

HOVEDOPPGAVE VÅREN 2002

**UTVIKLING AV ESTIMATOR FOR
PRODUKTSAMMENSETNING I EN
DESTILLASJONSKOLONNE**

Utført av Sigurd Myhre Hellem

Faglærer: Professor Morten Hovd



Norges Teknisk- Naturvitenskaplige Universitet

Fakultet for Naturvitenskap og Teknologi

Institutt for Kjemisk Prosessteknologi

Sammendrag

Oleon Scandinavia AS destillerer i sitt anlegg i Sandefjord fettsyrer fra ulike råvarer, som fisk og kokosoljer. Ved destillasjon av den behandlede kokosoljen, som dannet et flerkomponentsystem, ble det stilt høye krav til renhet av enkelte produkter, samtidig med et ønske om minimaliserte tap av de mest verdifulle komponentene. En mer optimal regulering av sammensetningen dannet dermed utgangspunktet for denne oppgaven.

I denne hovedoppgaven ble det utviklet PLS estimatorer for produktsammensetningene i en destillasjonskolonne. Produktstrømmen over topp skulle tilfredsstille spesifikasjonene ved 95 % og 99 % (masse) laurinsyre ($C_{12}H_{24}O_2$). Estimatorene viste å kunne gi gode estimater for sammensetningen eller innholdet av laurinsyre innenfor de rammene som ble vurdert i oppgaven. Avhengigheten av de sterkt forsinkede GCL (gass-væskekromatografi) analysene, som tok om lag 45 minutter å gjennomføre, kan dermed reduseres ved å anvende de utviklede estimatorene.

Estimatorene ble utviklet ved å kun benytte data fra statiske simuleringer av destillasjonen i Hysys™. Destillasjonen ble foretatt under vakuumbetingelser i en kolonne med strukturerte pakninger. Trykket i kolonnen varierte både i toppen av kolonnen, som ble regulert av vakuumsystemet, og ved det totale trykkfallet over kolonnen. Modelleringen av kolonnen ble gjort ved en platekolonne, og resultatene fra simuleringene stemte overens med tilgjengelige trend- og analyseverdier fra prosessen.

I oppgaven ble det vist at estimatene ble betydelig forbedret ved å inkludere trykkmålinger i tillegg til temperaturmålinger i estimatoren. Den forbedrede estimeringsevnen kom hovedsakelig av den store trykkvariasjonen som forelå i prosessen. Temperatur ble målt seks steder i kolonnen, men kun fire av disse ble anbefalt brukt i estimatorene. Estimatorene som ble foreslåtte inneholdt målinger av trykket i topp og bunn, og de fire temperaturmålingene som var plassert over fødepunktet. Det optimale antallet faktorer i PLS estimatoren ble vurdert å være fire.

En endelig verifisering av resultatene ved å anvende estimatoren i anlegget ble ikke foretatt. et Det kan dermed ikke trekkes en endelig slutning om estimatoren også vil være tilstrekkelig til å bli anvendt ved regulering av produktsammensetningen til destillatet.

En reguleringssløyfe for kolonne var p.t lukket. Reguleringen av nivå i refluksakkumulatoren viste instabilitet, hvor manipuleringen av refluksstrømmen til kolonnen fører til oscillasjoner i denne strømmen. Tuning av nivåregulatoren for å hindre instabilitet blir derfor anbefalt å gjennomføre. Lukking av åpne sløyfer for destillasjonskolonnen blir deretter anbefalt, hvor en verifisert estimator eventuelt kan benyttes.

Innholdsfortegnelse

1	INNLEDNING	1
1.1	BAKGRUNN.....	1
1.2	OPPGAVEBESKRIVELSE.....	2
2	INDIREKTE SAMMENSETNINGSREGULERING	3
2.1	REGULERING MED EN TEMPERATUR.....	3
2.1.1	<i>Trykk kompenserte temperatur</i>	3
2.1.2	<i>Problemer forbundet med regulering av en temperatur</i>	4
2.1.3	<i>Plassering av målepunkt</i>	4
2.2	REGULERING AV SAMMENSETNING MED FLERE TEMPERATURMÅLINGER.....	5
2.2.1	<i>Enkel flertemperaturregulering</i>	5
2.2.2	<i>Modellbasert flertemperaturregulering</i>	6
3	MULTIVARIAT KALIBRERING OG ANVENDELSE INNEN REGULERING	7
3.1	ET ESTIMERINGSPROBLEM.....	7
3.1.1	<i>Estimator for en binær blanding</i>	7
3.1.2	<i>Transformerte variable</i>	8
3.2	VALG AV MÅLEPUNKTER.....	9
3.3	ORDINÆR MINSTE KVADRATERS (OLS) METODE.....	10
3.4	REGRESJONSMETODER FOR LATENTE VARIABLE (LVR).....	11
3.4.1	<i>Latente variable</i>	11
3.4.2	<i>Singulærverdidekomponering (SVD)</i>	12
3.4.3	<i>Prinsippal komponent regresjon (PCR)</i>	12
3.4.4	<i>Prinsippal minste kvadraters metode (PLS)</i>	13
3.4.4.1	Algoritme for ortogonaliserte PLS for en y-variabel [Ref XII].....	13
3.4.4.2	Antall PLS faktorer i modellen.....	15
3.4.4.3	Full prediksjon.....	15
3.4.4.4	Kort prediksjon.....	16
3.4.5	<i>Vekting av variable</i>	16
3.4.6	<i>Sammenlikning av PCR og PLS</i>	17
4	PROSESSBESKRIVELSE	18
4.1	FETTSYRENE.....	18
4.1.1	<i>Fettsyrer i kokosolje</i>	18
4.1.2	<i>Destillasjon av fettsyrer</i>	19
4.2	OVERSIKT OVER PROSESSEN.....	20
4.2.1	<i>Prosessanlegget</i>	20
4.2.2	<i>Prosessparametere</i>	22
4.3	PROSESSMODELLEN.....	23
4.3.1	<i>Prosessmodellen i Hysys</i>	23
4.3.2	<i>Termodynamikk</i>	24
4.3.3	<i>Antall teoretiske trinn</i>	24
5	ESTIMATOR FOR PRODUKTSAMMENSETNINGEN	26
5.1	FORBEREDELSE.....	26
5.1.1	<i>Valg av kalibreringssett</i>	26
5.1.2	<i>Forhold som ble vurdert</i>	27
5.1.3	<i>Evalueringskriterier</i>	28
5.1.4	<i>Støynivået i målingene</i>	29

5.1.5	<i>Simuleringene</i>	29
5.1.6	<i>Plassering av temperaturmålingene i kolonnen</i>	31
5.2	ESTIMATOR FOR PRODUKTSAMMENSETNINGEN	32
5.2.1	<i>Notasjoner for estimeringsprobleme</i>	32
5.2.2	<i>Produktstrøm med 99 % C₁₂</i>	34
5.2.2.1	Modell uten trykk variasjon.....	34
5.2.2.2	Estimator med temperaturmålinger	34
5.2.2.3	Estimator med temperatur og trykk målinger uten støy (n ₀).....	35
5.2.2.4	Estimator med temperatur og trykk målinger med støy (n ₁₋₆).....	38
5.2.3	<i>Variasjon i antall teoretisk trinn i kolonnen (n₀)</i>	41
5.2.4	<i>Estimator for 95 % renhet av C₁₂ topp-produktet</i>	42
5.2.5	<i>Følsomhet for C₁₀ i fødestrømmen til kolonne 2</i>	43
5.3	ESTIMATORENE FOR 99 % OG 95 % C ₁₂ I PRODUKTSTRØMMEN	44
5.3.1	<i>Estimatoren for 99 % C₁₂ i produktstrømmen</i>	44
5.3.2	<i>Estimatoren for 95 % C₁₂ i produktstrømmen</i>	45
5.4	TEST AV ESTIMERINGSEVNEN TIL DEN BESTE ESTIMATOREN PÅ HISTORISKE DATA.....	46
6	PROSESSBETRAKTNING	50
6.1	REGULERINGSSTRUKTUR FOR KOLONNE 2	50
6.2	BETRAKTNINGER AV REGULERINGEN.....	51
7	DISKUSJON	52
7.1	MODELL AV KOLONNEN I HYSYS	52
7.1.1	<i>Antall komponenter i fødestrømmen til kolonnen</i>	52
7.1.2	<i>Kilder til modellfeil i Hysys</i>	52
7.1.2.1	Valg av termodynamikk	52
7.1.2.2	Antall teoretiske trinn i destillasjonskolonnen.....	53
7.1.2.3	Trykket i kolonnen	53
7.1.2.4	Plassering av temperaturmålingene	54
7.1.3	<i>Tilgjengelig informasjon om prosessen</i>	54
7.1.4	<i>Neglisjerte effekter</i>	55
7.2	ESTIMATOR FOR PRODUKTSAMMENSETNINGEN	55
7.2.1	<i>Modell uten trykk variasjon</i>	55
7.2.2	<i>Modell med trykk variasjon</i>	55
7.2.2.1	Estimator med temperaturobservasjoner	55
7.2.2.2	Estimator med temperatur- og trykkobservasjoner	56
7.3	SANNSYNLIG PRESTASJONSEVNE I PROSESSEN	58
7.4	VIDERE ARBEID.....	58
8	KONKLUSJON	60
9	SYMBOLLISTE	61
	REFERANSER	63
	OVERSIKT OVER BILAGENE	65

1 Innledning

1.1 Bakgrunn

Destillasjon er en av de viktigste enhetsoperasjonene i prosessindustrien. I mange anlegg påløper opp mot 30-40% av de totale investeringskostnadene og energikostnadene denne enhet. Endringer og forbedringer i den operative driften av en destillasjonskolonne kan gi et betydelig utslag på de totale kostnadene. Separasjon av blandinger i destillasjonskolonner krever energi, og den tilførte mengde energi for en kolonne øker eksponentielt med separasjonsgraden. Dårlig regulering av produktsammensetningene til destillasjonskolonnen vil gi et høyere energiforbruk eller energitap, grunnet sannsynligheten for lengre driftsperioder med høyere renhet enn påkrevd. En tettere regulering av sammensetningen vil kunne gi et lavere energiforbruk og dermed lavere driftskostnader.

Måling av produktsammensetningene er forbundet med nøkkelordene pålitelighet, nøyaktighet og tilgjengelighet. De fleste analysatorer, slik som analyse med gasskromatografi (GC), gir sterkt forsinkede målinger. Forsinkelsen ligger typisk i intervallet 10 til 20 minutter, men kan også være opptil 45 minutter. Tilgjengeligheten reduserer derfor den oppnåelige ytelsen for prosessregulering basert på analyse av produktsammensetningen. Analysatorene er i tillegg komplekse instrumenter, og er derfor arbeids- og ressurskrevende.

Problemene forbundet med analysatorene har ledet til anvendelse av alternative målinger og metoder, hvilket blir brukt til indirekte å regulere produktsammensetningen med større nøyaktighet. Det vanligste alternative til analysatorene er regulering med en temperatur i kolonnen i følge Buckley et. al.[Ref I]. Dette innebærer regulering med temperaturen på ett gitt sted i kolonnen, hvor temperaturen da implisitt gir sammensetningen av produktstrømmene. Temperaturmålinger er karakterisert ved å være pålitelige og rimelige, og ha en neglisjerbar måleforsinkelse, men temperatur er ikke en nøyaktig indikator for produktsammensetningen. For et system med konstant trykk og med flere enn to komponenter vil ikke en temperatur indikere eller gi en unik sammensetning for blandingen. Flere sammensetninger kan gi den samme kokepunktstemperaturen.

Problemene forbundet med regulering med en-temperatur har ledet til anvendelse av flere temperaturmålinger i kolonnen. Alternative metoder for anvendelse av flere temperaturer for indirekte sammensetningsregulering har blitt foreslått av blant annet Luyben [Ref I], regulering med temperaturprofil, og Mejdell [Ref II], regulering med regresjonsestimator for sammensetningen. Mejdell og Skogestad [Ref III] har anvendt en statistisk partiell minste kvadraters (PLS) regresjonsestimator basert på simuleringer av prosessen, og implementerte denne estimatoren i en destillasjonskolonne med gode resultater. Mejdell og Andersson [Ref IV] anvendte historiske data da de implementerte en statistisk PLS estimator på en kolonne for etanfjerning.

1.2 Oppgavebeskrivelse

Oleon Scandinavia AS destillerer i sitt anlegg i Sandefjord fettsyrer fra ulike råvarer, som spenner fra fiskeprodukter til kokosoljer. Destillasjon av den behandlede kokosoljen er det som blir inngående vurdert i denne oppgaven. Oleon har høye krav til renhet av enkelte produkter, samtidig som det er ønskelig å minimalisere tap av de mest verdifulle komponentene. Denne oppgaven tar for seg sammensetningen i topp-produktet for en av kolonnene, hvor det er ønskelig med høy renhet (95 - 99 masse %) av laurinsyre ($C_{12}H_{24}O_2$). Det blir beregnet sammensetningsestimatorer ved PLS regresjon, som er basert på temperatur- og trykkmålinger. Estimatorene er ment å kunne erstatte eller redusere avhengigheten av nåværende sammensetningsmålinger. Dagens målinger av sammensetningen blir gjort ved GCL (gass-væskrokromatografi), hvor en analyse tar 45 minutter. De analyserte sammensetningene blir benyttet til å manuelt regulere destillatuttaket for kolonnen. En estimator vil gi et online estimat for sammensetningen og kan benyttes til direkte eller manuell regulering av sammensetningen. I denne hovedoppgaven blir statistiske simuleringer av kolonnen benyttet til å dekke typiske variasjoner i alle inngangsvariablene og prosessforstyrrelsene. Regresjonsestimatorene blir testet med forskjellige forstyrrelser og en enkelte anbefalinger blir gitt. Reguleringsstrukturen p.t. i anlegget og eventuelle forbedringer av reguleringen blir også kort vurdert i denne oppgave.

2 Indirekte sammensetningsregulering

2.1 Regulering med en temperatur

Temperaturen som en indikator for sammensetningen er basert på antagelse om termodynamisk likevekt mellom væske- og gassfasen innenfor trinnene i destillasjonskolonnen. Ved et gitt konstant trykk, P , er det ved likevekt for en binær blanding en enkel væske sammensetning, x , for som gir den korresponderende kokepunktstemperaturen, T .

$$T_b = f(x, P) \quad (2.1)$$

For et system med flere enn to komponenter vil ikke en temperatur indikere eller gi en unik sammensetning for blandingen, fordi flere sammensetninger kan gi den samme kokepunktstemperaturen.

2.1.1 Trykk kompenserte temperatur

Trykket i en kolonne vil kunne variere med forstyrrelser som virker på systemet og dermed påvirke kokepunktstemperaturen. Trykk endringene kan være et resultat av endret settpunkt, unøyaktigheter i reguleringen eller målingene av trykket i kolonnen, eller forandret trykktapet i kolonnen forårsaket av endret massestrøm for føden til kolonnen. Temperaturmålingene kan matematisk kompenseres for variasjonen i trykket. Den målte temperaturen, T , blir da referert til det varierende målte trykket, P , og kompensert ved ett referansetrykk, P_b . En lineær kompensasjon for trykket blir gitt på følgende form for en binær blanding:

$$T_b = T - \left. \frac{\partial T}{\partial P} \right|_x (p - p_b) \quad (2.2)$$

Den partiell deriverte er den inverse kurve for trykket i gassfasen ved normal produkt sammensetning. Antoine likningen gir en sammenheng mellom trykket og temperaturen for rene komponenter, med antatt ideelt system og gyldighet av Raoult's og Dalton's lover kan trykkkompenserte temperaturer bli estimert ved

$$\text{Antoine likningen: } \log(P) = A - \frac{B}{T + C} \quad (2.3)$$

$$\text{Raoult's lov: } P_i = x_i P_i^*(T) \quad (2.4)$$

$$\text{Dalton's lov: } P = \sum_i P_i^*(T) \text{ og } P_i \equiv y_i P \quad (2.5)$$

hvor A , B og C er parametere som er tilpasset fra eksperimentelle data, T er temperatur, P_i er partialtrykket og y_i er molfraksjon av komponent i .

2.1.2 Problemer forbundet med regulering av en temperatur

Typiske problemer som er forbundet med temperaturmålingene vil være støy og bias i målingene, hvilket også inkluderer unøyaktighet i trykkmålingene og dermed eventuelt de trykk-kompenserte temperaturene. Andre forhold som vil føre til temperaturvariasjon er pulser i massestrømmen i kolonnen og utilstrekkelig blanding på trinnene. Slutlig vil temperaturavvik forekomme hvis likevekt ikke blir oppnådd på trinnene.

Forholdet mellom sensitivitet og støy for et gitt målepunkt, vil være en kritisk faktor. Hvis en endret sammensetning resulterer i en liten temperaturendring sammenliknet med trykkfluktuasjonen og måleusikkerheten, vil regulering av sammensetningen ved anvendelse av en temperatur være vanskelig. Temperaturvariasjonen vil være liten ved høyere renhet av produktene i topp- og/eller bunnstrømmene i destillasjonskolonne. Temperatur endringen vil være større nærmere midten av kolonne, og målepunktene i kolonnen blir derfor oftest flyttet til deler av kolonnen hvor sensitiviteten er høyere. Motvekten for en plassering av målepunktet langt fra toppen/bunnen av kolonnen er en svakere korrelasjon mellom temperaturmålingene og produktsammensetningen. Effekten av endret sammensetning i den motsatte enden og følsomhet for føden vil også øke. Direkte kontroll av en temperatur nære midten av kolonnen på settpunktsnivået vil derfor ikke kunne holde produktsammensetningen på ønsket nivå. Modellbasert prediktiv regulering (MPC) av prosessen basert på temperaturmåling og prediksjon av forstyrrelsen i prosessen ville kunne redusert problemene forbundet med regulering av en temperatur. MPC for regulering av fettsyredestillasjon har blitt forsøkt av Ohshima et. al. [Ref V].

2.1.3 Plassering av målepunkt

De viktigste kriteriene, i følge Shinsky [Ref VI], for plassering av målepunktet er den beste kombinasjonen av sensitivitet, representativitet for produktsammensetningen, uavhengighet for endret sammensetning og refluksentalpi, og følsomhet for dynamiske endringer. I et flerkomponent system bør ikke punktet være følsomheten innflytelsen av ikke-nøkkel komponenter.

Andre kriterier som blir ved plassering av målepunkt er vurdering av de dynamiske og lineære egenskapene til reguleringsløyfen, og minimum påvirkning mellom sløyfene. Mejdell hevder at de nevnte kriteriene vanligvis er av mindre betydning, da de normalt blir håndtert av reguleringsystemet.

Destillasjon av flerkomponentsystemer gjør temperaturregulering vanskeligere. I enden av kolonnen vil temperaturen bli dominert av separasjonen mellom ikke-nøkkel og nøkkel komponent. Effekten av varierende innhold av ikke-nøkkel komponenter vil være størst i trinnene rundt fødepunktet. Ofte finnes det et område i kolonnen hvor innholdet av ikke-nøkkelkomponenter er forholdsvis konstant. De tidligere nevnte problemene ved temperaturregulering i multikomponentdestillasjon kan dermed bli omgått eller redusert ved å velge et sted i kolonnen hvor nivået av nøkkel-komponentene er tilnærmet konstant

Det er foreslått flere metoder for utvelgelse av beste plassering av temperaturmålepunkt, og på grunnlag av de tidligere nevnte kriteriene kan en generell metode settes opp. Det første kriteriet favoriserer plassering av temperatursensoren bort fra produktstrømmen ved høy produktrenhet, mens det andre punktet vil implisere en plassering like ved produktstrømmen. Et endelig valg vil være en middelvei for plasseringen.

2.2 Regulering av sammensetning med flere temperaturmålinger

2.2.1 Enkel flertemperaturregulering

Regulering med en enkelt temperatur gir som nevnt i avsnitt 2.1.2 generelt sett ikke nøyaktige resultater. Endringer i sammensetning, innhold av ikke-nøkkelkomponenter, trykkendringer, pulser i massestrømmen og utilstrekkelig blanding på trinnene vil alle påvirke temperaturen. I tillegg vil regulering av sammensetningen basert på en enkelt måling være mer følsom for støy, enn hvis flere målinger ble anvendt.

Noen av problemene ovenfor kan bli motvirket ved å inkludere flere målinger. Høy frekvent støy kan fjernes ved å filtrere signalene. Effekten av trykkvariasjon kan kompenseres ved å anvende trykkmålinger eller ved å bruke differensielle temperaturer. Innflytelsen av ikke-nøkkel komponenter kan reduseres ved å plassere temperaturmålingene i områder av kolonnen hvor innholdet av komponentene er nær konstant. Regulering av en trinntemperatur på en konstant verdi vil ikke være tilstrekkelig til å holde produktsammensetningen på et gitt nivå. En løsning på dette problemet er å inkludere flere temperaturmålinger for bedre indikasjon på produktets sammensetning.

Differensielle temperaturer kan kompensere for effekten av trykk variasjon, men ikke for endret trykkfall i kolonnen. Trykket har omtrent samme virkning på alle temperaturer i kolonnen ved total trykk endring, og en ekstra temperatur måling på et trinn med nær konstant sammensetning vil da kunne følge trykkvariasjonen. Temperaturdifferansen vil være et selektivt mål for sammensetningsendringer.

En ekstra differensiert tempertur, hvilket gir *dobbelt differensierte* temperaturer har blitt foreslått for å håndtere endrede trykkfall i kolonner. Industriell anvendelse av metoden er rapportert å være vellykket. Dobbelt differensierte temperaturer har også blitt foreslått å kunne spore endringer i fødesammensetningen for flerkomponentsystemer. I den sammenhengen var en differensiell temperatur selektiv sensitiv mhp nøkkel-komponenten, og den andre var sensitiv for den viktigste ikke-nøkkel komponenten. Temperaturdifferansene ble så brukt til å korrigere settpunktet for to vanlige regulatorer.

Regulering ved *temperaturprofil* i kolonnen har blitt foreslått av Luyben [Ref I] for kolonner med skarpe temperaturprofiler, hvilket vil si høy relativ flyktighet og høy produktrenhet. Profilets midtpunkt vil forflytte seg opp og ned som følge av endret produktsammensetning, da masseoverføringen hovedsakelig vil foregå i en kort seksjon av kolonnen. Summasjon av målte temperaturer på ulike trinn i kolonnen vil indikere senteret for profilen, og ved kontroll av produktsammensetningen skal regulatoren holde senteret på et angitt trinn i kolonnen.

