uger havner i en ligevægtsposition karakteriseret ved mindre og tilfældige variationer (formentlig i ligevægt med den anvendte turnover). Plottet antyder et mindre spring til ligevægtspositionen efter prøve 11 som er mandagsprøven fra uge 3 og dermed at noget er sket med olien under opbevaringen weekenden over. Vi bemærker også det lille sving opad for de to prøver 3 og 4, hvilket skyldes at anisidintallet er højest på dette tidspunkt i friteringsprocessen. Det bringer os videre

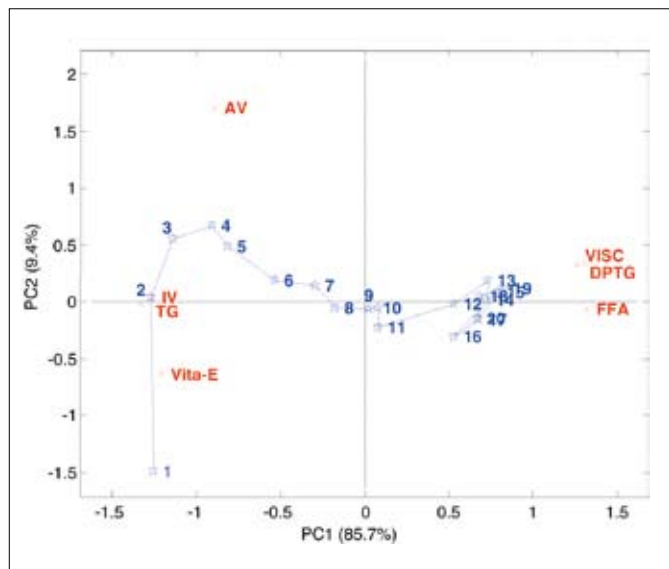


Figur 2. NIR spektre målt direkte på fritureolierne. Spektrene er farvet efter måledag (fra den nye planteolie ved måledag 1 (cyan) til den brugte fritureolie ved måledag 20 (magenta)).

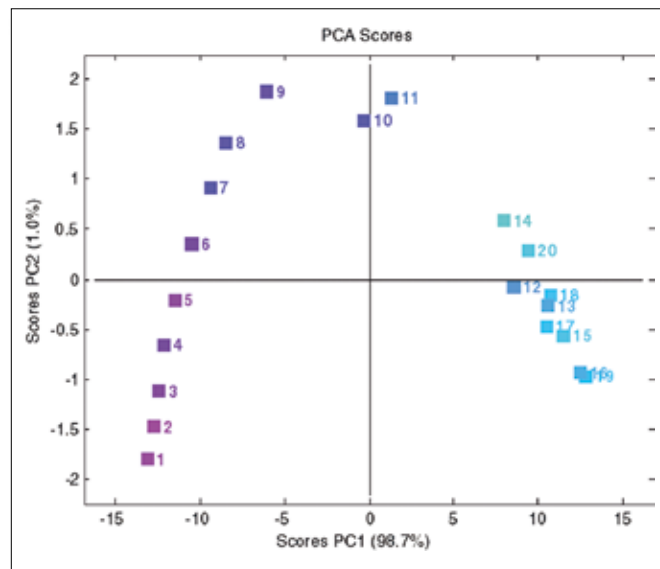
til loadings, der viser at komponent to især er spændt ud af AV. Praktisk taget alle de øvrige kvalitetsparametre er stærkt interkorrelerede: når TG falder, så falder IV og videre når TG og IV falder, så stiger VISC, FFA, DPTG. Det er af forskellige årsager ganske normalt at man indenfor fødevarerindustrien benytter denne slags overbestemmelse, men i lyset af PCA'en bør det diskuteres, om det ikke vil være fordelagtigt fx at øge frekvensen på TG-målingen på bekostning af de øvrige kvalitetsparametre (som jo fundamentalt variationsmæssigt giver samme information om lige nøjagtig denne proces).

PCA på spektrale data

Et hurtigt kig på de kemiske data (figur 1) giver forholdsvis meget information om den dynamiske variation i processen. De spektrale data på den anden side, er langt vanskeligere at fortolke umiddelbart (figur 2). Imidlertid er der potentielt langt



Figur 3. Bi-plot fra en PCA på de kemiske data. Udviklingen over tid fremgår klart (1 svarer til mandag i uge 1, 20 svarer til fredag i uge 4).



Figur 4. PCA scoreplot for NIR spektre. Prøverne er farvelagt efter iodtal, der varierer fra 33 til 73 (magenta er høj, cyan er lav).

mere information at hente i de spektrale data, og en vigtig pointe er, at vi kan visualisere og kvantificere denne information fuldstændig analogt til PCA modellen på de kemiske data. Vi har altså et simpelt værktøj til også at visualisere og forstå mere komplekse data.

For de spektrale data er PCA modellen

$$X_{mean} = TP' + E$$

hvor X og E har dimensionen 20×1050 og data er centre-rede. T har dimensionen $20 \times \text{antal PC}$, og P har dimensionen $1050 \times \text{antal PC}$. Antallet af komponenter der beregnes skal igen være rigeligt til den første eksplorative undersøgelse, f.eks. 10.