2.2.2 Modellbasert flertemperaturregulering

En mer rigorøs metode for regulering av produktsammensetningen i en destillasjonskolonne er i følge Mejdell [Ref II] en temperaturbasert estimator for prediksjon av sammensetningen. Mejdell vurderte i sin avhandling flere estimatorene, et dynamisk Kalman filter, en statisk Brosilow Inferentiell Estimator og en statisk Prinsippal Komponent Regresjon Estimator (PCR). En statisk Partiell minste kvadrat estimator (PLS) ble også vurdert mot resultatene fra de andre estimatorene. Mejdell fant at den statiske PCR estimatoren gav resultater som var sammenliknbare med resultatene med det dynamiske Kalman-filteret for destillasjonskolonnen i lukket sløyfe. Dette skyldtes at temperaturene og sammensetningen har liknende dynamiske respons. PLS estimatoren gav tilsvarende eller noe bedre resultater enn PCR estimatoren, men er en noe mer kompleks estimator.

Den statiske Brosilow estimatoren er basert på en åpen sløyfe modell og anvender temperaturmålinger i tillegg til målinger av refluksstrøm (L) og mengde oppkok (V) til å estimere forstyrrelsene fødestrøm (F) og molfraksjonen av lett komponent i føden (z_F). De estimerte forstyrrelsene anvendes i beregning av produktsammensetning. Estimatoren ble funnet å være veldig følsom for støy i inngangsmålingene og modell feil for dårlig kondisjonerte systemer eller systemer med høy renhet i produktene.

Kalman-filteret som gav de beste resultatene, men krever mye mer arbeid ved modelleringen og tuning av modellen. Mejdell anbefaler derfor de statiske regresjonsestimatorene, som enklere kan gi en modell eller estimator for sammensetningen på grunnlag et riktig valgt kalibreringssett med temperaturer og sammensetninger. Kalibreringssettet kan enten være simuleringresultater fra en modell av anlegget, historiske data fra anlegget med målte temperatur, analyserte sammensetninger og eventuelt supplerende målinger, eller en kombinasjon av de to nevnte. Mejdell og Andersson [Ref VII] anvendte historiske data da de implementerte en statisk PLS estimator på en kolonne for etanfjerning. Kano et. al. [Ref VIII] anvendte en dynamisk PLS estimator ved å bruke simulerte tidsseriedata under regresjonen. Den dynamiske estimatoren inneholdt målinger av temperaturen på det gitte og utvalgte tidligere tidspunkt og i tillegg målinger av refluksstrømmen, effekt tilført koker og trykket i kokeren. Den dynamiske modellen ble rapportert å fungere bedre enn den statiske modellen. Den lineære PLS regresjonen har også blitt foreslått av Liu et. al. [Ref IX] forbedret ved ulineær PLS basert på neuralnettverk metodikk.

3 Multivariat kalibrering og anvendelse innen regulering

Kalibrering er av Martens og Næs [Ref X] definert som ”anvendelse av empiriske data og forhåndskunnskap til kvantitativt estimering av informasjon Y fra tilgjengelige målinger X , gjennom en matematisk transfer funksjon”. Multivariat kalibrering innebærer dermed fastsettelse av hvordan en skal bruke de målte variablene $x_1, x_2 \dots x_n$ samtidig i en kvantifisering av den søkte verdien av y .

3.1 Et estimeringsproblem

3.1.1 Estimator for en binær blanding

Problem blir definert som en binær blanding ved konstant trykk, konstant trinneffektivitet, og føde og refluks som mettet væske. En spesifisering av sammensetningen for føden, destillatet og bunnproduktet vil gi en unik stasjonær tilstandsprofil for temperaturene i kolonnen.

$$\theta^{i,M \times 1} = f(z_F, y_D, x_B) \quad (3.1)$$

hvor i, M er antall temperaturmålinger θ i kolonnen. Den inverse av uttrykket blir søkt, og er gitt ved den følgende sammenhengen for den søkte destillatsammensetningen

$$\hat{y} = \hat{y}_D = g(\theta^{i,M \times 1}) \quad (3.2)$$

En forenklet uttrykk for estimatoren vil være

$$\hat{y}_D = K\theta \quad (3.3)$$

K vil ha en transponert dimensjonen sammenliknet med θ hvis antall avhengig y variable skal estimeres er en. Den statiske estimatoren som blir beregnet kan også inneholde andre uavhengige variable enn temperatur. Mejdell [Ref II] har i sin avhandling funnet at ved å innføre informasjon og variable for forstyrrelser og innganger i tillegg til temperaturer, vil estimatoren presterer noe bedre ved stasjonære betingelser. Den dynamiske ytelsen blir derimot sterkt forringet ved å anvende andre inngangsvariable enn temperaturer. Mejdell angir at den reduserte ytelsen kommer av ulik dynamisk oppførsel for temperaturer og for eksempel direkte inngangsverdier for trykket i en estimator.

En utvidelse av den binære blandingen, bestående av kun nøkkel komponenter, med ikke-nøkkel komponenter, tunge og lette, vil danne en multikomponent blanding. Formålet med separasjonen vil være uendret, det vil si å splitte de to nøkkel-komponentene. En pseudobinær molfraksjon basert på nøkkel komponentene kan hvis ønskelig bli dannet ved følgende formel,

$$x' = \frac{x_L}{x_L + x_H} \quad (3.4)$$

hvor x' er den pseudobinære molfraksjonen i væskefasen på et gitt trinn i destillasjonskolonnen, x_L er molfraksjon av lett nøkkelkomponent og x_H er molfraksjon av tung nøkkelkomponent. x' erstatter den virkelige molfraksjonen i en multikomponent blanding.

3.1.2 Transformerte variable

Sammensetningen og temperaturprofilen for kolonnen er ulineære funksjoner av driftsbetingelsene. En metode for å hanske med den ulineære naturen, er å introdusere andre faktorer eller variable som kan fange den ulineære oppførselen. En enklere metode innebærer en ulineær transformasjon av hver enkelt variabel. En logaritmisk transformasjon av produktsammensetningen har blitt foreslått av flere artikkelforfattere [Ref XI], som en effektiv linearisering av den dynamiske og statiske responsen av de uavhengige variablene L , V , F , z_F etc. For en binær blanding vil uttrykket den transformerte for destillatet og eventuelt bunnproduktet være:

$$Y_D = \ln(1 - y_D) \text{ og } X_B = \ln(x_B) \quad (3.5)$$

Temperaturen er ofte nesten en lineær funksjon av sammensetningen. En linearisering av sammensetningsprofilen i kolonne er også en vanlig fremgangsmåte. Molfraksjonen på gitt trinn(i) i kolonnen.

$$X = \ln\left(\frac{x_{i,L}}{x_{i,H}}\right) \quad (3.6)$$

Temperaturen kan med basis i ovenstående likning transformeres for å linearisere temperaturresponsen og temperaturprofilen. Utlegning av uttrykket finnes i Mejdells doktoravhandling [Ref II].

$$L_T = \ln\left(\frac{\theta - T_L^b}{T_H^b - \theta}\right) \quad (3.7)$$

Trinntemperaturen er θ , T^b er kokepunktstemperaturen for rene henholdsvis tung og lett nøkkelkomponent. Denne lineariseringen har vist å kunne gi en nesten lineær temperaturprofil ved noen høy renhets destillasjoner. Istedenfor kokepunktstemperaturer for rene komponenter, kan en anvende en referansetemperatur (vanligvis en målt temperatur) nær toppen eller bunnen av kolonnen. Temperaturen i endene av kolonnen vil for en binær blanding være veldig nær kokepunktstemperaturen for ren komponent under forutsetning om høy renhet av produktene ($y_D \approx 1$ og $x_B \approx 1$). Det vil si at likningen over kan uttrykkes ved

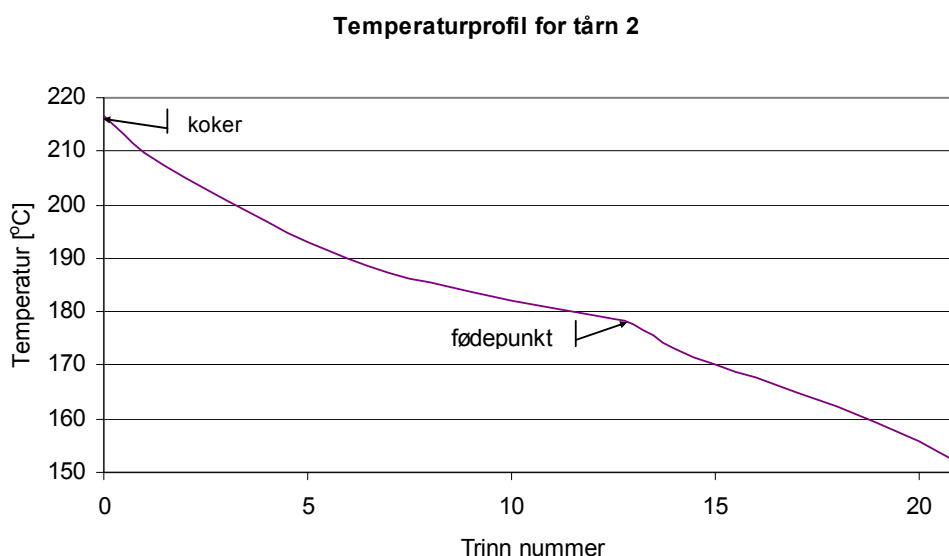
$$L_\theta = \ln\left(\frac{\theta - \theta_L}{\theta_H - \theta}\right) \quad (3.8)$$

Likning (3.8) krever i tillegg til en temperatur måling, to ekstra målte temperaturer som benyttes som referansetemperaturer. For en multikomponentblanding må referansetemperaturene være plassert noe bort fra endene, hvor innholdet av ikke-nøkkelkomponentene er forholdsvis konstant fra trinn til trinn. Temperaturdifferansene mellom referansetemperaturene og kolonne endene må behandles separat for å unngå å ta logaritmen til negative tall. En kan unngå dette ved å ta absolutt verdiene til differansene. For å unngå større påvirkning fra støy i temperaturer nære referanse temperaturen, bør også en minste tillatelig differanse være et kriterium. Ved å anvende referanse temperaturer istedenfor kokepunktstemperaturer oppnår en også en trykkkompensasjon av temperaturmålingene, og i tillegg til en kompensasjon for ikke-nøkkelkomponenter i følge Mejdell [Ref II].

3.2 Valg av målepunkter

Antallet og plasseringen av temperaturmålingene, og eventuelt andre målinger slik som trykk, i en destillasjonskolonne er viktig for å kunne gi gode estimat for den her søkte produktsammensetningen. En temperaturmåling på hvert trinn i kolonnen kan være en ønsket situasjon, men vil ikke være en garanti for en god estimator eller et trivielt estimeringsproblem.

Temperaturene målt i kolonnen vil ikke være uavhengige, og dermed vil antall frihetsgrader for problemet være mindre enn antall temperaturer hvis flere temperaturer i kolonnen anvendes (se avsnitt 2.2). Figur 3-1 viser en temperaturprofil for den simulerte destillasjonskolonnen som blir vurdert i denne oppgaven. Temperaturer på de ulike trinnene i kolonnen varierer ikke uavhengig av hverandre, og viser at antall frihetsgrader er mindre enn antall trinntemperaturer. Likning (3.1) gir antall frihetsgrader (3) for et binært system ved mindre forstyrrelser og konstant trykk. Flerkomponentsystemer og systemer med trykkvariasjon har flere frihetsgrader, men allikevel ett distinkt antall frihetsgrader færre enn antall mulige temperaturobservasjoner i kolonnen.



Figur 3-1 En temperaturprofil for tårn 2, hvor kokeren har trinn nummer null

Ett mindre antall temperaturer, som følge av antall tilgjengelige frihetsgrader i estimatoren, vil gi en dårligere estimator. Mejdell [Ref II] viste for de vurderte casene at mindre enn tre temperaturer for et binært system og mindre enn fire temperaturer for et flerkomponent-system ville kunne gi dårlige resultater. Med perfekte målingene ville beste plasseringer av målepunktene være mot enden av kolonnen. Inneholder målingene støy anbefaler Mejdell å plassere målingene nær midten av kolonnen. Hvis det er få temperaturmålinger kan plasseringen av målingene være av stor betydning for estimeringsresultatet. Et større antall målinger vil kunne bedre estimatoren sin kvalitet siden temperaturene vil inneholde mer uavhengig informasjon og effekten av målefeil blir redusert. I tillegg vil estimatoren være mindre sensitiv for plassering av målepunkt, endrede driftsbetingelser og vil bedre modellere avvik fra linearitet (ulineære responser). De fleste Industrielle kolonner har i følge Mejdell få temperaturmålinger i forhold til antall trinn, typisk antall er en pr. femte trinn. Det blir derfor foreslått å anvende alle de tilgjengelige målepunktene i industriell kolonner i estimatoren.

3.3 Ordinær minste kvadraters (OLS) metode

Destillasjon er slik som mange andre kjemiske prosesser et multivariabelt system karakterisert ved mange variable tilstander er med og beskriver produksjonsprosessen. Destillasjon er multivariabel fordi det er mange termodynamiske tilstandsvariable som påvirker hvert enkelt trinn i prosess, og det er sterk interaksjon mellom trinnene. Ordinær minste kvadraters metode (OLS), eller ofte kalt multippel lineær regresjon (MLR), er den vanlige metoden fra statikken som benytter minste kvadraters metode og er gitt ved

$$y = Xb + \varepsilon \quad (3.9)$$

hvor X -variablene er de uavhengige sentrerte måledataene, y er en eller flere sentrerte avhengige variable og ε er residualet. Sentrering blir foretatt for å fjerne konstantleddene. Regresjonsvektoren finnes ofte ved uttrykket for den generaliserte inverse av X , hvor vektoren da blir uttrykt ved

$$b = (X^T X)^{-1} X^T y \quad (3.10)$$

hvor $(X^T X)^{-1}$ er den inverse kvadratiske kovariansmatrisen. Metoden minimerer residualet ($\varepsilon^T \varepsilon$), hvilket er ekvivalent med en maksimering av tilpasningen til y . Dette innebærer at all variasjon i kolonnerommet til X , som varierer med y , vil bli inkludert i estimatet for b . Hvis det er støy i X som kovarierer med y , vil denne bli inkludert i b og føre til overtilpassing og dårlig prediksjon for fremtidige y -verdier (\hat{y}). I tillegg, for å kunne estimere b , krever OLS at X -verdiene er lineært uavhengige, dvs. $X^T X$ har full rang ($R < \min(n, q)$) og at antall variable, K , er mindre eller lik antall observasjoner eller prøver, N , ($N \geq K$).

En estimator som benytter flere temperaturmålinger i en destillasjonskolonne oppfyller ikke kravene som ble beskrevet i dette avsnittet, da temperaturmålingene i kolonnen vil være sterkt kolineære. Andre regresjonsmetoder som benytter for eksempel latente variabler kan redusere antall "variable" og dermed muliggjøre gode prediksjoner.

3.4 Regresjonsmetoder for latente variable (LVR)

3.4.1 Latente variable

For å unngå problemer med rang forbundet med OLS metoden kan en approksimere den målte datamatriksen X med en annen matrise med samme dimensjoner, men med lavere rang. Avhengigheten mellom variablene og objektene gjør at en kan finne at en kan finne et lite antall underliggende eller latente variable som kan forklare all systematisk informasjon i datasettet. De latente variablene kan trekkes ut fra X ved hjelp av en dekomponeringsmetode. Likning (3.11) viser en dekomponeringsformel, hvor X splittes i en informasjonsdel og en støydel (E).

$$X = TP^T + E \quad (3.11)$$

Innenfor kjemometri kalles T skårmatrisen og P ladningsmatrisen. Produktet av TP^T inneholder den systematiske informasjon og har samme dimensjoner som X , men denne matrisen har rang A . A vil være lik antall latente variable som er ekstrahert fra det opprinnelige datasettet, og vil på grunn av kolinearitet i dataene være mye mindre enn $\text{rang}(X)$. I eksplorativ analyse er formålet å tolke dataene, og å beskrive den systematiske variasjon med så få latente variable som mulig ved hjelp av såkalt prinsipal komponent analyse (PCA). De latente variablene i PCA kalles for prinsipale komponenter (PC), og er lineære kombinasjoner av de opprinnelige variablene. Den første prinsipalkomponenten (PC1) er den lineærkombinasjonen som forklarer mest mulig av variansen i X . PC2 vil deretter forklare mest av variansen etter at informasjonen forklart av PC1 er fjernet osv. (se avsnitt 3.4.2 for nærmere forklaringer av fremgangsmåten).

Skårvektorene t_a blir beregnet ved å projisere objektene i X ned på PC-ene,

$$t_a = Xw_a \quad (3.12)$$

hvor w_a er enhetsvektorene eller vektene, og er for PCA lik ladningsvektorene. Estimatoren kan beskrives som en lineær kombinasjon av de latente variablene, og kan uttrykkes ved følgende likning

$$\hat{y} = K_a t_a \quad (3.13)$$

K_a er en mindre matrise med a parametere for hver utgang, og har i følge Mejdell [Ref II] typisk 3 til 5 parametere for binære destillasjoner og 4 til 6 parametere for multikomponent destillasjoner. Faren for overtilpasning av regresjonsmodellene er et problem som blir adressert nærmere i senere avsnitt.

En annen metode for å trekke ut latente variable er å fjerne målinger i X , og dermed komme frem til mindre antall nesten uavhengige variable som latente variable. Denne metoden er selvsagt ikke en optimal løsning, da ekstra målinger vil forbedre estimeringen. Mejdell foreslår derfor å benytte, som tidligere nevnt, alle temperaturmålingene og å benytte latent variabel regresjonsmetoder som PLS og PCR.

3.4.2 Singulærverdidekomponering (SVD)

I OLS blir den generelle inverse av X søkt i regresjonen, men dette leder ofte til matriser som ikke lar seg invertere på grunn av singulærverdier. For å unngå dette problemet kan den generelle pseudoinverse for X benyttes, hvor retninger med lave singulær verdier (små PC-er) med samme størrelsesorden som ”målefeilen” blir fjernet.

$$Y = XK^T \quad (3.14)$$

$$K^T = X^\dagger Y \quad (3.15)$$

Den pseudoinverse X^\dagger blir lettest funnet ved å benytte SVD på X . I tillegg til å minimere kvadratet av differansen mellom virkelig y og estimert \hat{y} (dvs. $(y - \hat{y})^2$), blir også K minimalisert i løsningen.

En $n \times q$ matrise X , med $n \geq q$, kan ved singulærverdidekomponering skrives som

$$X = U\Sigma V^T \quad (3.16)$$

eller som summen av rang l matriser

$$X = u_1\sigma_1v_1^T + u_2\sigma_2v_2^T + \dots + u_m\sigma_mv_m^T \quad (3.17)$$

hvor m er rangen av X . Hvis $m < \min(n, q)$, er både kolonner og rader i X lineært avhengige (singulær). Kolonnevektorene $U, u_1 \dots u_m$ og $V, v_1 \dots v_m$, er ortogonale, i.e. $U^T U = V^T V = I_m$. U er egenvektorene til $X^T X$ og V er egenvektorene til XX^T . $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m, 0, \dots, 0)$ og singulærverdiene har $\sigma_1 \geq \sigma_2 \dots \geq \sigma_m > 0$. Den pseudoinverse av X blir

$$\Theta^\dagger = V\Sigma^{-1}U^T \quad (3.18)$$

Den inverse av Σ betyr at de minste singulære verdiene, σ_m , blir de største i den pseudoinverse. Sensitiviteten i den pseudoinverse for feil (for eksempel støy) i X kan bli stor hvis σ_m er liten. Forholdstallet $\gamma(X)$ vil da være stort:

$$\gamma(X) = \frac{\sigma_1}{\sigma_m} \quad (3.19)$$

I PCR er målet å redusere denne følsomheten for feil ved å velge $k \leq m$ ledd i likning (3.17).

3.4.3 Prinsippal komponent regresjon (PCR)

Likning (3.17) kan skrives på en alternativ måte

$$X = t_1 p_1^T + t_2 p_2^T + \dots + t_m p_m^T \quad (3.20)$$

De prinsipale komponentene, m , har skårvektoren (latent variabel), $t_m = u_m \sigma_m$, og ladningsvektoren, $p_m = v_m$. Kun de k første leddene som kan skilles fra målefeil beholdes, matrisene $P^{q \times k}$ og $T^{n \times k}$ vil da inneholde de viktigste informasjonsretningene og dermed $X \approx X_k = TP^T$. Minste kvadraters løsning for $Y = TK_t^T$ er

$$K_t = Y^T T (T^T T)^{-1} \quad (3.21)$$

P er ortogonal, og den fulle oversikten for estimatoren blir som følger

$$K_{PCR} = Y^T (X_k)^\dagger = Y^T T (T^T T)^{-1} P^T \quad (3.22)$$

3.4.4 Prinsipal minste kvadraters metode (PLS)

I PCR anvendes de prinsipale komponenter fra eksempelvis SVD som latente variable for X i regresjonen. De latente variablene blir beregnet uavhengig av om de samvarierer med informasjonen i Y . I PLS, eller ofte kalt PLS-regresjon, forsøker man derimot å dekomponere variablene i X -matrisen slik at de latente variablene best mulig beskriver Y -matrisen eller vektoren. Standard algoritmen for kalibrering av partiell minstekvadraters regresjonskomponenter, dvs. faktorer, er den ulineære iterative partiell minste kvadraters eller NIPALS (nonlinear iterative least squares).

3.4.4.1 Algoritme for ortogonaliserte PLS for en y-variabel [Ref XII]

For ytterligere beskrivelse av de matematiske begrepene som benyttes i PLS regresjonen som er beskrevet i punktene under, anbefales bøkene gitt i referanse X og XII, og for praktisk anvendelse i industrien kan referanse XIX anbefales.

1. De skalerte inngangsvariablene X og y blir først sentrert

$$X_0 = X - 1\bar{x} \text{ og } y_0 = y - 1\bar{y} \quad (3.23)$$

Velg A_{\max} til å være større enn antall forventede fenomener i X . For hver faktor $a = 1, \dots, A_{\max}$ utføres steg 2 -6.

2. Anvend den gjenværende variansen i y for å finne ladningsvekten w_a , ved minste kvadrats (LS) og lokal ”modell”

$$X_{a-1} = y_{a-1} w^T + E \quad (3.24)$$

og skaler vektoren til lengde 1. Løsningen er

$$\hat{w}_a = c X_{a-1}^T y_{a-1} \quad (3.25)$$

hvor c er skaleringsfaktoren som gjør at lengden av \hat{w}_a er lik 1, dvs.

$$c = \left(y_{a-1}^T X_{a-1} X_{a-1}^T y_{a-1} \right)^{-0.5} \quad (3.26)$$

3. Estimerer skårvektoren \hat{t}_a ved å bruke den lokale ”modellen”

$$X_{a-1} = t_a \hat{w}_a^T + E \quad (3.27)$$

LS løsningen er, siden $\hat{w}_a^T w_a = 1$,

$$\hat{t}_a = X_{a-1} \hat{w}_a \quad (3.28)$$

4. Estimerer x -ladningene p_a ved å bruke den lokale ”modellen”

$$X_{a-1} = \hat{t}_a p_a^T + E \quad (3.29)$$

hvilket gir LS løsningen

$$\hat{p}_a = \frac{X_{a-1}^T \hat{t}_a}{\hat{t}_a^T \hat{t}_a} \quad (3.30)$$

5. Estimerer den y -ladningen q_a ved å bruke den lokale ”modellen”

$$y_{a-1} = \hat{t}_a q_a + f \quad (3.31)$$

hvilket gir løsningen

$$\hat{q}_a = \frac{y_{a-1}^T \hat{t}_a}{\hat{t}_a^T \hat{t}_a} \quad (3.32)$$

6. Beregner \hat{E} og \hat{f} residualene ved å subtrahere den estimerte effekten av denne faktoren

$$\hat{E} = X_{a-1} - \hat{t}_a \hat{p}_a^T \quad (3.33)$$

$$\hat{f} = y_{a-1} - \hat{t}_a \hat{q}_a^T \quad (3.34)$$

Residualene etter a -faktorer kan deretter bli vurdert ved å benytte ulike statistiske målverdier, summering \hat{e}_{ik} for objektene i og variablene k , og summering av \hat{f}_i for objekt i . Ved å bytte ut X_{a-1} og y_{a-1} med de nye residualene \hat{E} og \hat{f} og øke a med 1. Danner nye X og y ved å sette

$$X_a = \hat{E} \quad (3.35)$$

$$y_a = \hat{f} \quad (3.36)$$

$$a = a + 1 \quad (3.37)$$

3.4.4.2 Antall PLS faktorer i modellen

Bestemmer A , antall gjeldende PLS faktorer som skal benyttes i kalibreringsmodellen. Beregner \hat{b}_0 og \hat{b} for A PLS faktorer,

$$\hat{b} = \hat{W} \left(\hat{P}^T \hat{W} \right)^{-1} \hat{q} \quad (3.38)$$

$$b_0 = \bar{y} - \bar{x}^T \hat{b} \quad (3.39)$$

hvor \hat{W} er matrisen av \hat{w}_a -vektorene for a faktorer og \hat{P} er matrisen av \hat{p}_a -vektorene.

De beregnede konstantene blir benyttet i estimatoren

$$\hat{y} = 1\hat{b}_0 + X\hat{b} \quad (3.40)$$

På grunnlag av den kalibrerte PLS modellen estimeres y -verdier for fremtidige objekter enten ved *full* prediksjon eller ved *kort* prediksjon.

3.4.4.3 Full prediksjon

For hvert nye objekt $i = 1, 2, \dots$ utføres de etterfølgende stegene 1-3

1. Skalerer inngangsdataene x_i tilsvarende som for kalibreringsvariablene, og beregner deretter

$$x_{i,0}^T = x_i^T - \bar{x}^T \quad (3.41)$$

\bar{x} er senteret fra de kalibrerte objektene. For faktor $a = 1, \dots, A$ utføres steg 2 til 3.

2. Søker \hat{t}_a som gitt i likning (3.28), dvs. $\hat{t}_a = X_{a-1} \hat{w}_a$.

3. Beregner nye residualer

$$x_{i,a} = x_{i,a-1} - \hat{t}_{i,a} \hat{p}_a^T \quad (3.42)$$

hvis $a < A$, økes a med 1 og fortsetter beregningen fra punkt 2, ellers ($a = A$) forsetter beregningen i punkt 4.

4. Predikerer y_i ved

$$\hat{y}_i = \bar{y} + \sum_{a=1}^A \hat{t}_{i,a} \hat{q}_a \quad (3.43)$$

og beregner eventuelle statistiske mål for vurdering av resultatene

3.4.4.4 Kort prediksjon

Alternativt til den fulle prediksjonen kan \hat{y} beregnes ved å bruke regresjonskoeffisientene \hat{b}_0 og \hat{b} fra likning (3.39) og likning (3.38), som dermed gir

$$\hat{y}_i = \hat{b}_0 + x_i^T \hat{b} \quad (3.44)$$

3.4.5 Vekting av variable

A priori vekting eller skalering av variablene kan være viktig i alle metoder hvor minste kvadraters metode (LS) blir brukt til å konsentrere informasjon til lavere ordens fenomener. Individuell vekting av X -variablene regresjonen kan gjøres etter følgende generelle likning

$$x_{i,k} = x_{i,k,0} \times \text{vekt}_k \quad (3.45)$$

Vekting av variablene kan hindre målingene med størst relativt endring å få dominerende vekt i estimatoren. Forskjellig typer variable i en X -matrise kan vektes slik at de får for eksempel får samme størrelsesorden. Den vanligste vekten er den inverse av standard avviket, hvor den totale initiale variansen for hver variabel i kalibreringssett blir drevet til 1. s_{ci} er standard avviket for verdiene i variabel, i , i kalibreringssettet og vekten (W) for variabel, i , blir gitt ved

$$W_{1,i} = \frac{1}{s_{ci}} \quad (3.46)$$

Temperaturene i destillasjonskolonnen vil ved vekting med standard avviket, bli skalert til samme den størrelsesorden. Endringen i de uskalerte temperaturene mot enden av kolonnen vil være små, og temperaturene mot enden av kolonnen vil dermed få store vekter ved denne skaleringen. Støynivået i temperaturene i enden av kolonnen kan i forhold til informasjonsinnholdet være relativ høyt, sammenliknet med temperaturer målt andre steder i kolonnen. Standardiseringen kan dermed forsterke innflytelsen av støy i kalibreringssettet, og en balansert vekting av variablene basert på kunnskap om støynivået i variabelsettet har blitt foreslått av Mejdell [Ref II] ved vekt funksjonen

$$W_{2,i} = \frac{1}{s_{c,i}} \frac{s_{c,i} - s_{e_k,i}}{s_{c,i}} \quad (3.47)$$

hvor $s_{e_k,i}$ er et estimat av det generelle støynivået i kalibreringssettet, definert som kvadratroten av den gjenværende variansen etter k faktorer (dvs. kvadratroten til det diagonalelement i av $\left(\frac{1}{n}(E_k^T E_k)\right)$). Variansen vil inkludere alle modell- eller datafeil, dvs. både støy og feil pga. ulinearitet. For vekten gjelder $s_{c,i} = s_{e_0,i} \geq s_{e_k,i}$ og $W_{2,i} = W_{1,i}$ hvis det ikke er støy, men gir ingen vekt til målinger hvor all variasjon skyldes støy.

3.4.6 Sammenlikning av PCR og PLS

I denne oppgaven blir som tidligere nevnt kun PLS regresjon benyttet, men både PLS og PCR regresjon har blitt benyttet med gode resultater for estimering av sammensetningen av produktstrømmene til destillasjonskolonner [Ref II, VII, VIII]. PLS og PCR har ulike styrker og svakheter når det gjelder anvendelse og eventuelle fallgruver knyttet til bestemmelse av antall faktorer i regresjonen. Noen punkter knyttet til denne problemstilling blir behandlet i dette avsnittet.

Det har i følge Esbensen [Ref XII] blitt hevdet at valg av PLS eller PCR regresjon ofte er et spørsmål om personlig skjønn. Mange statistikere foretrekker PCR fordi den er statistisk mer studert, mens ingeniører innen anvendt forskning benytter modellering med PLS pga at den er begrepsmessig lett forståelig. Mange kjemometrikere foretrekker også PLS fordi den kan gi en noe enklere modell å tolke og anvende, hvilket kan være et resultat av færre faktorer i estimatoren.

Modellen ut i fra PLS og PCR regresjon kan i følge Martens [Ref X] noe forenklet bli delt inn i grupper. X -data med høy nøyaktighet, for eksempel analyseresultater fra et nær infrarød (NIR) spektrometer. Spektrofotometeret har *høy* nøyaktighet og forenkler den grafiske tolkningen av PLS regresjonen, da den Y relevante informasjonen i X vil være i de første faktorene. Det kan allikevel forekomme at optimale antall bilineære faktorer i PLS regresjonen tilsvarer antall komponenter i PCR. Data med *lav* nøyaktighet kan føre til at PCR feiler mens PLS vil kunne gi brukbare resultater. PLS vil for Y -dataene med mye støy oftere gi en overtilpasset modell sammenliknet med PCR. PLS er mer følsom for overtilpassing, og et optimalt antall PLS faktorer er derfor mer avgjørende for resultatet enn for PCR. Valideringen av modellene vil da være nøkkelen til et optimalt antall faktorer i estimatoren.

Prediksjon med PLS og PCR estimatorene vil være på virket av eventuelle feilkilder. Kildene til prediksjonsfeil [Ref X] eller modell feil kan være forårsaket av problem med tilpassing av kalibreringsmodellen. Modellen vil i slike tilfeller ikke representere dataene i kalibreringssettet på en fullverdig måte, og den fremtidige prediksjonen kan bli dårlig. Modellkalibrering omfatter gjerne en lokal tilpassing til de ukjente, som er den mer eller mindre lineære X - Y funksjonen, og vil derfor alltid til en hvis grad inneholde modellfeil.

Mangel på representativitet kan oppstå når kalibreringssettet ikke dekker det totale området for variabiliteten til de fremtidige objektene. Evnen til prediksjon for fremtidige objekter som ligger utenfor kalibreringsområdet vil dermed være befengt med stor usikkerhet.

Tilfeldig støy i kalibrerings- og prediksjonsdataene vil virke på estimatet for kalibreringsparameterne, og dermed redusere prediksjonsevnen. Prediksjonsnøyaktigheten kan allikevel være bedre enn eksaktheten av data brukt i kalibreringssettet. Multivariat kalibrering kan kansellere noe av den tilfeldige støyen i kalibreringssett forutsatt at antall kalibreringsobjektet er høyere enn antall uavhengige kalibreringsparametere som skal bli estimert.

4 Prosessbeskrivelse

Oleon Scandinavia AS destillerer i sitt anlegg i Sandefjord fettsyrer fra ulike råvarer, som spenner fra fettsyrer fra fisk til kokosoljer. Destillasjon av den behandlede kokosoljen er det som blir inngående vurdert i denne oppgaven. Prosessanlegget for destillasjon eller fraksjonering av de ulike komponentene består av tre destillasjonskolonner i serie (se Figur 4-2), hvor destillasjonen foregår under vakuum betingelser.

4.1 Fettsyrene

4.1.1 Fettsyrer i kokosolje

Vegetabiliske fettsyrer består for det meste av monoene syrer, som er fettsyrer med et like antall karbonatomer. Kokosolje eller kokosfett inneholder flere ulike fettsyrer med kjedelengder fra C_6 til C_{20} og hovedsakelig mettede fettsyrer

Tabell 4-1 Innholdet av de enkelte fettsyrene i kokosolje, angitt med vanlig og IUPAC navn.

Fettsyre	Navn (IUPAC)	Masse % i kokosolje
$C_6H_{12}O_2$	kapronsyre (heksansyre)	<1.2
$C_8H_{16}O_2$	kaprylsyre (oktansyre)	3.4 – 15
$C_{10}H_{20}O_2$	kaprinsyre (dekansyre)	3.2 – 15
$C_{12}H_{24}O_2$	laurinsyre (dodekansyre)	41 – 56
$C_{14}H_{28}O_2$	myristinsyre (tetradekansyre)	13 – 23
$C_{16}H_{32}O_2$	palmitinsyre (n-heksadekansyre)	4.2 – 12
$C_{18}H_{36}O_2$	stearinsyre (n-oktadekansyre)	1.0 – 4.7
$C_{18}H_{34}O_2$	oljesyre (9-oktadekaen syre)	3.4 – 12
$C_{18}H_{32}O_2$	linolsyre (2,9-oktadekadien)	0.9 – 3.7
$C_{20}H_{40}O_2$	ikosansyre	<1

Tabellen over (Gunstone et. al [Ref XIII]) viser den naturlige variasjonen for fettsyresammensetningen i kokosolje, som er karakterisert ved sitt innhold av kortkjedede fettsyrer ($C_6 - C_{12}$), og det lave innholdet av umettede syre.

Fettsyrene som utvinnes av kokosoljen blir industrielt anvendt i for eksempel vaskemidler. Estere av fettsyrene har et vidt bruksområde innenfor næringsmiddelindustrien, som ingredienser i for eksempel iskrem og margariner, og innenfor farmasi og kosmetikk bransjen som tilsetning i kremer og kappselede legemiddelprodukter.

4.1.2 Destillasjon av fettsyrer

Effektiviteten i en destillasjonsprosess blir avgjort av forskjellene i kokepunkt for de ulike komponentene. Kokepunktene til fettsyrene blir karakterisert av kjedelengden og ikke av graden av mettede fettsyrer. Destillasjon vil derfor kunne skille fettsyrene C_6 og oppover i enkelt komponenter og i mettede og umettede fraksjoner for syrene som har en konjugert syre.

Den vanligste destillasjonen av fettsyrer foregår ved redusert eller evakuert trykk. Selv ved redusert trykk vil destillasjon av fettsyreblandinger bli utført ved moderat høye temperaturer. Høye temperaturer kan føre til polymerisering, syklisering og stereomutasjon av de umettede fettsyrene, men den viktigst faren er dekomponering av de organiske forbindelsene ved høye temperaturer. Destillasjon ved vakuumbetingelser vil senke den maksimale temperaturen (ved kokeren) i kolonnen, og dermed redusere faren for dekomponering av stoffene.

Tabell 4-2 Kokepunkt for noen fettsyrer ved 1 kPa.

Fettsyre	Kokepunkt [$^{\circ}C$] ved 1 kPa
$C_6H_{12}O_2$	94.6
$C_8H_{16}O_2$	121.3
$C_{10}H_{20}O_2$	145.5
$C_{12}H_{24}O_2$	167.4
$C_{14}H_{28}O_2$	187.6
$C_{16}H_{32}O_2$	206.1
$C_{18}H_{36}O_2$	224.1

Relativ flyktighet, α , er et vanlig utgangspunkt for vurdering av eventuelle vansker ved destillasjonen. Relativ flyktighet kan ved antagelse om ideelt system defineres ved

$$\alpha_{AB} = \frac{P_A}{P_B} \quad (4.1)$$

hvor P_A er metningstrykket til komponent A ved blandingens temperatur.

Tabell 4-3 Relativ flyktighet for laurinsyre, myristinsyre og palmitinsyre.

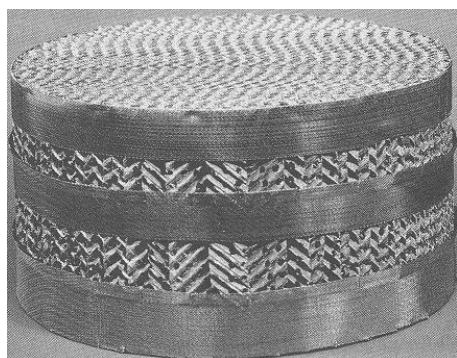
Temperatur [°C]	Relativ flyktighet, α	
	$C_{12} - C_{14}$	$C_{12} - C_{16}$
150	3.13	10.14
160	2.97	8.99
170	2.84	8.07

Verdiene i Tabell 4-3 angir den relative flyktigheten for fettsyrene i den øvre delen av kolonnen, og er beregnet ved Antoine likningen tilpasset eksperimentelle data (Vapor-Liquid Equilibrium Data Collection [Ref XIV]). Den relative flyktigheten viser at destillasjonen er forholdsvis enkel, hvilket resulterer i forholdsvis høy teoretisk renhet av produktet ved destillasjon.

4.2 Oversikt over prosessen

4.2.1 Prosessanlegget

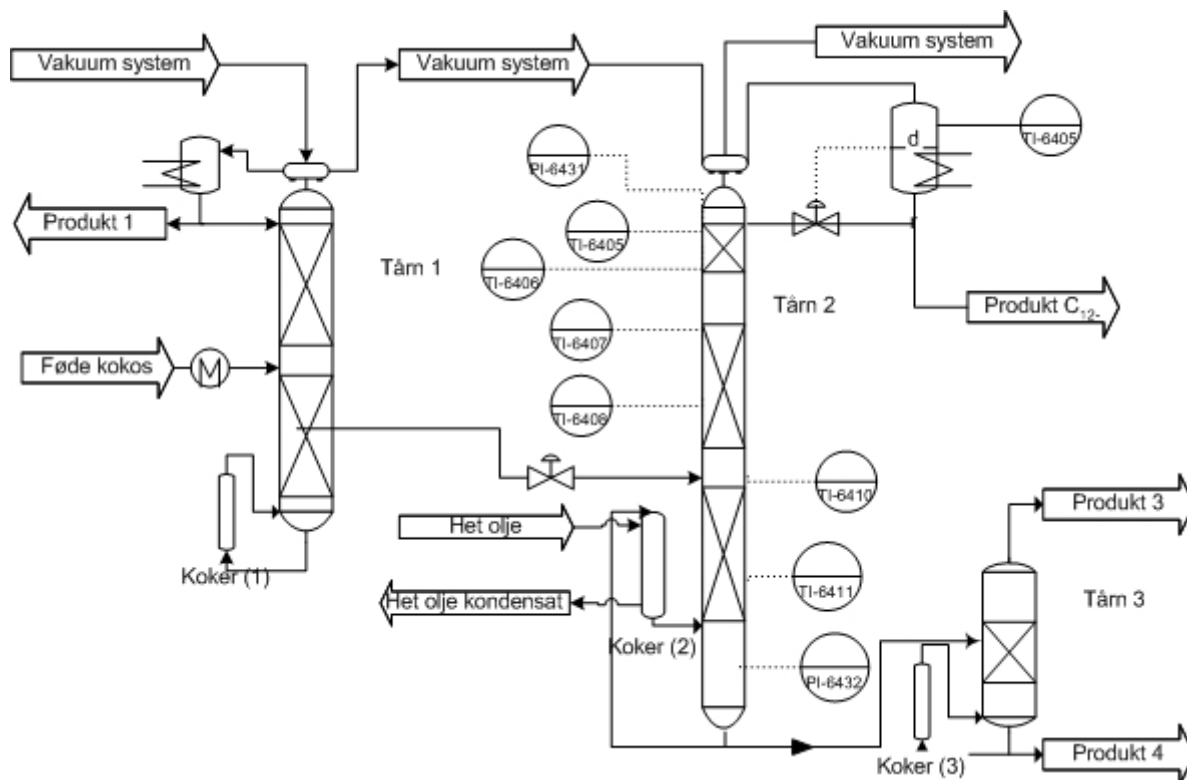
Prosessanlegget består av tre kolonner i serie. Destillasjonen foregår under vakuum og kolonnene har et vakuumsystem drevet av et vandampsystem med høyere trykk. Anlegget og kolonnene er av eldre dato, men har i nyere tid fått nye innmater. Pakningen i kolonnene er av typen strukturert pakning, og består av vertikalt stilte rifleformede plater. Strukturerte pakninger er fordelaktig i vakuumdestillasjon pga et lavere trykktap pr. teoretiske trinn eller meter i kolonnen. Et eksempel for en pakning er gitt i figuren nedenfor.



Figur 4-1 En strukturert pakning

Tårn 2 i anlegget, som blir behandlet i denne oppgave, har tre seksjoner med strukturerte pakninger. Hver seksjon har i forkant av pakningene en væskespreder, som fordeler væskestrømmen (L) i kolonnene over pakningstverrsnittet. Tilsvarende er det etter hver pakningsseksjon plassert en væskesamler. Den nedre (strippe) seksjon av kolonnen, hvor forholdet mellom massestrøm av væske og gass er større enn for den øvre (rektifisering) seksjonen, har mindre tverrsnitt enn den øvre delen.

Råvaren, kokosolje, består av tunge og lette komponentene fra C_6 til C_{22} . Oljen blir først mettet ved hydrogenering i en reaktor. De mettede fettsyrene blir deretter fraksjonert i de tre kolonnene i serie. Fettsyrene blir fødet til tårn 1 og den letteste fraksjonen, C_6 til C_{10} , danner strøm angitt som produkt 1 i prosessoversikten. Bunnfraksjonen fra kolonne 1 er fødestrøm til tårn 2.



Figur 4-2 Prosessflytdiagram med temperatur og trykk TAG for tårn 2

I tårn 2 er det ønsket å få 99 % (masse %) laurinsyre i toppstrømmen. Føden til kolonnen blir antatt å være væske ved sitt kokepunkt. Kokeren for kolonnen står på sitt maksimale effekt. Væsknivået i bunnen av kolonnen blir statisk regulert, ved at bunnproduktet fra kolonnen flyter direkte til tårn 3 ved et overløp (i det virkelige anlegget står kolonne 3 direkte under kolonne 2). Innstillingen for overløpsplaten til kolonne 2 er konstant. Dampstrømmen (V) i kolonnen blir total kondensert, og noe underkjølt, i kondensatoren. Kondensatet går deretter til refluxakkumulatoren. Refluksen til kolonnen blir regulert av væsknivået i akkumulatoren, hvor væsknivået i akkumulatoren blir holdt konstant. Det eksterne refluxforholdet ligger i ved dagens drift i området [1 – 1.5]. Renheten av C_{12} i destillatet blir i dagens anlegg regulert ved at operatørene regulerer destillat mengden for å kunne nå den spesifiserte sammensetningen. Den ønskede renhet av C_{12} har sin maksimale verdi i dagens anlegg ved 99 % C_{12} .

Bunnstrømmen fra kolonne 2 som hovedsakelig består av de tyngste komponentene C_{14} til C_{22} , blir i den siste kolonnen fraksjonert i to produkter. Bunnstrømmen og dens produktstrømmer og de tilhørende enhetene vil ikke bli nærmere behandlet i denne oppgaven.

4.2.2 Prosessparametere

Føden til tårn 2 består som nevnt hovedsakelig av fettsyrer fra C_{12} til C_{22} . En ideell fødeblanding til kolonnen ville kun inneholde disse komponentene, men ved den nåværende driften av anlegg vil føde også inneholde de lettere fettsyrene C_6 til C_{10} . Innholdet av C_{10} i produktstrømmen fra tårn 2 vil være av avgjørende betydning for den oppnådde renhet av C_{12} . Destillasjonen må derfor betraktes som et multikomponentsystem. Destillasjonen vil ha to nøkkelkomponenter, C_{12} og C_{14} , lette ikke-nøkkelkomponenter bestående teoretisk av tre fettsyrer, $C_6/C_8/C_{10}$, og de tunge ikke-nøkkelkomponentene fra C_{16} til C_{22} . Fødens konkrete sammensetning vil være avhengig av driften av tårn 1 og mulig variasjon i sammensetningen av de tilførte hydrogenerte kokosoljene. Tabell 4-4 viser den gjennomsnittlige sammensetning føden hadde i perioden 23. – 28. januar 2002 og den ønskede sammensetningen til kolonnen. Innholdet av C_{10} lå i perioden langt over den ønskede verdien, og dermed ble heller ikke den ønskede renheten av C_{12} i produktstrømmen oppnådd.

Tabell 4-4 Sammensetning for fødestrøm til kolonne 2, normal og ønsket.