I figur 2 ses de rå NIR-data, hvor den tykke blå linje viser spektret af den nye ubrugte fritureolie, der blot er blevet varmet op og den tykke røde linje viser den olie, der kasseres efter fire ugers fritering. Det bemærkes at den røde linie ikke repræsenterer et ekstremum, men at det er tydeligt, at se at olien er blevet mere farvet (400 – 800 nm), og at der hen over friteringen er kommet en del opslemmet materiale, der spreder lyset (stigende offset).

I figur 4 ses tidsudviklingen i PCA score-plottet for NIR-data. Det ses, at udviklingen understøtter udviklingen i de kemiske data, men der er forskelle, der er værd at notere: dynamikken i de to forløb er forskellig. Ligesom for de traditionelle kvalitetsparametre øges variationen hen over processen og ligesom før sker der et »hop« til en ligevægtsposition efter mandagsprøven i uge 3: 11. Det bemærkes dog, at NIR data viser et mere roligt procesforløb, hvilket skyldes NIR-spektroskopis fortræffelige reproducerbarhed samt at »hoppet« mellem prøve 11 og ligevægtspositionen er blevet meget stort, således at der med NIR-brillerne på ikke er den mindste tvivl om et processkift.

Hvis man kun havde monitoreret processen med en enkelt kvalitetsparameter som IV eller TG, så ville man formentlig ikke have observeret dette skift (se figur 1). Og dette er præcis kernen i erkendelsen af den amerikanske Food and Drug Administration når de i deres Process Analytical Technology (PAT) vejledning skriver ”Traditional one-factor-at-a-time experiments do not effectively address interactions between products and process variables”. Det er endda muligt, at processkiftet ville være blevet overset selv med det oligovariate PCA med flere standard kvalitetsparametre. Først med kombinationen af den multivariate spektroskopi og kemometri fremgår processkiftet tydeligt.

Outro

Den generelle strategi når, der er målt mange forskellige datasæt på det samme sæt prøver, er først at beregne PCA modeller på hvert datasæt for sig. Efter disse er inspiceret og tolket individuelt, kan de sammenlignes kvalitativt som illustreret for fritureolie data. Næste skridt er at udvikle en kvantitativ model mellem datasættene. Kemometriske metoder velegnet til dette vil vi se nærmere på i de kommende klummer.

Kilder

[1] SB Engelsen. Explorative Spectrometric Evaluations of Frying Oil Deterioration. JAOCS, 74(12), 1495-1508, 1997.

Fra træflis til biobrændstof

Forskere fra Biomass Gasification Group har produceret det miljøvenlige brændstof bio-metanol på basis af træ. Deres næste skridt er at lave bio-DME - dimethylether. På sigt regner forskerne med at de kan lave metanol og DME på planteaffald som halm, risskaller og andre besværlige biobrændsler.

- I sig selv er DME et godt diesel-brændstof. Den har en god CO_2 økonomi og afgiver meget få miljøfarlige partikler under forbrænding. Når stoffet så laves på basis af termisk forgasning af biomasse, bliver det CO_2 neutralt. Samtidig er DME nemt at transportere, fordi det opfører sig som flaskegas, siger adjunkt Jesper Ahrenfeldt ved DTU Mekanik.

Dimethylether er allerede en del af vores hverdag som drivgas i spraydåser, men man kan ikke bare hælde DME på en almindelig dieselmotor, for det kan brændstofpumpen ikke holde til. Stoffet anvendes dog på det mere eksperimentelle plan som brændstof i for eksempel DTU's økobil Dynamo.

Uspiselig biomasse

- Vi kan få meget energi ud af jordens uspiselige biomasse ved produktionen af bio-methanol og bio-DME. Samtidig kan vi også udnytte en energirest på 30 procent fra produktionen af bio-ethanol. Vi regner med, at vi med held kan kombinere produktionen af bio-methanol og bio-ethanol, og vi har ansøgt om midler til formålet sammen med forskere fra DTU Biosys., forklarer Jesper Ahrenfeldt.

Omdannelsen af syntese-gas til forskellige brændstoffer er velkendte processer, som allerede foregår i industrien. Derfor har det været naturligt at inddrage DONG Energy og Haldor Topsøe A/S i projektet, som er finansieret af Energinet.dk.

Forskningen i forgasning har hidtil først og fremmest været rettet mod el- og varmeproduktion, og gruppen er i fuld gang med at commercialisere teknologien sammen med virksomheden Weiss A/S, der leverer udstyr til fjernvarmesektoren. I en årrække har forskergruppen kørt et lille kraftvarmeverk på DTU i Lyngby og Weiss har nu en opdateret og opskalaret version af anlægget i Hadsund.

Kilde: DTU

**SVINGMØLLE
MM 400**

Allround mølle til de mindre prøver, bl.a. DNA/RNA og XRF analyser. Tilbereder op til 20 prøver samtidigt.

SKANLAB

Tlf: 4738 1014 - www.skanlab.com

Vind en Mercedes sportsvogn med RETSCH, læs mere på www.retsch.com/grandprix