	Sammensetning [masse %]	
	Ønsket	Gjennomsnitt (januar 2002)
Fettsyre		
C_{10}	—	1.2
C_{12}	54	48.6
C_{14}	22	22.2
C_{16}	11	12.0
C_{18}	14	15.0
C_{20}	—	1.0

Figur 4-2 viste skjematisk basisprinsippene for den aktuelle kolonnen som ble vurdert. Kolonnen har sitt fødepunkt over den første seksjonen med strukturerte pakninger. Føden til kolonnen er bunnstrømmen fra kolonne 1, og den blir antatt at den er væske på sitt kokepunkt. Temperaturen i toppen av kolonnen, ved 5mbar, er ved ønsket destillatsammensetning omtrentlig 153 °C. Bunnproduktet forlater kolonnen ved 195 °C, ved 19 mbar trykk. Kokeren for kolonnen er en varmeveksler med olje som varmemedium. Dampstrømmen forlater kokeren med stor hastighet, noe som er vanlig i kolonner med strukturerte pakninger. Forholdstallet væske-/damphastighet for kolonnen er lite dvs. lave væske- og høy damphastighet.

Vakuumsystemet, som blir drevet av en trykkdrevet dampsløyfe, vil holde trykket i toppen av kolonnen omtrent ved 4.5-5.5 mbar. Trykket i bunnen av kolonnen vil være påvirket av trykkfallet over kolonnen. Trykkfallet er igjen avhengig av belastning og væske/damp raten. Trykkfallet over kolonnen vil ved normal drift være om lag 14 mbar.

Det er plassert åtte termometer i kolonnen. En måling er i refluksakkumulatoren, og gir en indikasjon på underkjølingen og en ved kokeren. En temperatur blir målt i toppen av kolonnen, og en temperatur måles ved fødepunktet. I rektifiseringsdelen av kolonnen er det ytterligere tre termometer. Under fødepunktet blir en temperatur målt. Den nærmere plasseringen av målingene blir beskrevet senere i avsnittet.

Trykket blir målt i toppen av kolonnen og i bunnen av kolonnen. Trykket blir ikke målt flere steder i kolonnen. Trykkprofilen for kolonnen, og dens avhengighet, kan derfor ikke vites nøyaktig.

Der er flere forstyrrelsene som vil påvirke kolonnen. Føden til kolonnen vil variere i sammensetning og massestrøm, etter drift av kolonne 1 og i tillegg endring av fødesammensetning og endring i massestrøm til kolonnen. Denne forstyrrelsen vil være en forholdsvis rask forstyrrelse, og vil derfor kunne av avgjørende betydning for responsen i kolonne 2. Trykket i kolonnen blir opprettholdt et vakuum system, men er ikke en regulert variabel. Trykkvariasjon vil dermed forekomme grunnet endringer i vakuumsystemet og endrede betingelser i kolonnen. Endringene av trykket i kolonnen har relativt større utslag i nedre del av kolonnen sammenliknet med øvre del av kolonnen. Vakuum systemet kan under gitte betingelser dra med seg fettsyrer inn i vakuumsystemet, og dette kan medføre store trykk utslag med kortere varighet. Refluksen i kolonnen blir regulert med mål om konstant nivå i refluksakkumulatoren, men refluksvæskestrøm varierer med store utslag med forholdsvis høy frekvens (nærmere omtalt i kapittel 6).

4.3 Prosessmodellen

4.3.1 Prosessmodellen i Hysys

Laurinsyre, C_{12} , er topp-produktet fra kolonne 2 i anlegget, og denne blir ønsket produsert med høy renhet. En estimator for topp-produktet i kolonnen med indirekte målinger for sammensetningen skulle beregnes. Forsøkene eller dataene for å kunne etablere en estimatoren må kunne spenne hele det forventede variabel eller forstyrrelse rommet med tilhørende responser i eksempelvis temperatur og trykk i kolonnen. Kolonne 2 ble derfor simulert statisk med en enkel modell i Hysys Plant™. Hysys tilbyr en tilnærming for prosessmodellering og simulering. Hysys inneholder en rekke enhetsoperasjoner som kan settes sammen til komplette prosess-systemer og anlegg. Kolonnen i anlegget i Sandefjord har strukturerte pakninger, men Hysys har ikke dette som en standardenhet. En standard platekolonne ble valgt som tilnærming for destillasjonskolonnen. Antall teoretiske trinn i den virkelige kolonnen og andre spesifikasjoner ble angitt for å kunne tilnærme modellen. Hysys inneholdt bibliotek over organiske og uorganiske forbindelser, og inneholdte også de mettede fettsyrene fra C_{10} til C_{20} .

Trykket i kolonnen ble antatt som en lineær sammenheng av trykket i toppen og bunnen av kolonnen. Trykket for føden ble antatt å være omtrent lik trykket ved kolonnens fødepunkt. Trykktapet over kolonnen blir dermed antatt å være lineært, men trykket blir ved normale betraktninger uttrykt ved oppkoket (V), og kan for eksempel ha en andreordens karakter.

4.3.2 Termodynamikk

Fettsyrer har en polar karakter som avtar med økende kjedelengde. Avviket fra idealitet vil dermed være omvendt proporsjonalt med kjedelengden. Damp – væske likevektsdata for binære blandinger av de ulike fettsyrene er undersøkt eksperimentelt og gjengitt i litteraturen [Ref XIV]. Resultater fra simulering av binære blandinger av fettsyrene med væske aktivitetskoeffisientmodellene; UNIQUAC, Wilson og NRTL, ble sammenliknet med de rapporterte eksperimentelle dataene. Manglende interaksjonsparametere ble beregnet med UNIFAC. Wilson og NRTL gav omtrent like resultater, men Wilson ble foretrukket i den endelige modellen av kolonnen. Aly og Ashour (Ref XV) foreslår anvendelse av "Perturbed Hard Chain" tilstandsligning, PCH-EOS, men denne er ikke en opsjon i Hysys™ og ble derfor ikke forsøkt.

4.3.3 Antall teoretiske trinn

Teoretiske trinn i kolonnen er beregnet av Morten Gulbrandsen [Ref XVI], og ble etterprøvd ved å standard teoretisk trinntallbestemmelse ved Fenske likningen på en "short cut kolonnen" i Hysys, og ved en mer rigorøs platekolonne i Hysys. Antall teoretiske trinn i en pakket kolonne blir normalt beregnet på grunnlag av HETP (høyde ekvivalente teoretiske trinn) beregninger. HETP for strukturerte pakninger kan beregnes ved empiriske korrelasjoner eller ved "tommelfinger regler". Metoder estimering av HETP for strukturerte pakninger er presentert av Kister et. al. [Ref XVII]. I boken blir de korrigererte "tommelfinger reglene" generelt anbefalt ved HETP beregninger.

HETP verdier vil generelt kun betraktes som et overslag for antall teoretiske trinn i kolonnen. HETP verdiene kan også bli endret ved endrede driftsforhold, og vil dermed kunne være en "dynamisk" verdi. En faktor som kan påvirke HETP er væske og damp strømmene i kolonnen. En kort oversikt for fastsettelse av riktige HETP verdier for kolonnen, og også eventuelt ulike verdier for de ulike komponentene, er gitt av Gill og Adranvala [Ref XVIII]. Den vanligste måten å angi en eventuell trinnvirkningsgrad i kolonnen er ved Murhpre likning [Ref XVII]. For binære blandinger vil virkningsgraden være identisk for de to komponentene. Trinnvirkningsgraden for flerkomponentseparasjon vil virkningsgraden være ulik for alle komponentene. Dette skyldes i følge Kister [Ref XVII] blant annet forskjellig diffusivitet for komponentene i gass og væskefasen og koblinger og interaksjon for masseoverføringene mellom de ulike komponentene.

Det teoretiske antall trinn i kolonnen, med et trinn i kokeren, ble anslått til 21 og dette antallet ble dermed brukt i modellen. Trinnvirkningsgraden var for alle komponentene lik en. Temperaturmålingene i kolonnen i forhold til de teoretiske trinnene, som ble brukt i modellen i Hysys, ble vurdert ved sammenlikning av driftsdata og analyser av strømmer i anlegget og resultater fra simulering i Hysys med tilsvarende innganger og driftparameterer. Den endelige plasseringen av temperatur- og trykkmålingene i kolonnen er gitt i Tabell 5-6.

5 Estimator for produktsammensetningen

Produkt estimatoren ble estimert på grunnlag av data fra simuleringene, og estimatorens følsomhet for ulike forstyrrelser ble testet. I det følgende kapitlet presenteres resultatene og forhold som ble vurdert under arbeidet med estimatoren. Dette innebærer blant annet valg av kalibreringssett og kriterier for estimatorens godhet.

5.1 Forberedelser

5.1.1 Valg av kalibreringssett

Valget av kalibreringssett for produkt estimatorer, og estimatorer generelt, vil ha avgjørende betydning for det oppnåelige resultatet. For å kunne predikere $\hat{Y} = f(X)$ med en ønsket nøyaktighet og pålitelighet, må kalibreringsdataene inneholde tilstrekkelig informasjon. Kalibreringssettet må derfor spenne alle retninger med systematisk variabilitet, inkludert interferenser, slik at de kan bli modellert. En strategi for valget av forsøkene som skulle inngå i kalibreringssett var derfor viktig. Et lineært problem med n variable som blir variert på to nivåer gir 2^n antall forsøk i et faktorforsøk.

Føden til tårn 2 inneholder i praksis syv komponenter. Utgangspunktet for simuleringene var allikevel at føden kun skulle inneholde fire komponenter, C_{12} til C_{18} . Dagens føde inneholder i varierende grad C_{10} , men målet for destillasjonen er 99 % renhet av C_{12} i topp-produktet. Hvis føden inneholder en for større mengde C_{10} , vil dermed ikke dette kunne nås. I fremtiden må derfor innholdet av C_{10} være på et meget lavt nivå, og ble dermed utelatt fra kalibreringssettet. Effekten av eventuell C_{10} i føden blir vurdert i avsnitt (5.2.5). Antall mulige kombinasjoner for de fire komponentene var $4^2 = 16$, hvis man så bort fra blandeproblematikken. Fødesammensetningen var reelt sett et blandeproblem, hvor de ulike komponentene kunne variere innenfor en øvre og nedre grense og med sum av massefraksjonene like en. Det er beskrevet flere metoder for valg av blandingsdesign eller forsøksdesign (Ref XIX). En forenkling av problemet ble gjort ved å begrense antall mulige kombinasjoner av sammensetningen. Føden ble antatt beskrevet ved en lett føde, relativt mye C_{12} i forhold til C_{14} , og en tyngre føde (Tabell 5-1). Lineære kombinasjoner av de to klassene ble antatt å kunne gjenspeile de ulike fødesammensetningene på en tilfredsstillende måte.

Trykket i kolonnen varierer med fødemengden, væske og damp strømmene i kolonnen, og produsert trykk av vakuumsystemet. I den statiske modellen av kolonnen var trykket ikke angitt som en funksjon av for eksempel dampmengden (V) i kolonnen, men kun en parameter i kolonnemodellen. Trykkvariasjonen og trykktapet over kolonnen var derfor en variable som ble brukt i forsøksdesignen.

Kolonnen hadde ytterligere to frihetsgrader etter at trykket i kolonnen er spesifisert. Massefraksjon C_{12} i toppen av kolonnen, som også er den søkte y -variabelen i estimatoren, var en naturlig variabel. C_{12} i bunnproduktet var den andre variabelen ble anvendt i forsøksplanen, og dermed var antall frihetsgrader reduserte til null og et fullt spesifisert system. Regulering av de samme parameterne ble anvendt av blant annet Mejdell [Ref III] for å kunne undersøke alle prosessutslag (svake og sterkt retninger).

Variablene som ble undersøkt ved forsøksdesignen var dermed fem, og antall forsøk var 32 i en full faktoriell design. Senterpunktet for designet ble i tillegg undersøkt, for å kunne avdekke ulinearitet i prosessresponsen. Hvis en blandingsdesign for føden og en vanlig design hadde blitt valgt for prosessvariablene ville dette medført et mer komplekst oppsett, men uten at dette nødvendigvis ville bedre resultatet i særlig grad.

Den stasjonære modellen impliserte en perfekt regulering av kolonnen. Målingene av de ulike variablene som eventuelt skulle inngå i estimatoren, var også selvsagt perfekte målinger av de ulike størrelsene. Eventuelle feil i simuleringresultatene var kun begrenset til numeriske feil. De virkelige målingene i kolonnen er derimot alltid beheftet med støy (avsnitt 5.1.4) og eventuelt et bias avvik for målingene. Resultatene fra simuleringene kunne allikevel inneholde en stor feilkilde, hvis de simulerte resultatene avvek fra de virkelige målte verdiene på grunn av modellfeil.

Estimatorene ble på grunnlag av kalibreringssettet beregnet i Matlab R12™ med en PLS rutine i PLS_Toolbox 2.1 fra firmaet Eigenvector Research [Ref XX]. Estimatorene ble vurdert med valideringsrutiner i fra den samme rutinepakken. The Unscrambler™ ble også brukt for i evalueringen av resultatene.

5.1.2 Forhold som ble vurdert

En estimator for innholdet av C_{12} i produktstrømmen i toppen av kolonnen, kunne anvende målinger som allerede fantes i kolonnen (i Sandefjord). Antallet temperaturmålinger i kolonnens seksjoner var seks, i tillegg til målinger i refluksakkumulator og bunnen av kolonnen. Plasseringen av målingene i kolonnen i forhold til de teoretiske trinnene i den modellerte kolonnen, var et punkt som ble undersøkt. En uriktig plassering av anvendte temperaturdataene i modellen, kunne introdusert en bias feil, men viktigst en feil på grunn av ikke linear temperaturprofil eller respons. Mejdell påpeker at kolonner med høy produktrenhet er ulineare prosesser, og at transformering og vekting av variablene kan være påkrevd. Graden av ulinearitet etter hvilken produktrenhet som ble oppnådd eller ble søkt, var derfor også et emne som ble undersøkt.

Mejdell [Ref II] undersøkte trykkvariasjon innenfor $\pm 10\%$ fra base casen, og konkluderte med at trykkmålinger ikke forbedret estimatoren. I den undersøkte destillasjonen av fettsyrene var det tildels store trykkvariasjoner i kolonnen. En eventuell forbedring av estimatoren ved anvendelse av trykkmålinger ble derfor undersøkt,

Estimatoren, som bygget på simulerte og dermed ideelle data, skulle kunne anvendes på en virkelig industriell kolonne. En vurdering av forstyrrelse som sannsynlige vil påvirke eller være avgjørende for ”godheten” av de estimerte verdiene, ble derfor foretatt. I tillegg ble det gjort en vurdering av andre forhold som ville kunne være svakheter og styrker for estimatoren. Tilslutt ble estimatoren forsøkt brukt på historiske data fra prosessen, og en fikk dermed en kvalitativ indikasjon på hvordan de estimerte verdiene forholdt seg i forhold til analyserte verdiene.

5.1.3 Evalueringskriterier

Modellene ble hovedsakelig vurderte etter to standard statistiske mål. Evalueringsmålene var tilnærmet like de Mejdell [Ref II], Kresta et. al. [Ref XXI] og Liu et. al. [Ref IX] benyttet. PLS regresjon er følsom for overtilpasning. Modellene ble validert ved kryssvalidering, hvor prøvesettet blir delt inn i undersett (eventuelt med en prøve i hvert sett ved små matriser). Deretter ble resultatet for det utelatte settet predikert fra modellen som var utledet fra resten av datasettet. Prediksjonsfeilen eller PRESS (prediction residual error sum of squares) ble beregnet ved

$$PRESS_k = \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (5.1)$$

hvor y er målingene av produktsammensetningen, \hat{y} er dens estimerte, N er antall målinger og k er antall faktorer i estimatoren. Et mål for modellens evne til å predikere nye prøver blir uttrykt ved RMSECV (root-mean-square error of cross variation)

$$RMSECV = \sqrt{\frac{PRESS_k}{N}} \quad (5.2)$$

RMSECV vil for utransformerte y -verdier ha samme måleenhet som y . Prosent varians fanget (Percent of explained variance, EPV) var det andre kriteriet estimatoren ble vurdert etter. Den er definert ved

$$EPV = \frac{SSTO - PRESS}{PRESS} \times 100\% \quad (5.3)$$

SSTO er den totale summen av kvadrerte beregnede feil (total sum of squared errors calculated) som ble beregnet ved

$$SSTO = \sum_{i=1}^N (y_i - \bar{y})^2 \quad (5.4)$$

hvor \bar{y} var gjennomsnittet av målingene.

5.1.4 Støynivået i målingene

Den forventede støyen i målingene kunne ikke bli estimert fra historiske data fra prosessen, da de registrerte dataene kun var trendverdier. Trendingen av verdiene innebar at nye endrede verdier innenfor en gitt grense \pm % ikke ble loggført før de endrede verdiene lå utenfor området.

Tekniske data for trykktransmitteren oppgav en nøyaktighet på \pm 0.1 – 0.2 %. Temperaturtransmitteren hadde en nøyaktig \pm 0.1 % innenfor det innstilte området.

Resultatene fra simuleringene ble i beregningene av estimatorene påført et hvitt støysignal fra 0 til 0.2 %.

5.1.5 Simuleringene

Simuleringene var statiske, og gjengav ikke kolonnens virkelige følsomhet for øket eller redusert fødemengde. Endret belastning og til dels fødesammensetning førte til endret endrede væske og dampstrømmer. Trykket var ikke angitt som en funksjon av oppkoket. Variasjonen i trykktapet overkolonnen ble dermed ikke direkte påvirket ved endret føde, men angitt som en variabel i simuleringene.

Kalibreringssettene for estimatorene trykk- og temperaturmålinger ble beregnet med prosessmodellen i Hysys. Kalibreringssettet inneholdt variasjoner i fødesammensetning, C_{12} i topp- og bunnprodukt, og trykkvariasjon.

Kalibreringssettet inneholdt forventede minimal og maksimal verdier for de undersøkte variablene. Fødesammensetningen ble som tidligere nevnt fastsatt ved en blanding av en lett og en tung sammensetning, gitt innenfor intervallet [0 1] (Tabell 5-1)

Tabell 5-1 Verdier for sammensetning av føden gitt innenfor intervallet [0 1]

Komponent	Minimum (-) massefraksjon	Maksimum (+) massefraksjon
C_{12}	0.45	0.54
C_{14}	0.25	0.22
C_{16}	0.14	0.11
C_{18}	0.16	0.13

Tabell 5-2 Design for simulert kalibrerings- og valideringssett for 97 % til 99.5 % C₁₂ i produktstrømmen

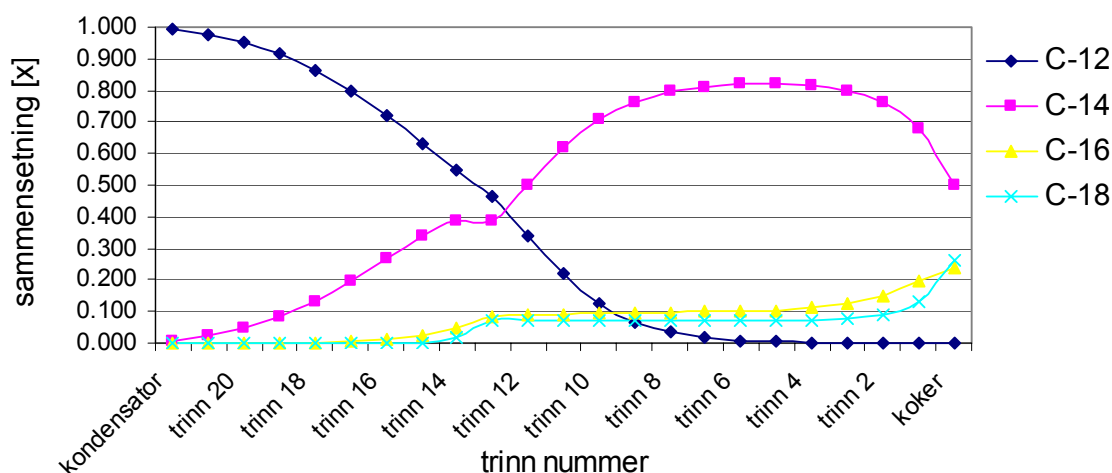
Variabel	Minimum (-)	Maksimum (+)	Enhet
Føde	0	1	
C ₁₂ i topp-produktet	0.97	0.995	Massefraksjon
C ₁₂ i bunnproduktet	.01	.04	Massefraksjon
Trykk i topp av kolonnen	4.5	5.5	mbar
Trykktap over kolonnen	10	18	mbar

Simuleringene som ble gjort var som beskrevet i tabellen ovenfor. I tillegg til punktene angitt i forsøksdesignet ble senterpunktet og tilfeldig utvalgte punkt innenfor de beskrevne maksimums og minimumsgrensene simulert (gitt i Bilag A). Resultatene fra simuleringene utgjorde kalibrerings- og valideringssett. Modeller for estimeringen med og uten trykkvariasjoner ble vurdert. I modeller hvor trykket ikke ble variert, var trykket på en middelvei. Tilsvarende kalibrering og valideringssett som ovenfor ble laget for produkt med 95 % C₁₂ (Tabell 5-3) og for kolonne med varierende antall teoretiske trinn i kolonnen. Kalibreringssettet for forskjellig antall teoretiske trinn (21 og 20) var tilsvarende som forsøkene gitt i Tabell 5-2, men med variasjon av antall teoretiske trinn. Figur 5-1 viser en typisk sammensetningsprofil for komponentene, C₁₂ til C₁₈, på trinnene i kolonnen.

Tabell 5-3 Design for simulert kalibrerings- og valideringssett for 93 % til 97 % C₁₂ i produktstrømmen

Variabel	Minimum (-)	Maksimum (+)	Enhet
Føde	0	1	
C ₁₂ i topp-produktet	0.93	0.97	Masse %
C ₁₂ i bunnproduktet	0.02	0.08	Masse %
Trykk i topp av kolonnen	4.5	5.5	mbar
Trykktap over kolonnen	10	18	Mbar

Sammensetning på trinnene i kolonnen



Figur 5-1 Typisk sammensetningsprofil for fettsyrene i kolonnen

5.1.6 Plassering av temperaturmålingene i kolonnen

De målte temperaturene i kolonnen i Sandefjord hadde en tilnærmet plassering på de teoretiske trinnene i prosessmodellen etter følgende fordeling:

Tabell 5-4 Plassering av målingene i destillasjonskolonnen (med TAG nummer i anlegget) på teoretisk trinn i prosessmodellen, og notasjon for måling i de beregnede estimatorene

Trinn i prosessmodell	TAG i anlegget	Notasjon i estimator	Måleenhet
Trinn 21	TI-6405	T ₁	°C
Trinn 19	TI-6406	T ₂	°C
Trinn 16	TI-6407	T ₃	°C
Trinn 15	TI-6408	T ₄	°C
Trinn 13	TI-6410	T ₅	°C
Trinn 5	TI-6414	T ₆	°C
Trinn 21	PI-6431	p ₁	mbar
Trinn 1	PI-6432	p ₂	mbar

Notasjonen som er angitt i Tabell 5-4 ble benyttet i angivelse av målingene i estimatorene. Det maksimale antallet temperaturmålinger i en estimator eller modell hadde øvre grense ved det tilgjengelige antallet målinger i kolonnen. Målingene av trykket i kolonnen ble benyttet i estimatoren når dette er spesifisert.

5.2 Estimator for produktsammensetningen

5.2.1 Notasjoner for estimeringsproblemene

Betegnelse for estimatorene med tilhørende inngangsvariable er listet i Tabell 5-5 og Tabell 5-6.

Tabell 5-5 Oversikt over estimatorene med tilhørende temperaturmålinger

Betegnelse for estimator	Målinger i estimatoren
Temp_1	alle temperaturer
Temp_2	T ₁ , T ₂ , T ₃ , T ₄ , T ₅
Temp_3	T ₁ , T ₂ , T ₃ , T ₄
Temp_4	T ₁ , T ₂ , T ₃ , T ₅
Temp_5	T ₂ , T ₃ , T ₄ , T ₅
Temp_6	T ₁ , T ₂ , T ₃

Tabell 5-6 Oversikt over estimatorene med tilhørende temperatur- og trykkmålinger

Estimator betegnelse	Målinger
TP1	alle temperturer og trykk
TP2	T ₁ , T ₂ , T ₃ , T ₄ , T ₅ , p ₁ og p ₂
TP3	T ₁ , T ₂ , T ₃ , T ₄ , p ₁ og p ₂
TP4	T ₁ , T ₂ , T ₃ , T ₅ , p ₁ og p ₂
TP5	T ₂ , T ₃ , T ₄ , p ₁ og p ₂
TP6	T ₁ , T ₂ , T ₃ , T ₄ , T ₅ , p ₁
TP7	T ₁ , T ₂ , T ₃ , T ₄ , p ₁
TP8	T ₁ , T ₂ , T ₃ , T ₄ , T ₅ , p ₂

Målingene i den virkelige prosessen vil alltid inneholde målestøy og måleusikkerhet. En tilnærming til den virkelige prosessen og kalibrering med historiske data, ble gjort ved å introdusere støy i X matrisen (temperatur- og trykkmålingene). Støy påført simuleringresultatene for temperatur- og trykkdataene dannet dermed de nye kalibrerings og valideringssettene. Tabell 5-7 viser en oversikt over betegnelsene som ble benyttet for de ulike støynivåene i målingene. Verdiene for sammensetningen ble alltid benyttet uten støy.

Tabell 5-7 Oversikt over betegnelse for datasett med støy i målingene

Betegnelse	Støy i måling (\pm)	
	Temperatur [$^{\circ}$ C]	Trykk [mbar]
n_0	0	0
n_1	0.1	0.01
n_2	0	0.02
n_3	0.2	0
n_4	0.1	0.02
n_5	0.2	0.01
n_6	0.2	0.02

A priori vekting eller skalering av variablene er viktig i alle metoder hvor minste kvadraters metode blir benyttet. Fire forskjellige skaleringer og transformasjoner ble benyttet i PLS regresjonen. Alle variable ble sentrerte før PLS regresjonen.

Tabell 5-8 Oversikt skaleringer og transformasjoner av variable

Betegnelse	X -matrise	y -vektor
S_1	ingen	ingen
S_2	$1/SD_X$	ingen
S_3	$1/SD_X$	$\ln(1-y)$
S_4	ingen	$\ln(1-y)$

5.2.2 Produktstrøm med 99 % C_{12}

5.2.2.1 Modell uten trykk variasjon

Trykket i kolonnen blir ikke regulert, og vil dermed kunne ha til dels stor påvirkning på de målte temperaturene i kolonnen. Simuleringene med konstant trykk i kolonnen ble foretatt for å kunne sammenlikne resultatene med simuleringene uten konstant trykk nivå. Kalibreringssettet var som for forsøkene gitt i Tabell 5-2, men uten trykkvariasjon. Modellene (se Tabell 5-5) som ble beregnet gav generelt lave RMSECV verdier ved få faktorer eller latente variable i PLS regresjonen. Modellene ble beregnet uten støy i målingene (n_o). Estimatorene basert på temperaturmålingene angitt ved $Temp_3$ med vektning med $1/SD_X$ (S_2) gav beste valideringsresultater.

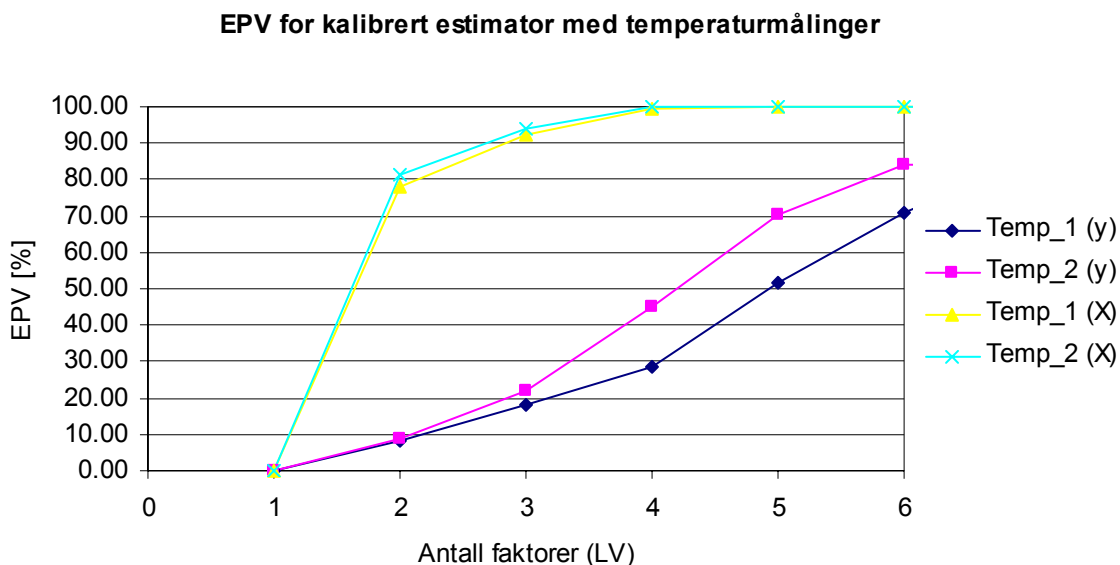
Tabell 5-9 Validerte verdier for estimatoren med inngangsvariable angitt ved $Temp_3$

Antall latent variabel (lv)	1	2	3
EPV, [%]	91.240	99.471	99.997
RMSECV [\pm m% C_{12}]	0.2	0.03	0.0001

Antall underliggende fenomener var dermed lite, og er kun ment å gi et bilde av situasjonen hvis trykket i kolonnen kunne blitt antatt konstant eller eventuelt med små endringer.

5.2.2.2 Estimator med temperaturmålinger

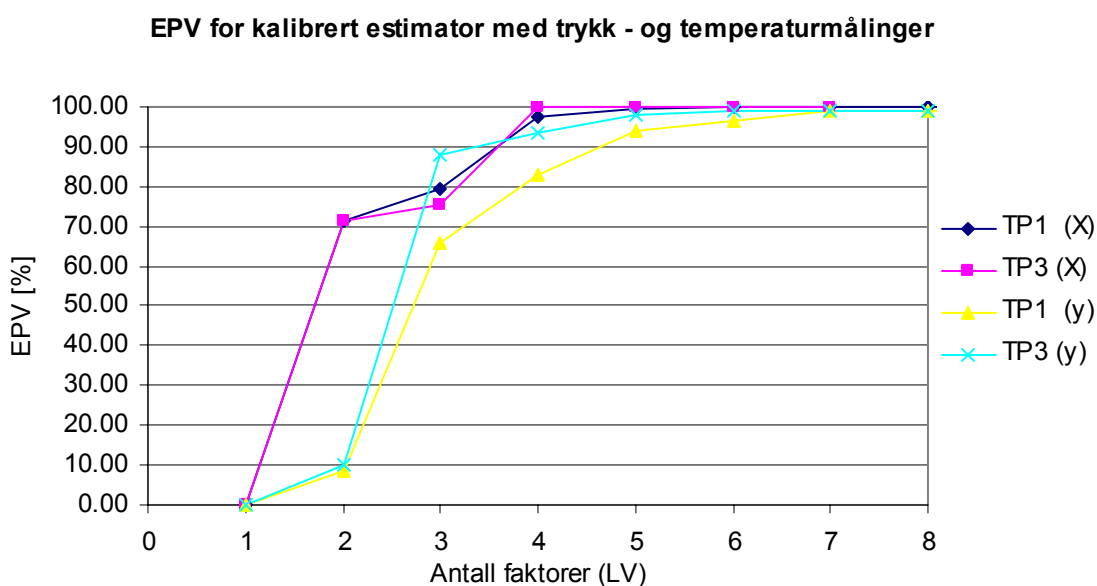
Destillasjonen ble deretter simulert med varierende trykket i kolonnen. Datasettet var som angitt i Tabell 5-2. Temperaturmålingene som inngikk i de beregnede estimatorene var tilsvarende som gitt i Tabell 5-5. Figur 5-2 viser den forklarte variansen i PLS regresjonen med målingene $Temp_1$ og $Temp_2$. Temperaturmålingene i modellene i Figur 5-2 var vektet med standardavviket (S_2). Den totale y variansen i kalibreringssettet ble ikke forklart og svært lite av X variansen gjenspeilte y dataene. Trenden var tilsvarende for de andre estimerte modellene, og modellene gav dermed generelt dårlige resultater kun med temperaturmålinger.



Figur 5-2 Forklart X og y-varians (C_{12}) i estimatorene ved målingene *Temp_1* og *Temp_2*

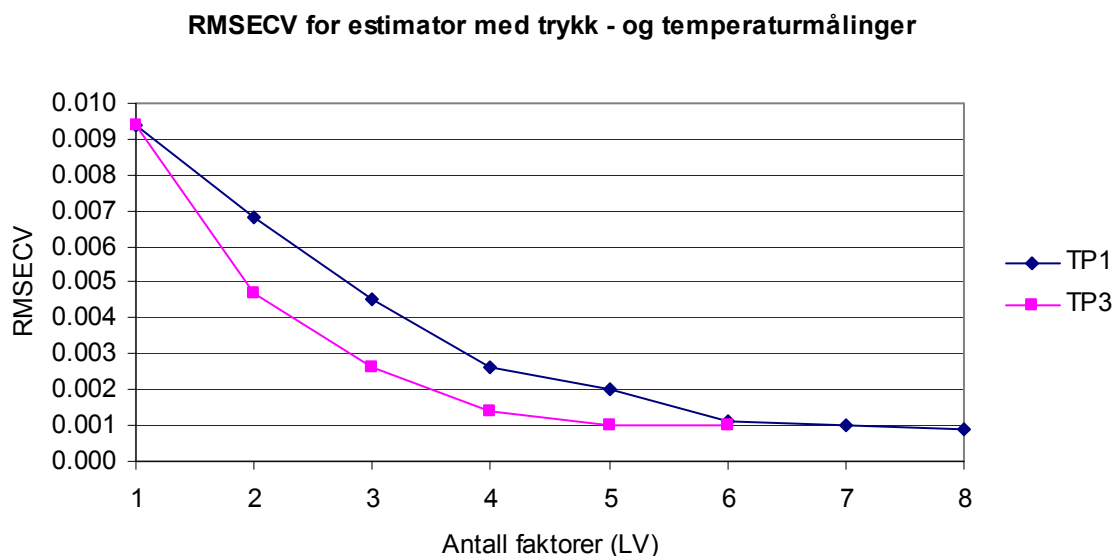
5.2.2.3 Estimator med temperatur og trykk målinger uten støy (n_0)

Kalibreringssettet for estimatorene var som gitt i Tabell 5-2. De kalibrerte og validerte modellene var enten perfekte målinger eller målinger påført støy, hvor dette er angitt. Estimatorene hadde i denne delen temperatur og trykkmålinger som inngangsvariable. X-matrisen (temperatur- og trykkmålingene) var vektet med standardavviket i de viste estimatorene (Figur 5-3 og Figur 5-4). Trykkvariasjonen i kalibreringssett var til dels stor, og trykkmålingene ble derfor antatt og bedre kunne beskrive variasjonen enn estimator med kun temperaturmålinger.



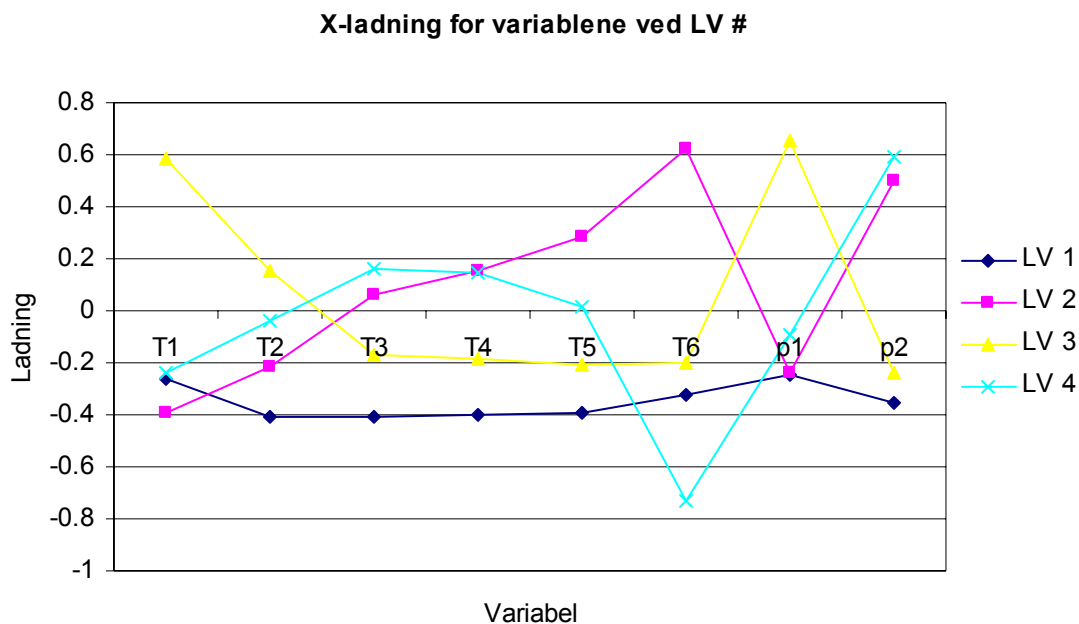
Figur 5-3 Forklart X og y varians ved antall faktorer i estimatorene med TP1 og TP3

Variansen for C_{12} i kalibreringssettet ble tilnærmet forklart 100 % ved 4 (*TP1*) og 6 LV (*TP3*), hvilket tilsvarte to mindre faktorer enn antall variable i estimatorene. RMSECV verdiene for de samme estimatorene gav følgende sikkerhet for fremtidige estimerte y :



Figur 5-4 RMSECV for modellene med målingene gitt ved TP1 og TP3

Et minimum for den forventede sikkerheten for estimerte verdier, i samme måleenhet som y , inntraff ved 5 faktorer for estimator med *TP3* målingene. En transformasjon av y gav ingen bedret forventet prediksjon (se Tabell 5-10). Antallet faktorer i estimatoren med målingene angitt ved *TP1*, hadde et optimum ved 6 faktorer. Ved 6 faktorer var den resterende variansen i X tilnærmet lik null (Figur 5-3) og dermed kunne en lav virkelig presisjonen være forventede for estimatoren pga av den økede modellusikkerheten. Figur 5-5 viser X -ladning for estimator *TP1* for de fire første latente variablene (LV). Den første faktoren beskriver mye av X -variansen, men lite av y -variansen. Den andre faktoren beskriver mye av y -variasjonen og gir følsomheten for endret sammensetning. Faktor tre gir endring i sammensetningen og innflytelsen av trykket på den målte temperaturen. Tabell 5-10 viser en oversikt over forklart varians (EPV) for de ulike estimatorene, og Bilag E gir en mer fullstendig oversikt en det som er presentert i dette avsnittet.


 Figur 5-5 X-ladningene for variablene i PLS estimator TP1 med n_0 og S_2 .

 Tabell 5-10 EPV % (y) for gitt antall faktor for estimatorene, skalering (S_i) og støy (n_0)

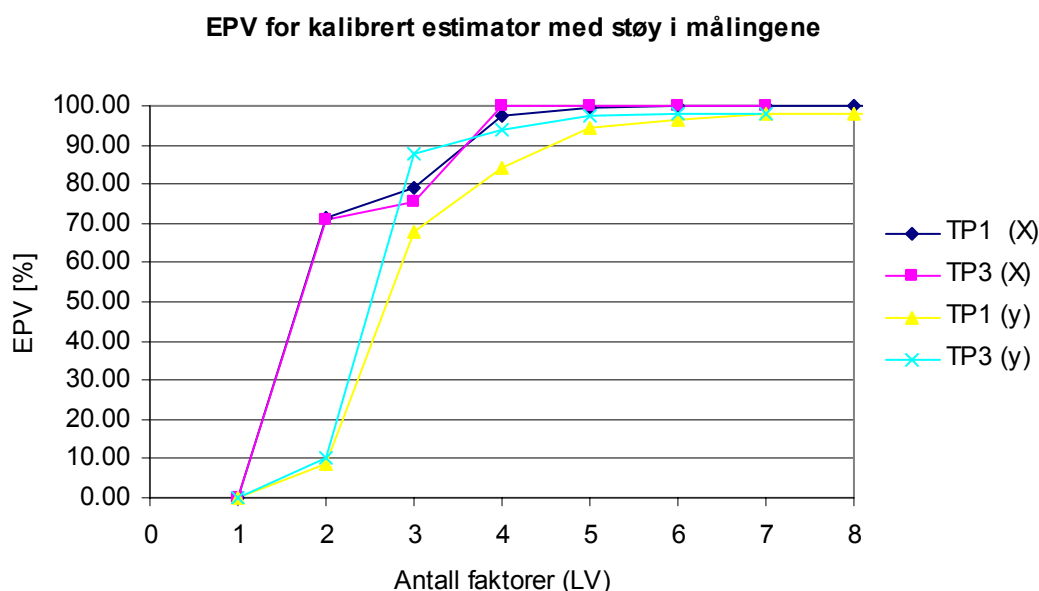
Estimator / S_i / n_0	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP1 / S_1 / n_0	6.38	31.41	77.80	80.39	87.43	98.33
TP1 / S_2 / n_0	8.47	66.00	82.80	93.76	96.63	99.03
TP1 / S_3 / n_0	8.29	70.31	83.95	95.29	97.17	97.46
TP2 / S_2 / n_0	8.90	93.54	93.78	95.96	98.98	99.13
TP3 / S_2 / n_0	9.86	88.07	93.52	98.08	99.06	99.17
TP4 / S_2 / n_0	9.42	93.80	94.62	96.06	99.01	99.15
TP5 / S_2 / n_0	10.29	93.26	93.42	97.97	98.89	
TP6 / S_2 / n_0	8.81	22.26	77.20	97.17	98.90	99.01
TP7 / S_2 / n_0	9.75	17.15	93.62	98.96	98.97	
TP8 / S_2 / n_0	8.85	29.38	87.68	91.43	98.00	98.04
TP3 / S_3 / n_0	9.77	85.46	94.65	97.20	97.79	98.30

Estimator / S_i / n_0	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP6 / S_3 / n_0	8.56	22.25	78.22	92.61	98.11	98.52
TP7 / S_3 / n_0	9.58	16.55	90.20	97.80	98.22	
TP8 / S_3 / n_0	8.56	31.32	91.49	94.40	96.17	97.68

Målingene i den virkelige prosessen vil alltid inneholde målestøy og måleusikkerhet. Resultatene fra de overnevnte modellene var basert på perfekte målinger, dvs. resultatene fra simuleringene var ikke påført et støy signal. En tilnærming til den virkelige prosessen og kalibrering med historiske data, ble gjort ved å introdusere støy i X matrisen (temperatur- og trykkmålingene). Støy i temperatur- og trykkdataene fra simuleringene ble derfor anvendt i PLS regresjonen i det følgende avsnittet.

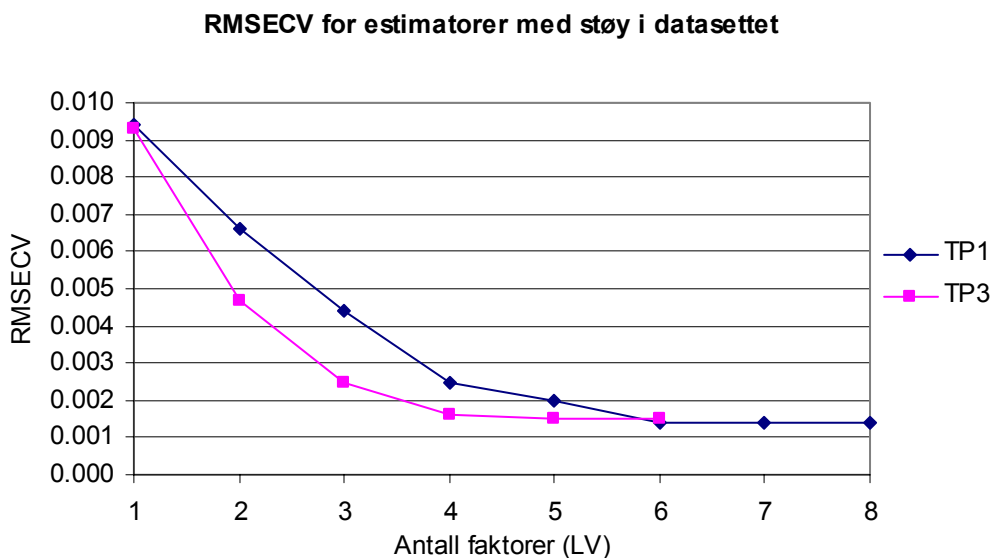
5.2.2.4 Estimator med temperatur og trykk målinger med støy (n_{1-6})

Støyen i datasettet (Tabell 5-2) ble påført verdiene ved en tilfeldig fordeling, som gitt i Tabell 5-7. Resultatene fra kalibreringen og valideringen av modellene med støy i datasettet er presentert i dette avsnittet. Bilag E gir en mer fullstendig oversikt en det som er presentert i dette avsnittet.



Figur 5-6 EPV for kalibrerte modeller ($TP1$ og $TP3$) med støy (n_i) i datasettet og skalering (S_2)

Den forklarte varians for y nådde sin høyeste verdi ved 7 LV for $TP1$, mens X variansen ble forklart ved 6 LV. Estimatoren med $TP3$ målingene hadde et maksimum for forklart y -variansen ved 5 latente variable.



Figur 5-7 RMSECV for de validerte estimatorene (*TP1* og *TP3* med n_4 og S_2)

Det optimale antall faktorer i estimatoren (*TP3*) var fra EPV og RMSECV verdiene 4 faktorer. Antall LV som burde inngått i estimatorene ble uforandret, men den validerte sikkerheten for estimerte verdier ble selvsagt lavere sammenliknet med resultatene kalibrering og validering uten støy i datasettet. Tabell 5-11 viser en oversikt over innflytelsen av støy i målingene for den forklarte variansen (EPV) for estimatorene.

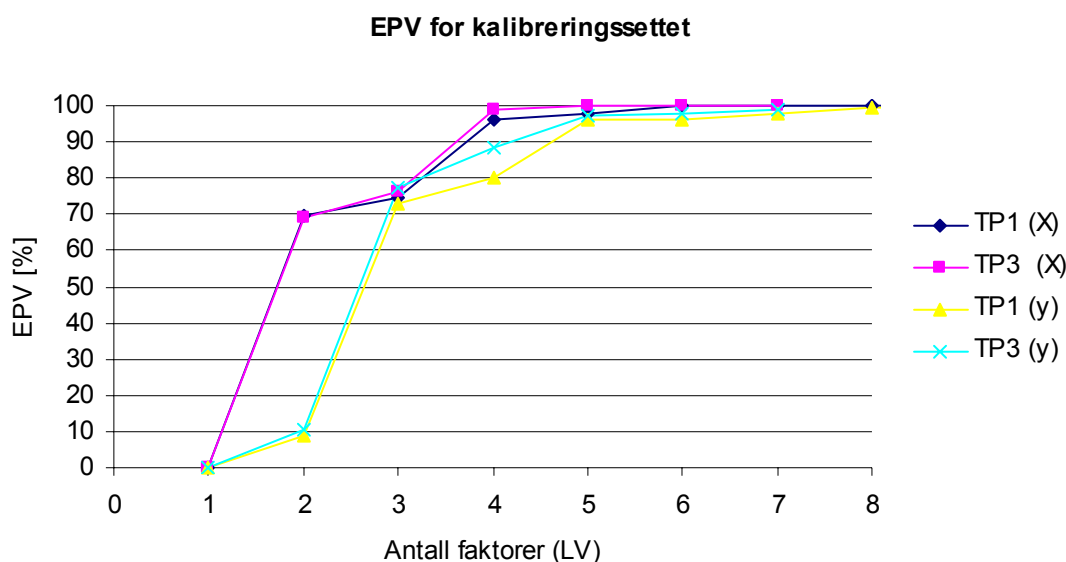
Tabell 5-11 EPV % (y) for gitt antall faktor for estimatorene, skalering (S_i) og støy (n_i)

Estimator / S_i / n_i	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP1 / S_1 / n_1	6.40	31.45	77.44	79.97	86.43	95.41
TP1 / S_2 / n_1	8.35	63.56	81.69	92.72	95.54	97.71
TP1 / S_3 / n_1	8.28	70.09	84.08	95.05	96.63	96.84
TP2 / S_2 / n_1	8.76	92.21	92.82	94.89	97.71	97.77
TP2 / S_3 / n_1	8.71	95.47	95.55	96.42	96.64	96.88
TP3 / S_1 / n_1	9.41	37.42	81.34	93.54	95.61	97.62
TP3 / S_2 / n_1	9.72	87.91	92.49	97.04	97.70	97.70
TP3 / S_3 / n_1	9.78	86.23	94.51	96.22	96.91	96.91
TP6 / S_2 / n_1	8.67	22.33	76.59	95.40	95.61	95.61

Estimator / S_i / n_i	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP6 / S_3/n_1	8.56	22.91	76.97	90.08	93.32	93.63
TP1 / S_2/n_4	8.56	67.85	84.00	94.44	96.63	98.17
TP1 / S_3/n_4	8.29	70.62	83.51	94.98	96.52	96.75
TP3 / S_2/n_4	10.00	88.00	94.01	97.60	98.02	98.02
TP3 / S_3/n_4	9.77	85.00	94.55	96.55	97.03	97.03
TP6 / S_2/n_4	8.90	23.37	77.94	94.04	94.60	94.62
TP1 / S_2/n_5	8.46	65.37	83.59	93.88	95.46	95.94
TP1 / S_3/n_5	8.24	69.88	82.57	93.78	95.61	95.98
TP3 / S_2/n_5	9.91	90.10	93.79	95.81	95.89	95.89
TP3 / S_3/n_5	9.69	83.60	93.05	95.17	95.44	95.44
TP1 / S_2/n_6	8.41	66.04	81.20	91.96	94.23	95.39
TP1 / S_3/n_6	9.77	85.60	91.69	94.97	95.18	95.19
TP3 / S_2/n_6	8.66	68.18	85.03	95.22	96.55	97.01
TP3 / S_3/n_6	10.19	87.94	94.19	96.59	96.95	96.95

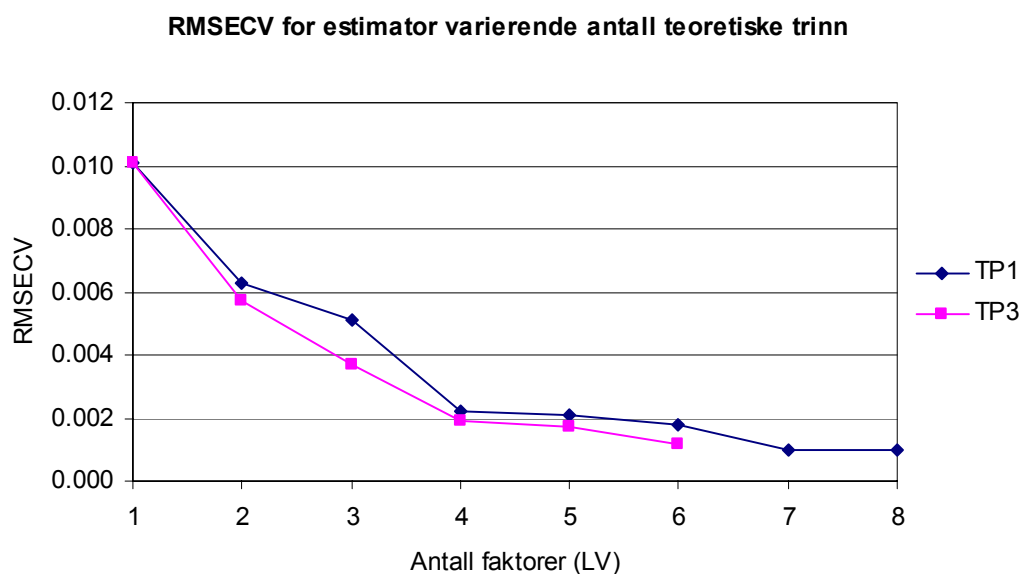
5.2.3 Variasjon i antall teoretisk trinn i kolonnen (n_0)

Antall teoretiske trinn i kolonnen ble bestemt ved en enkel analyse av kolonnen. Antall reelle trinn i kolonnen med strukturert pakning vil kunne variere med varierende last og driftsbetingelser. Simuleringer med forskjellig antall teoretiske trinn (21 og 20) ble foretatt med forsøkene gitt ved forsøksdesign i Tabell 5-2, men med variasjon av antall teoretiske trinn. Estimatorer med effekten av en eventuell endring i antall teoretiske trinn i kolonnen ble beregnet.



Figur 5-8 Forklart varians for estimatorene $TP1$ og $TP3$ (n_0 og S_2) med variasjon i antall teoretiske trinn

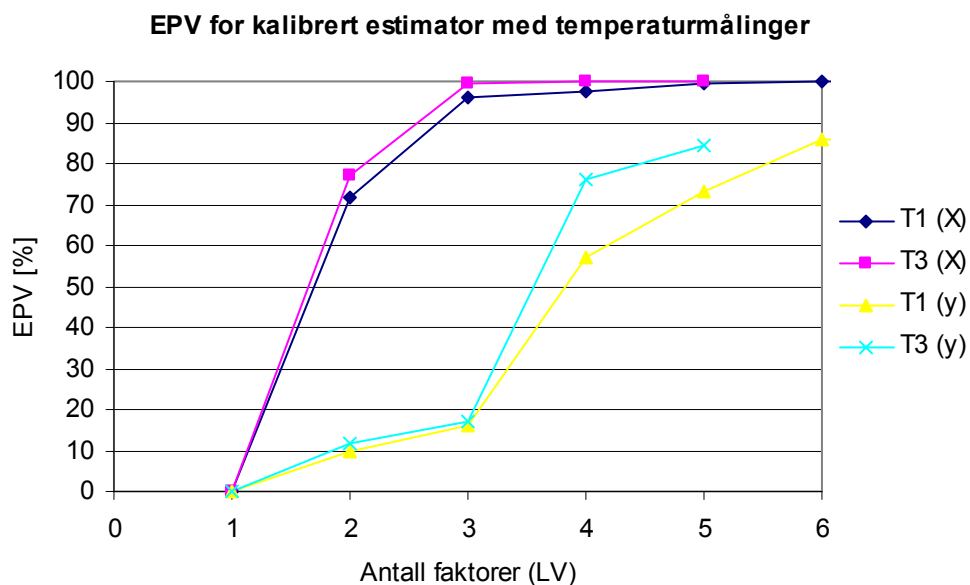
Den forklarte variansen lå som forventet, pga modellusikkerheten, lavere enn for estimatorene uten variasjon av antall teoretiske trinn.



Figur 5-9 RMSECV verdiene ved antall LV i estimatorene $TP1$ og $TP3$ (n_0 og S_2)

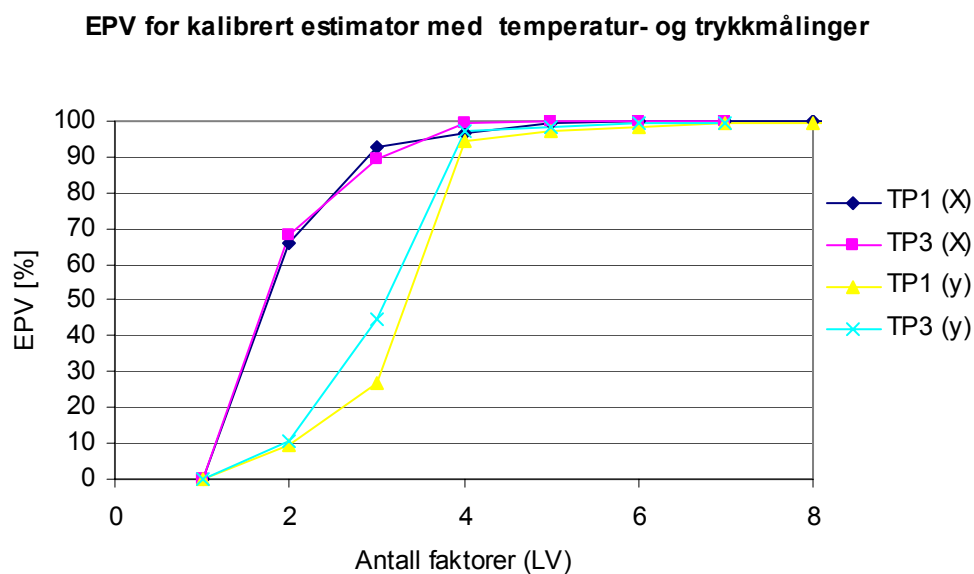
5.2.4 Estimator for 95 % renhet av C₁₂ topp-produktet

Spesifikasjonen for innholdet av C₁₂ i topp-produktet kunne bli redusert til 95 %. En estimator for denne spesifikasjonen ble derfor beregnet. Et lavere innhold av C₁₂ i produktstrømmen kunne tilfredsstillte andre krav og dermed bli solgt til andre marked enn det som var aktuelt for 99 % laurinsyre i produktet.

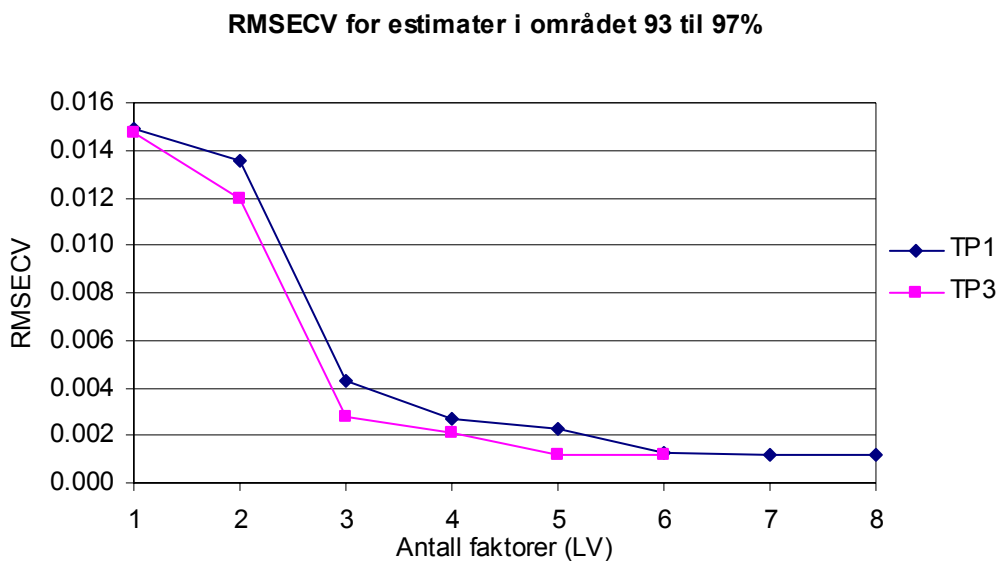


Figur 5-10 Estimator med temperaturmålinger for C₁₂ i produktstrømmen i området 93-97%

PLS modellene med kun temperaturmålinger viste, som i tidligere resultater, en lav forklart varians for innholdet av laurinsyre i destillatet. Resultatene for estimator med temperatur og trykk inngangsvariable viste en betydelig forbedring av forklart varians.



Figur 5-11 Estimator med trykk- og temperaturmålinger

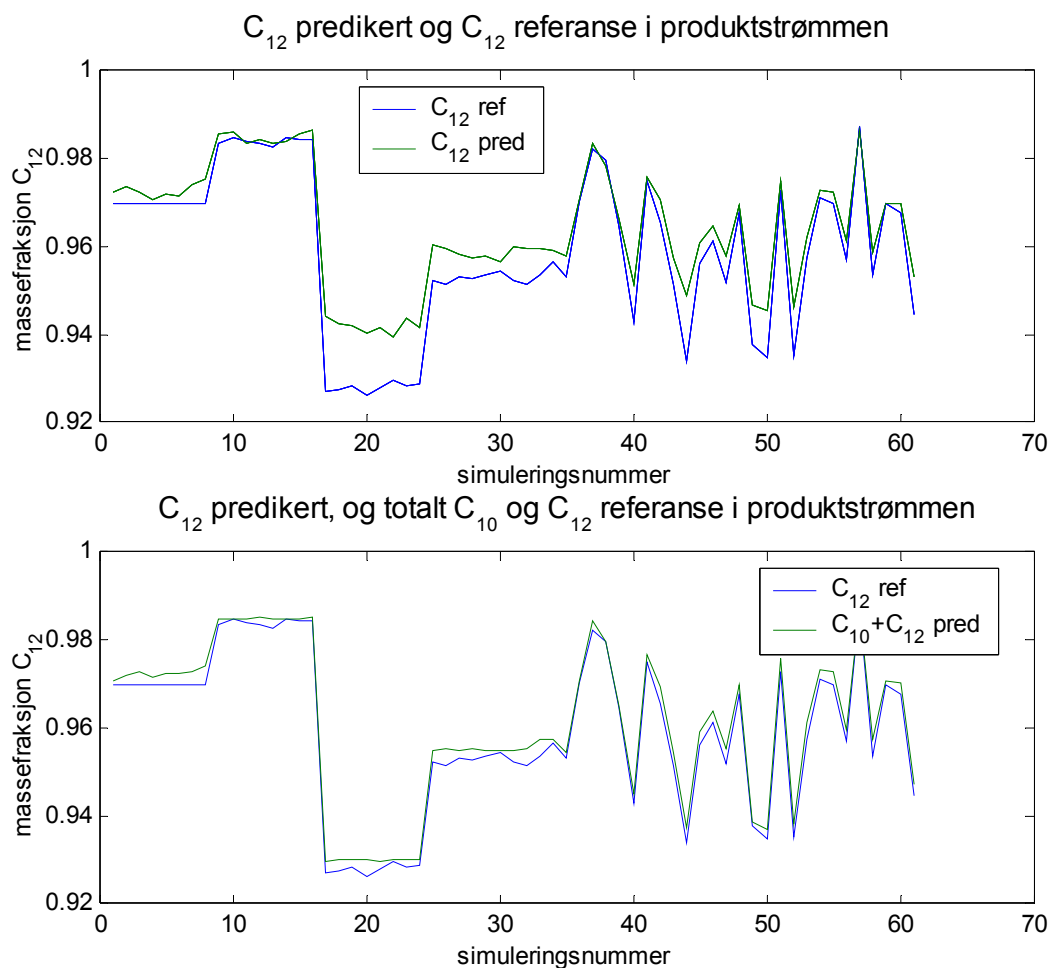


Figur 5-12 RMSECV verdier ved antall LV i estimatorene

Det optimale antall faktorer i estimatorene var fire for TP3 og seks for TP1.

5.2.5 Følsomhet for C_{10} i fødestrømmen til kolonne 2

Innholdet av C_{10} (kaprinsyre) ble i de foregående beregningene antatt lik null. Analyser fra januar 2002 viste at fødestrømmen inneholdt C_{10} i slike mengder at 99 % renhet av produktet ikke kunne bli oppnådd. To dager i februar 2002, hvor det forelå analyseresultater av strømmene, var det om lag 99 % C_{12} i produktstrømmen. Fødestrømmen til tårn 2 inneholdt også for disse to dagene en liten mengde C_{10} . Fødestrømmen kunne derfor også i fremtiden antas å inneholde noe C_{10} . PLS-estimatorens følsomhet for C_{10} i føden ble derfor undersøkt ved å estimere C_{12} (\hat{y}) i produktet på grunnlag av resultater fra simuleringer med C_{10} i føde- og produktstrømmen.



Figur 5-13 Predikerte og referanse verdier for massefraksjon C₁₂ i produktstrømmen ved C₁₀ i føden. (øverst) virkeliginhold av C₁₂ (blå) og C₁₂ estimert (rød), (nederst) virkeliginhold av C₁₂ + C₁₀ (blå) og C₁₂ estimert (rød)

De estimerte verdiene for sammensetningen gav et totalt estimat for innholdet av C₁₂ og C₁₀ i produktstrømmen. Estimatoren var ikke kalibrert for innhold av C₁₀ i føden, og de estimerte verdiene lå dermed utenfor forventningsrommet til estimatoren.

5.3 Estimatorene for 99 % og 95 % C₁₂ i produktstrømmen

De ulike estimatorene ble vurdert ved de presenterte kriteriene (forklart varians og RMSECV). De ulike estimatorene var vektet og skalert eller transformert i ulik grad. Estimatorene som ut fra kalibrering og validering av resultatene fra PLS regresjonene ble vurdert å være best for de to renhetsgradene av destillatet, blir presentert i denne delen.

5.3.1 Estimatoren for 99 % C₁₂ i produktstrømmen

Målingene som inngikk i den beste estimatoren var angitt ved

$$\hat{y} = f(T1, T2, T3, T4, p1, p2) \quad (5.5)$$

Inngangsvariablene i estimatoren var sentrerte og vektet med $1/SD_X$ (standardavviket for variablene i X , dvs. temperatur- og trykkmålingene, se ellers avsnitt 3.4.5). C_{12} , massefraksjon, var sentrert. Verdiene for sentrering og vekting av variablene er gitt i tabellen nedenfor.

Tabell 5-12 Gjennomsnitt og standardavvik for variablene i estimatoren

Variabel / Måling	Gjennomsnitt \bar{x} / \bar{y}	Standardavvik SD_x
T ₁	153.37	1.4915
T ₂	158.69	1.6697
T ₃	168.04	2.4739
T ₄	170.56	2.6403
p ₁	5.00	0.4075
P ₂	19.06	3.2963
C ₁₂	0.9831	

PLS estimatoren for kort prediksjon ble i avsnitt 3.4.4.4 gitt ved

$$\hat{y}_i = \hat{b}_0 + x_i^T \hat{b} \quad (5.6)$$

hvor $\hat{b}_0 = \bar{y}$ og $\hat{b}_{i=1..6}$ var gitt ved

$$\hat{b} = \begin{bmatrix} -0.4493 & -2.705 & -0.2918 & 0.7539 & 2.086 & 1.148 \end{bmatrix}, (\hat{b}_{i=1..6}) \quad (5.7)$$

Estimeringen av C_{12} kunne eventuelt være gjort ved en fullstendig PLS prediksjon (avsnitt 3.4.4.3. Verdier og eksempel for fullstendig prediksjon er gitt i bilag D.

5.3.2 Estimatoren for 95 % C_{12} i produktstrømmen

Målingene som inngitt i den beste PLS estimatoren for 95 % C_{12} var tilsvarende som for estimatoren gitt for 99 % C_{12} i likning (5.5).

Tabell 5-13 Gjennomsnitt og standardavvik for variablene i estimatoren for 95 % C₁₂ i produktet

Variabel / Måling	Gjennomsnitt \bar{x} / \bar{y}	Standardavvik SD_x
T ₁	154.17	1.5974
T ₂	160.31	1.7134
T ₃	169.54	2.4451
T ₄	171.81	2.6551
p ₁	5.00	0.4281
p ₂	19.28	3.3769
C ₁₂	0.9516	

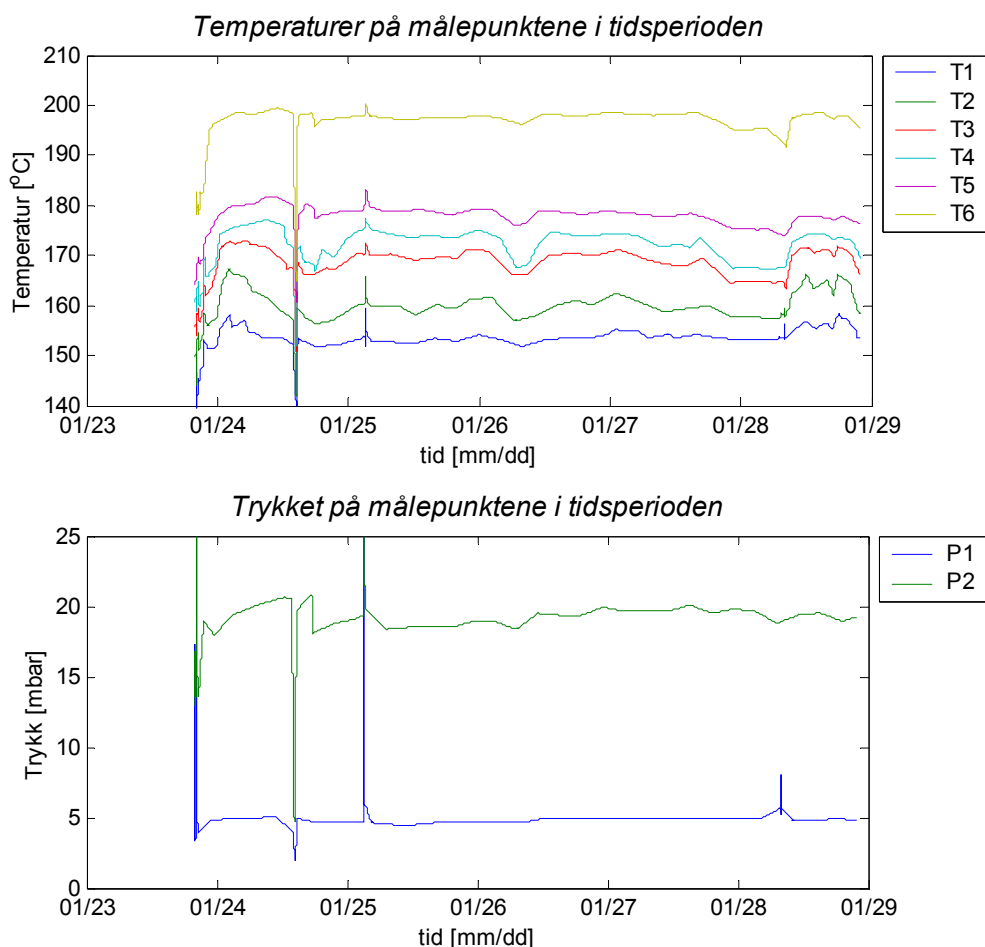
\hat{y} var gitt ved likning (5.6), hvor $\hat{b}_0 = \bar{y}$ og $\hat{b}_{i=1..6}$ var gitt ved

$$\hat{b} = \begin{bmatrix} -1.541 & -3.474 & 0.4044 & 0.5043 & 3.527 & 1.208 \end{bmatrix} \times 10^2, \left(\hat{b}_{i=1..6} \right) \quad (5.8)$$

Estimeringen av C₁₂ kunne eventuelt være gjort ved en fullstendig PLS prediksjon (avsnitt 3.4.4.3). Verdier og eksempel for fullstendig prediksjon er gitt i bilag C

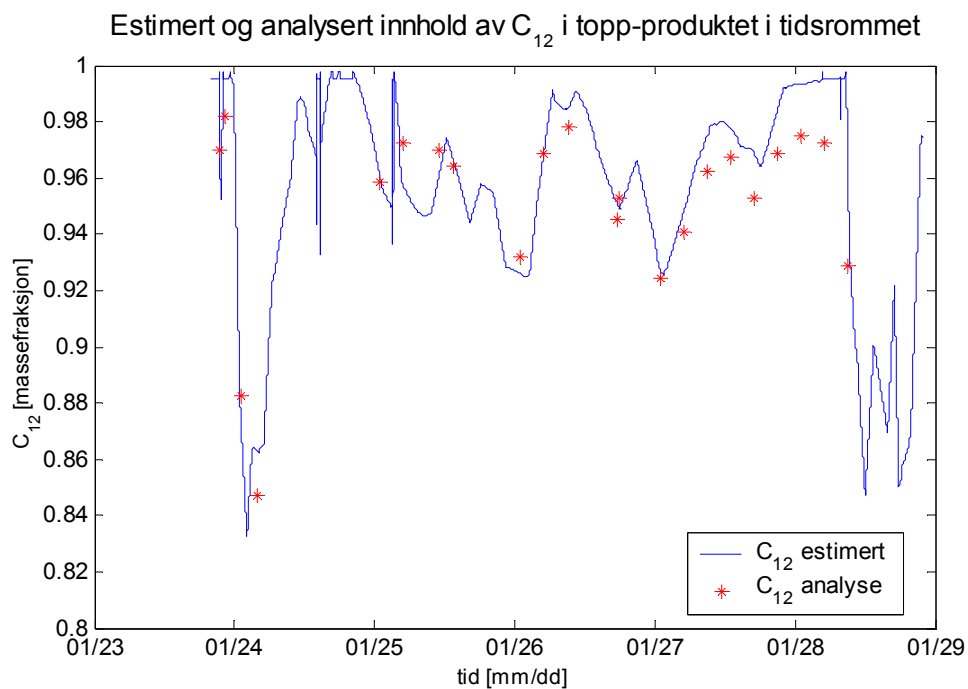
5.4 Test av estimeringsevnen til den beste estimatoren på historiske data

Det forelå, som tidligere nevnt, historiske data for tårn 2 for perioden 23. til 28. januar 2002. De historiske dataene (logg for trykk og temperatur målingene) gav en indikasjon på prosessdynamikken, når de ble sett i sammenheng med plot av historiske data for prosessstrømmene. De loggede verdiene var ikke de eksakte verdiene på tidspunktet, men var trend verdier som forklart i kapittel 5.1.4. Figur 5-14 viser temperatur og trykk på målepunktene i den aktuelle perioden.



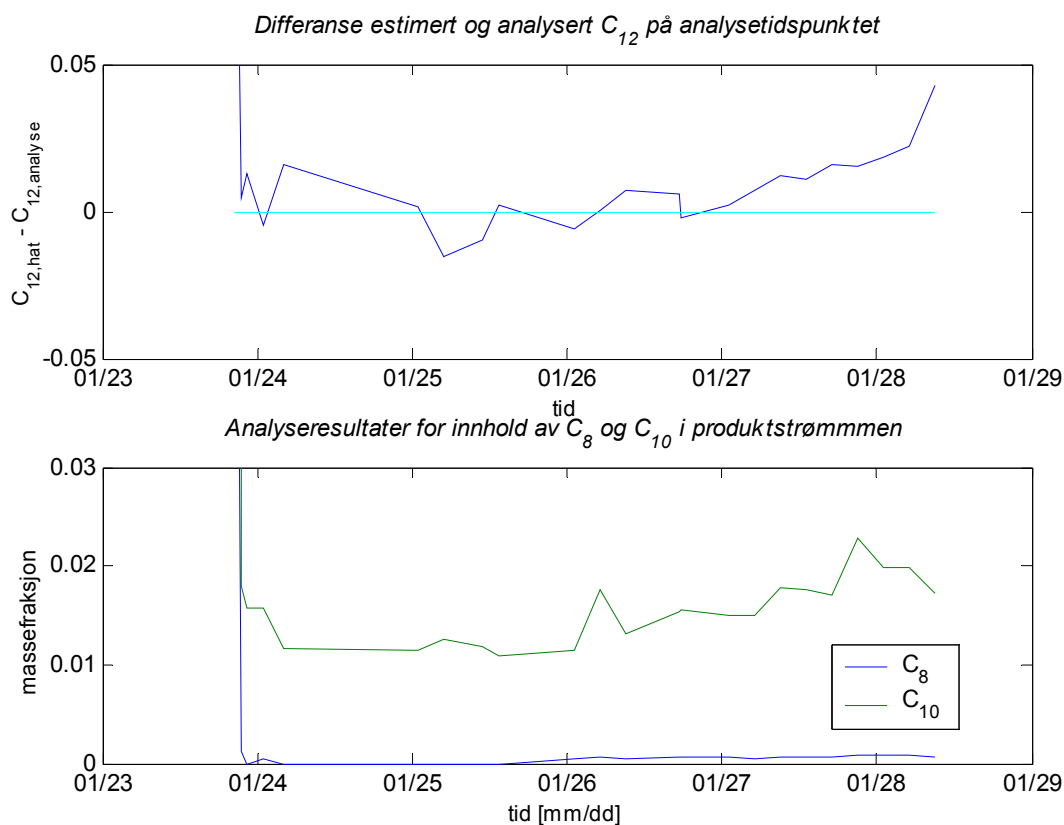
Figur 5-14 Historiske data for prosessen i tidsperioden 23. til 28. januar 2002

Estimerte verdier for C_{12} i destillatet i den overnevnte perioden ble beregnet ved å anvende estimator for 99 % C_{12} , som ble gitt i avsnitt 5.3.1, og de historiske dataene for perioden (plottet i Figur 5-14). I plottede historiske dataene ble det observert enkelte pulser i trykkmålingene og en viss fluktuasjon i målte temperaturene. I Figur 5-15 ble de estimerte verdiene plottet i sammen diagram som de analyserte verdiene for C_{12} i produktstrømmen.



Figur 5-15 Estimerte verdier for C_{12} ved historiske dataene, og analysert innhold av C_{12}

De estimerte verdiene avvek noe fra de analyserte verdiene, men avviket mellom de to verdiene synes allikevel å være akseptabelt tatt i betraktning datagrunnlaget. De estimerte verdiene synes å ligge noe før de analyserte verdiene. Differanse mellom \hat{y} og y og en mulig underliggende årsak for noe av avviket blir vist i Figur 5-16.



Figur 5-16 (Øverst) Differanse mellom estimert og analysert C_{12} på analysetidspunktet. (Nederst) analysert innhold av C_8 og C_{10} i produktstrømmen

Det analyserte innholdet av C_8 og C_{10} i produktstrømmen (nederst i figuren Figur 5-16) viste at det var mye C_{10} i produktet i perioden 23. til 28. januar 2002. Differansen mellom \hat{y} og y viste en viss kovarians med innholdet av lette komponenter i produktet. Dette samsvarte med resultatene i avsnitt 5.2.5.

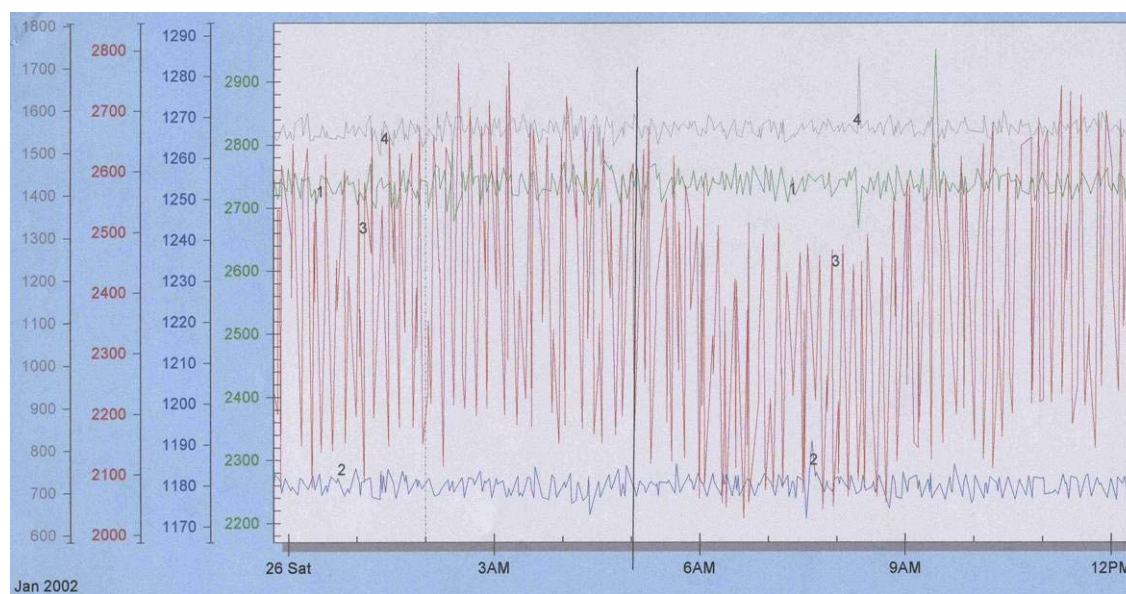
6 Prosessbetraktning

6.1 Reguleringsstruktur for kolonne 2

I forbindelse med arbeidet med oppgaven ble det observert noen fenomener eller problemer som ikke direkte har tilknytning til den definerte oppgaven, men som allikevel vil ha direkte innflytelse på vurderingen av estimatoren og den endelige konklusjonen. Problemene blir derfor presentert i korte trekk i dette kapitlet.

En skjematisk oversikt over anlegget ble gitt i Figur 4-2. Reguleringsstrukturen for kolonne-2 i dagens anlegg består av nivåregulering av refluksakkumulatoren. Nivået i akkumulatoren ble regulert ved å justere refluksstrømmen til kolonnen. De to vanligste formene for nivåregulering av kolonnen er LV-konfigurasjonen (L og V uavhengige variable) og DV-konfigurasjonen (D og V uavhengige variable). DV var konfigurasjonen som ble benyttet for regulering av nivået i refluksakkumulatoren.

Historiske data for prosesstrømmen i anlegget blir vist i figuren nedenfor, hvor fødestrømmen var kalkulert på grunnlag av en massebalanse for kolonne 1.



Figur 6-1 Målte og kalkulerte verdier for føde-(grønn), bunn-(grå), topp-(blå) og refluks- (rød) strømmer 26. januar 2002 (kl. 24.00 til kl.12.00)

6.2 Betraktninger av reguleringen

Plottet av de historiske dataene i Figur 6-1 viste noen klare tendenser. Refluksstrømmen oscillerte med noe som lignet på sinusformet respons. Frekvensen var i følge plottet om lag ni minutter og amplituden var $\pm 10\%$ av det som hadde vært en stasjonærverdi for prosessen. Denne oscillasjon kunne skyldes feil tuning av regulatoren eller feil valg av reguleringsstruktur. Hovedårsaken for oscillasjonen kunne, hvis en antar at regulatoren for sløyfen var en PI-regulator, være for kort integral tid og for lav proporsjonal virkning [Ref VI]. En ren P-regulator kunne isteden vært anvendt med tilstrekkelig stort proporsjonalvirkning. Oscillasjonen ville da bli redusert, men et avvik lite avvik ($\pm 5\%$) ville kunne oppstå for væsknivået i akkumulatoren [Ref XXII]. Kokeren for kolonnen står på sin maksimale belastning, og en tettere regulering av refluksstrømmen vil sannsynlig senket den maksimale belastningen på kokeren.

7 Diskusjon

Beregning av estimator for sammensetningen i produktstrømmen i toppen av kolonne 2, med hensyn på laurinsyre (C_{12}), var hovedoppgaven i denne hovedoppgaven. Kilder til eventuell usikkerhet i estimatoren, og dermed fremtidige estimerte verdier, vil være en kombinasjon av flere årsaker. I den første delen av diskusjonen blir det derfor drøftet fundamentale elementer vedrørende simuleringen av destillasjonskolonnen. Videre blir estimatorene som ble presentert i avsnittene 5.2 - 5.4 vurdert.

7.1 Modell av kolonnen i Hysys

7.1.1 Antall komponenter i fødestrømmen til kolonnen

Antallet komponenter i fødestrømmen til kolonnen ble under defineringen av oppgave og i simuleringene begrenset til fire fettsyrer ($C_{12} - C_{18}$). Analyser av sammensetningen gjort i januar 2002 viste at massestrømmen virkelig inneholdt både tynge og lettere komponenter enn de som inngikk i simuleringene. Fraksjonen av lettere komponenter enn C_{12} varierte da innenfor området $\langle 0.001, 0.018 \rangle$ (se bilag F). Målet for separasjonen i kolonnen 2 er 99 %, masseprosent, laurinsyre i produktet. Dette medfører at komponenter lettere enn C_{12} i fødestrømmen må senkes til et minimum, men denne forutsetningen er p.t. ikke oppfylt. Den fremtidige sammensetningen vil måtte oppfylle begrensningene gitt for estimatoren.

7.1.2 Kilder til modellfeil i Hysys

7.1.2.1 Valg av termodynamikk

De fysiske egenskapene til fettsyrene ble i modellen beregnet med væske aktivitetsmodellen Wilson likningen, og manglende parametere ble beregnet med UNIFAC. Damp – væske likevektsdata for binære blandinger av de ulike fettsyrene var undersøkt eksperimentelt og gjengitt i litteraturen [Ref XIV]. De eksperimentelle dataene ble sammenliknet med tilsvarende simulerte sammensetninger, og overensstemmelsen med de eksperimentelle dataene ble vurdert å være tilfredsstillende for oppgaven. En noe bedre tilnærming av de fysiske egenskapene kunne antagelig blitt oppnådd hvis tilstandslikningen, PHC-EOS, som ble foreslått av Aly et. al. [Ref XV] var blitt anvendt. Kriteriene for Aly var anvendelse av samme termodynamiske modell for under- og overatmosfæriske trykk. Applikasjonen og behovet for et bredere gyldighetsområde avvek fra kriteriet i denne oppgaven, som kun krevde gyldighet innenfor vakuuområdet. Aly etterlyser mer nøyaktig termodynamisk informasjon som bedre kunne beskrive den kjemiske assosiasjonen for komponentene. Modellfeilen ved valg av termodynamisk modell (Wilson) ble dermed vurdert å være en lite påvirkende feilkilde i med tanke på den nevnte alternative termodynamikken (PHC-EOS) og den totale feilkilden i prosessmodellen og estimatorene.

7.1.2.2 Antall teoretiske trinn i destillasjonskolonnen

Antall teoretiske trinn i destillasjonskolonnen var tidligere bestemt og benyttet i simulering av destillasjonskolonnen av Morten Gulbrandsen [Ref XVI]. Simuleringen var benyttet i diverse arbeid internt i bedriften. Antallet teoretiske trinn i kolonnen ble derfor antatt å være tilfredsstillende tilnærmet. Antallet teoretiske trinn i kolonnen ble også undersøkt ved simuleringer beskrevet i avsnitt 4.3.3. Gill et. al. [Ref XVIII] har beskrevet at antall trinn i kolonnen beregnet ved HETP (høyde ekvivalente teoretiske trinn) kunne variere for de ulike komponentene som ble separert i kolonnen. Trinnvirkningsgraden for separasjon av flerkomponentblandinger vil være forskjellig for de ulike komponentene (avsnitt 4.3.3) [Ref XVII]. Denne effekten ville være en ekvivalent til ulike HETP-verdier for komponentene, som var påpekt av Gill. Fastsettelse av eventuelt ulikt antall teoretiske trinn for komponentene og anvendelse i simuleringmodellen ble allikevel ikke vurdert i oppgaven, da det ikke ble funnet noe teoretisk grunnlag som indikerte størrelsesorden for en slik antagelse. Utgangspunktet for oppgaven var også modellering av kolonnen med ideelle trinn, tilsvarende lik antall teoretiske trinn, og dermed trinnvirkningsgrad lik en. Antallet teoretiske trinn kan allikevel avvike fra det antatte antall teoretiske trinn i kolonnen. Dette kan enten skyldes beregningsfeil, feil i antagelsene, eller fysiske forhold som gass og væske last til systemet. Ved lavere massestrømmer i kolonnen vil dermed antall teoretiske trinn bli redusert i forhold til utleggingspunktet. Dette ble derfor vurdert i avsnitt 5.2.3.

7.1.2.3 Trykket i kolonnen

Trykkprofilen i kolonnen ble antatt å være en lineær sammenheng av trykket toppen og bunnen av kolonnen. Trykket ble ikke målt flere steder i kolonnen. Antagelser om den virkelige trykkprofilen for kolonnen, og dens avhengighet var derfor vanskelig å gjøre fra de tilgjengelige dataene fra anlegget. Tverrsnitt for den nedre delen er mindre enn den øvre delen av kolonnen, men viktigst er væskestrømmen størst i denne delen av kolonnen. Denne delen av kolonnen vil derfor sannsynligvis ha en relativt større trykkfall sammenliknet med toppseksjonen. Denne effekten var dessverre vanskelig å kvantifisere, og dermed ble dette ikke benyttet i modellen, med den feilen dette eventuelt vil medføre. Trykkfallet over kolonnen vil normalt være av høyere orden enn en (1), uttrykt ved for eksempel massestrømmen for oppkoket, og dermed en ulinear karakter. Kolonnen er delt inn i tre seksjoner med strukturerte pakninger, og trykkfallet vil kunne være forholdsvis større ved væskesamlerne og fordelerne som står før og etter pakningsseksjonene.

Den lineære trykkprofilen, som var uavhengig av føderaten, vil kunne føre til en relativ stor feilkilde i modellen, hvis denne avviker betraktelig fra den virkelige trykkprofilen. Simuleringer med spesifisering av fødesammensetningen og sammensetningen i topp- og bunnproduktet, med utgangspunkt i analyserte sammensetninger og sammenlikning med målte temperaturer i anlegget, viste ingen for store avvik mellom de målte temperaturene og temperaturene på tilsvarende trinn fra simuleringene. Modellfeilen i forbindelse med trykkprofilen for kolonnen blir derfor vurdert å ikke være en u håndterlig feilkilde. Den mest innlysende effekten av et lineært trykkfall over kolonnen er en mindre ulinearitet for temperaturene på trinnene i kolonnen, som danner den såkalte temperaturprofilen.

7.1.2.4 Plassering av temperaturmålingene

Temperaturen blir målt på flere punkter i den pakkede kolonnen. En riktig plassering av målepunktene i forhold til de teoretiske trinnene, som ble benyttet i den modellerte kolonnen, vil kunne være av avgjørende betydning for det endelige resultatet eller overførbarheten av de beregnede estimatorene. Plasseringen av temperaturmålingene i toppen av kolonnen, ved fødepunktet, i refluksakkumulatoren og i bunnen av kolonnen ble gjort med stor sikkerhet. Plasseringen av målepunktene som befant seg mellom fødepunkt og toppen av kolonnen ble implementert med noe usikkerhet. Punktene var i utgangspunktet plassert med tanke på en jevn distribusjon av observasjonene, og ikke i forhold til eventuelle teoretiske trinn. Overensstemmelsen mellom virkelig plassering og plassering i modellen ble prøvd ved simuleringer (se avsnitt 5.1.6) og omtrentlig plassering foretatt av Gulbrandsen [Ref XVI]. Målepunktene og plasseringen av disse på de teoretiske trinnene i blir vurdert å være en kilde til feil, som ikke kunne bli redusert uten mer rigorøs modellering av kolonnen. Med rigorøs blir det ment en modellert kolonne med trinnvirkningsgrad for komponentene eller eventuelt kolonne med strukturert pakning. Differanse mellom virkelig og tilpasset plassering av temperaturmålepunktene vil føre til en systematisk bias feil og en feil pga ulineære respons.

7.1.3 Tilgjengelig informasjon om prosessen

Modellering av kolonne-2 anlegget ble gjort med bakgrunn i kjent eller tilgjengelig informasjon om kolonnen og prosessen. Verifiseringer av modellen og modellresultatene ble også foretatt på grunn av tilgjengelig informasjon. Mangelfull eller manglende data og kommunikasjon kan i denne oppgaven ha ført til noe dårlig forutsetninger for å lage en modell med lavest mulig modellusikkerhet eller eventuelt forventet nøyaktighet av resultatene. Empiriske data (resultatene) og forhåndskunnskap ble benyttet til kvantitativt å estimere av informasjon Y fra tilgjengelige målinger X , gjennom en matematisk transfer funksjon". Multivariat kalibrering innebar dermed fastsettelse av hvordan en skal bruke de målte variablene $x_1, x_2 \dots x_n$ samtidig i en kvantifisering av den søkte verdien av y . Den eventuelle usikkerhet som oppsto i prosessmodellen vil være forplantet i estimatorene som ble beregnet ved PLS regresjon.

7.1.4 Neglisjerte effekter

Estimatorene som ble beregnet ved PLS regresjonen gav generelt gode resultater. Simuleringsmodellen for kolonnen var statistisk hvor effekten av variasjon av noen parametere og variable var neglisjert. Føden ble i alle tilfeller antatt å være væske på sitt kokepunkt og trykket i fødestrømmen lik trykket på fødepunktet i kolonnen. Effekten av underkjølingen av refluxen til kolonnen ble ikke tatt med i modellen. Trykket i kolonnen ble antatt å variere lineært (diskutert i avsnitt 7.1.2.3.), og i tillegg kunne det forekomme lokale trykkfall og liknende i kolonnen. De nevnte variasjonene ville kunne påvirke temperaturprofilen i kolonnen, hvilket også ble nevnt av Mejdell [Ref II], men det ble antatt at en eventuell variasjon av de nevnte tilfellene ville hatt liten total effekt på resultatene. Dette vil være gyldig for målingene, som inneholder støy, hvor effekten av variasjonen vil være vanskelig å skille fra annen støy i målingene.

7.2 Estimator for produktsammensetningen

7.2.1 Modell uten trykk variasjon

Estimatoren for sammensetningen basert på prosess-simuleringer uten trykkvariasjon og støy gav som ventet lav prediksjon, da trykket ikke var en frihetsgrader i destillasjonen og det er lite ulinearitet. Estimator ved *Temp_3* gav de beste resultatene, med 99.5 % EPV allerede ved to latente variable.

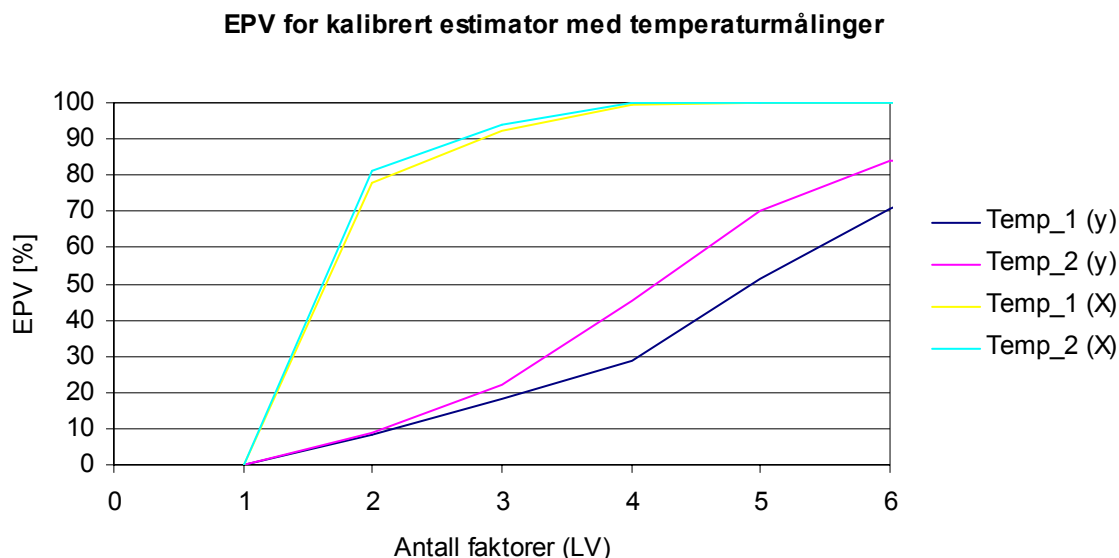
Tabell 7-1 Validerte verdier for estimatoren med inngangsvariable angitt ved *Temp_3*

Antall latent variabel (lv)	1	2	3
EPV, [%]	91.240	99.471	99.997
RMSECV [\pm m% C ₁₂]	0.2	0.03	0.0001

7.2.2 Modell med trykk variasjon

7.2.2.1 Estimator med temperaturommålinger

Estimatorene som kun benyttet temperaturmålinger i estimatorene viste lav forklart varians (EPV). Temperaturen blir målt på seks punkter i kolonnen, med fire av disse over fødepunktet. Korrelasjonen mellom sammensetningen og temperaturen avtar med økende avstand til toppen av kolonnen og dermed topp-produktstrømmen. Temperaturen under fødepunktet korrelerte derfor dårlig den observerte sammensetning, og førte dermed ikke til en bedre modell. En transformering av variablene, som har blitt foreslått av Mejdell [Ref II], ville sannsynlig ikke forbedret estimatene. Hvis antallet temperaturmålinger i kolonnen hadde vært høyere, kunne situasjonen blitt vurdert annerledes.



Figur 7-1 Forklart X og y-varians (C_{12}) i estimatorene ved målingene *Temp_1* og *Temp_2*

7.2.2.2 Estimator med temperatur- og trykkobservasjoner

Estimator med både temperatur og trykkmålinger viste langt høyere EPV enn tilsvarende uten trykkmålinger. Den forklarte variansen ved en faktor i PLS regresjonen var lav (se Tabell 7-2), hvilket kom av trykkvariasjonen i kolonnen. Variasjon i trykket vil ikke være sterkt korrelert med sammensetningen, og trykkvariasjonen vil hovedsakelig bli forklart i den første faktoren. *TP1* var basert på alle tilgjengelige målinger i kolonnen, mens *TP3* hadde målinger for trykk og temperaturer overføddepunktet. *TP3* viste generelt høyest verdier for den forklarte variansen, og det optimale antallet faktorer i estimatoren var fire. Flere faktorer enn fire ville medført lite forbedring av EPV (+ 1 %). En generell anbefaling for antall faktorer i PLS regresjonen er at ytterligere faktorer i estimatoren bør forklare 2 % eller mer hvis den skal være signifikant eller vil forbedre resultatene.

Tabell 7-2 EPV % (y) for gitt antall faktor for estimatorene (99 % C_{12}), skalering (S_i) og støy (n_o)

Estimator / S_i / n_o	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP1 / S_2/n_0	8.47	66.00	82.80	93.76	96.63	99.03
TP1 / S_3/n_0	8.29	70.31	83.95	95.29	97.17	97.46
TP3 / S_2/n_0	9.86	88.07	93.52	98.08	99.06	99.17
TP3 / S_3/n_0	9.77	85.46	94.65	97.20	97.79	98.30

Målingene i den virkelige prosessen inneholder alltid målestøy og måleusikkerhet. Resultatene i tabellen ovenfor var for målinger uten støy. En tilnærming til den virkelige prosessen og kalibrering med historiske data, ble gjort ved å introdusere støy i X matrisen (temperatur- og trykkmålingene). Den forklarte variansen ble noe redusert ved introduksjon av støy i målingene, men resultatene var fremdeles akseptable. Metoden for validering av resultatene kan gi et noe optimistisk overslag for de fremtidige prediksjonene, men selv om EPV var noe lavere enn beregnet ville resultatene fortsatt være plausible.

Tabell 7-3 EPV % (y) for gitt antall faktor for estimatorene (99 % C_{12}), skalering (S_i) og støy (n_i)

Estimator / S_i / n_i	EPV % (y) for gitt antall faktor					
	1	2	3	4	5	6
TP1 / S_2 / n_1	8.35	63.56	81.69	92.72	95.54	97.71
TP1 / S_2 / n_4	8.56	67.85	84.00	94.44	96.63	98.17
TP1 / S_3 / n_6	9.77	85.60	91.69	94.97	95.18	95.19
TP3 / S_2 / n_1	9.72	87.91	92.49	97.04	97.70	97.70
TP3 / S_2 / n_4	10.00	88.00	94.01	97.60	98.02	98.02
TP3 / S_2 / n_6	8.66	68.18	85.03	95.22	96.55	97.01

Skaleringen av variablene hadde innvirkning på resultatene, og skalering av X -matrisen med standardavviket gav best resultater. Transformasjon av sammensetningen gav i liten grad bedre resultater. Dette kan skyldes kolonnen hadde en forholdsvis lineær profil (Figur 5-1), og transformasjon vil dermed ikke kunne bedre resultatene i særlig grad. Transformasjon av sammensetningen vil kunne hindre at estimatet for massefraksjon C_{12} overstiger en, og dermed være utenfor det mulige området.

Antall reelle trinn i kolonnen med strukturert pakning vil kunne variere med varierende last og driftsbetingelser. Forsøk med variasjon av antall teoretiske trinn derfor utført. Mer robust estimator var formålet med å inkludere effekten av en eventuell endring i antall teoretiske trinn i kolonnen. Plassering av målepunktene i kolonnen ved 20 trinn ble gjort ved å justere plasseringen med et trinn. Dette vil sannsynligvis ikke gi et helt korrekt bilde av endringen, da plasseringen vil avvike for punktene mellom føde og topptrinnet. Resultatene fra beregningen var i seg selv tilfredsstillende, men vil inneholde stor usikkerhet pga plasseringen målepunktene.

Resultatene for 95 % C_{12} i produktstrømmen gav sammenliknbare med estimator for 99 % C_{12} , og problemstillingene som ble diskutert for estimatoren for 99 % C_{12} vil også gyldig for estimatoren for 95 % C_{12} .

7.3 Sannsynlig prestasjonsevne i prosessen

Estimatorene som ble presentert i avsnitt 5.3 har ikke blitt testet med en dynamisk modell av kolonnen. Dynamiske simuleringer av kolonnen ville kunne gitt en indikasjon på de forventede prestasjonsevnene til estimatorene. Forutsetningene for å lage en dynamisk modell ble vurdert å ikke tilstede, dette skyldes hovedsakelig grunner tilsvarende de som var gitt i avsnitt 7.1.3. Det var også forutsatt at estimatorene skulle bli testet direkte under driften av anlegget, men dette har p.t. dessverre ikke blitt gjort pga stort arbeidspress og dermed tidsnød hos Oleon Scandinavia AS.

I avsnitt 5.4 ble estimatoren testet ved å estimere sammensetningen på grunnlag av historiske dataene og sammenliknet med de analyserte sammensetningene. Sammenfallet mellom de analyserte og de estimerte verdiene var akseptabelt tatt i datagrunnlaget. Dataene fra prosessen var kun trend verdier og dermed grovt filtrert, og manglet dermed var noe informasjon om endringene i målingene av temperatur og trykk. Den kanskje viktigste grunnet til avviket mellom de analyserte og estimerte verdiene var at estimatoren ble forsøkt på et datasett som lå utenfor kalibreringsområdet for estimatoren. Innholdet av C_{12} lå i deler av perioden under 97 % (se Figur 5-15) og dermed under det nivået som var forutsatt for estimatoren. Produktet inneholdt også en betydelig fraksjon av lette komponenter (nederst i Figur 5-16), og forutsetningen for estimatoren var at fødestrømmen (Bilag F) til kolonnen ikke skulle inneholde betydelig mengder av lettere komponenter enn C_{12} . Differansen mellom \hat{y} og y viste en kovarians med innholdet av lette komponenter i produktet (Figur 5-16).

Estimatorene ble beregnet på grunnlag av statiske simuleringer av kolonnemodellen, og denne modellen har gitt resultater som samsvarer med den virkelige kolonnen. Mejdell [Ref II] har vist at en statiske estimator er tilstrekkelig for destillasjonskolonner. Det virker også rimelig at dette også skal være tilfelle for den undersøkte kolonnen. Den mangelfulle reguleringen av kolonnen, som ble omtalt i kapitel 6, kan påvirke negativt den forventede nøyaktigheten til estimatoren.

7.4 Videre arbeid

Bakgrunnen for utviklingen av estimatoren var i første omgang å kunne estimere sammensetningen i passive stilling til prosessen, hvilket innebærer at estimatoren ikke skulle bli anvendt til regulering. Estimatoren som ble presentert i kapitel 5 vil kunne estimere sammensetningen i produktstrømmen, men en endelig verifisering av resultatene mangler.

Implementering av estimatoren online i prosessen vil være første steget videre, hvis fokus kun blir satt på estimatoren. Mejdell et. al.[Ref III] undersøkte PLS estimatører beregnet på grunnlag av data fra simuleringer og eksperimentelle forsøk. Det ble funnet at estimatoren som kun var basert på simulerte data kunne være følsom for et misforhold mellom resultatene simuleringer og eksperimentelle forsøk. Hvis estimatoren skulle vise seg å gi gode estimater over lengre tid eller eventuelle feil enkelt kunne bli korrigert, kan en videre anvendelse av estimatoren i regulering av prosessen bli vurdert. En oppdatering av estimatoren med data fra prosessen i tillegg til resultater fra simuleringer, vil sannsynlig kunne lede til en mer robust estimator. En kombinerende av prosessdata og simuleringsdata vil kunne danne et kalibreringssett som dekke alle ”retninger” i rommet for de uavhengige variablene (y_D , x_B og forstyrrelser) og dermed tilstrekkelig eksitasjon av variablene i følge Mejdell.

En oppdatering av den statiske modellen med temperatur- og trykkdata fra stasjonærverdier for kolonnen vil som tidligere nevnt kunne gi en bedre estimator enn en som kun er basert på simuleringer. Observasjonene som ble presentert i kapittel 6 viste at refluksstrømmen oscillerte med forholdsvis stor amplitude. Virkelige stasjonære verdier for kolonnen vil dermed trolig ikke eksistere. En rekonfigurering av regulatorinnstillingen vil dermed som forklart i avsnitt 6.2 være det første en bør gjøre før en eventuell oppdatering av regresjonsmodellen blir gjort. Stabilisering ved en riktig regulering av refluksnivået vil da kunne gi anvendelige resultater. Problemer i forbindelse med regulering av destillasjonskolonne 2, er etter det som er kjent gjeldende for hele anlegget.

Det foreslåtte videre arbeidet med sammensetningsestimatorene vil være knyttet til problemene forbundet med situasjonen p.t., med tanke på regulering eller manglende regulering av kolonnen og det totale prosessanlegget, og er som følger

- implementere sammensetningsestimatorene i destillasjonskolonnen
- stabilisere reguleringsløyfer som svinger
- lukke reguleringsløyfer som er åpne
- oppdatere sammensetningsestimatorene
- vurdere anvendelse av oppdaterte sammensetningsestimatorene i lukkede reguleringsløyfer

8 Konklusjon

I denne hovedoppgaven ble det utviklet PLS estimatorene for produktsammensetningene i en destillasjonskolonne. I kolonne ble det destillert et flerkomponent system av fettsyrer fra kokosolje. Produktstrømmen over topp skulle tilfredsstille spesifikasjonene ved 95 % og 99 % (masse) laurinsyre ($C_{12}H_{24}O_2$). Estimatorene viste å kunne gi gode estimater for sammensetningen eller innholdet av laurinsyre innenfor de rammene som ble vurdert i oppgaven. Avhengigheten av de sterkt forsinkede GCL (gass-væskrokromatografi) analysene, som tok om lag 45 minutter å gjennomføre, kan dermed reduseres ved å anvende de utviklede estimatorene.

Estimatorene ble utviklet ved å kun benytte data fra statiske simuleringer av destillasjonen i Hysys™. Destillasjonen ble foretatt under vakuum betingelser i en kolonne med strukturerte pakninger. Trykket i kolonnen varierte både i toppen av kolonnen, som ble regulert av vakuumsystemet, og ved det totale trykkfallet over kolonnen. Modelleringen av kolonnen ble gjort ved en platekolonne, og resultatene fra simuleringene stemte overens med tilgjengelige trend- og analyseverdier fra prosessen.

I oppgaven ble det vist at estimatene ble betydelig forbedret ved å inkludere trykkmålinger i tillegg til temperaturmålinger i estimatoren. Den forbedrede estimeringsvekten kom hovedsakelig av den store trykkvariasjonen som forelå i prosessen. Temperatur ble målt seks steder i kolonnen, men kun fire av disse ble anbefalt brukt i estimatorene. Estimatorene som ble foreslåtte inneholdt målinger av trykket i topp og bunn, og de fire temperaturmålingene som var plassert over fødepunktet. Det optimale antallet faktorer i PLS estimatoren ble vurdert å være fire.

En endelig verifisering av resultatene ved å anvende estimatoren i anlegget ble ikke foretatt. Det kan dermed ikke trekkes en endelig slutning om estimatoren også vil være tilstrekkelig til å bli anvendt ved regulering av produktsammensetningen til destillatet.

En reguleringsløyfe for kolonne var p.t lukket. Reguleringen av nivå i refluksakkumulatoren viste instabilitet, hvor manipuleringen av refluksstrømmen til kolonnen fører til oscillasjoner i denne strømmen. Tuning av nivåregulatoren for å hindre instabilitet blir derfor anbefalt å gjennomføre. Lukking av åpne løyfer for destillasjonskolonnen blir deretter anbefalt, hvor en verifisert estimator eventuelt kan benyttes.

Trondheim, 7. juni 2002

Sigurd Myhre Hellem

9 Symbolliste

Tabell 9-1 Symbolliste for prosessvariable og parametere

Symbol	Forklaring
α_{AB}	relativ flyktighet for komponentene A og B
z_F	molfraksjon lett komponent i kolonnen
y_i	molfraksjon av komponent i i gassfasen
x_i	molfraksjon av komponent i i væskefasen
V	molar dampstrøm i kolonnen [kmol/h]
T_b	kokepunkt for en komponentblanding [$^{\circ}\text{C}$]
P_b	referansetrykk ved en temperatur [mbar]
P_A	metningstrykket til komponent A [mbar]
L	molar væskestrøm i kolonnen
F	mengde føding til kolonnen [kmol/h]

Tabell 9-2 Symbolliste for multivariat kalibrering

Symbol	Forklaring
\hat{y}	estimert verdi for sammensetningen
\bar{x}	gjennomsnitt av variabel
a	antall latente variable eller faktorer i PLS modellen
A_{\max}	maksimalt antall latente variable eller faktorer i PLS modellen
b	regresjonskoeffisient i PLS regresjon
c	skaleringsfaktor i PLS
E	støydel ved latent variabel regresjon

Symbol	Forklaring
f	feil i lokal PLS modell
L_T	transformert temperatur, basis kokepunkt
L_θ	transformert temperatur, basis
P	ladningsmatrisen
s	standardavviket (SD) til variabel
T	skårmatrisen
t_a	skårvektorene
$T_{L/H}^b$	kokepunkt lett (L) eller tung (T) komponent
W	vekt for variabel
w_a	enhetsvektorene
X	matrise for "uavhengige" variable, for eksempel målt temperatur
Y	matrise eller vektor for avhengige variable (sammensetning)
Y_D	transformert produktsammensetning
γ	forholdstall for singularverdier
ε	residual ved lineærregresjon
θ	temperaturmålinger i kolonnen [$^{\circ}\text{C}$]
$\theta_{L/H}$	referansetemperatur i kolonnen [$^{\circ}\text{C}$]
σ	singularverdi

Referanser

- I) Buckley, Page S, William L. Luyben, Joseph P. Shunta, Design of Distillation Column Control Systems, Edward Arnold (1985)
- II) Mejdell, Thor, Estimators for Product composition in distillation columns, Doktor Ingeniøravhandling, NTH-Trykk (1990)
- III) Mejdell, Thor, Sigurd Skogestad, Composition Estimator in a Pilot-Plant Distillation Column Using Multiple Temperatures, Industrial & Engineering Chemistry Research, **30**, 2555-2564 (1991)
- IV) Mejdell, Thor, Bengt-Ove Andersson, Using temperature profile for product quality estimation on a distillation column, ISA Transactions **33**, 27-34 (1994)
- V) Ohshima, Masahiro, Hiromu Ohno, Lori Hashimoto, Mirkiro Sasajima, Masayuki Maejima, Keiichi Tsuto, Tadaharu Ogawa, modell predikative control with adaptive disturbance prediction and its application to fatty
- VI) Shinskey, F.G., Distillation Control – For Productivity and Energy Conservation, 2nd.ed. McGraw-Hill (1984)
- VII) Mejdell, Thor, Bengt-Ove Andersson, Using temperature profile for product quality estimation on a distillation column, ISA Transactions **33**, 27-34 (1994)
- VIII) Kano, Manabu, Koichi Miyazaki, Shinji Hasebe, Iori Hashimoto, Inferential Control system of Distillation Compositions Using Dynamic Partial Least Squares Regression, DYCOPS-5, 5th IFAC Symposium on Dynamics and Control of Process Systems, Corfu, Greece, June 8-10 (1998)
- IX) Liu, June, Kwanggi Min, Chonghun Han, Kuu Soo Chang, Robust Nonlinear PLS Base on Neural Networks and Application to Composition Estimator for High-Purity Distillation Column, Korean Journal of Chemical Engineering, **17**(2), 184-192 (2000)
- X) Marten, Harald, Tormod Næs, Multivariate Calibration, John Wiley & Sons, (1989)
- XI) Joshep, B, B. Brosilow, Inferential control of processes, AIChEJ. **24**, 485-509 (1978)
- XII) Esbensen, Kim, Suzanne Schönkopf, Tonje Midtgaard, Multivariate Analysis in Practice, CAMO, (1996)
- XIII) Gunstone, Frank D., John L. Harwood. Fred B. Padley, The Lipid Handbook, Chapman and Hall (1986)
- XIV) Gmehling, J., U. Onken, W. Arlt, Vapor-Liquid Equilibrium Data Collection, DEMCHEMA, (1977)

- XV) Aly, G., I. Ashour, Applicability of the Perturbed Hard Chain Equation of State for Simulation of Distillation Processes in the Oleochemical Industry. Part I: Separation of Fatty Acids, Separation Science and Technology, **27**(7), 955-974 (1992)
- XVI) Gulbrandsen, Morten, Teknisk Sjef ved Oleon Scandianvia AS i Sandefjord, Personlig kontakt (2002)
- XVII) Kister, Henry Z., Distillation Design, McGraw-Hill, (1992)
- XVIII) Gill, Eugene V., Rustum J. Adranvala, The Use of Computer Simulation in the Design of Fatty Acid Distillation Columns, with a Comparison to Actual Operating Results A Applications II, teknisk program AIChE (1999)
- XIX) Anvendelse av Kjemometri innen Forskning og Industri, Sjefsredaktør: Rangnar Nortvedt, Norsk Kjemisk Selskaps Faggruppe for Kjemometri og Svenska Kemistsamfundet, Tidsskriftforlaget Kjemi AS (1996)
- XX) Wise, Barry M., Neal B. Gallagher, PLS_Toolbox 2.1 for use with Matlab™, Eigenvector Research, Inc. (2000)
- XXI) Kresta, J.V., T. E. Marlin, J. F. MacGregor, Development of Inferential Process Models Using PLS, Computers & Chemical Engineering, **18**(7), 597-611 (1994)
- XXII Seborg, Dale E., Thomas F. Edgar, Duncan A. Mellichamp, Process Dynamics and Control, John Wiley & Sons (1989)