

Tillståndsovervakning med hjälp av Multivariat dataanalys

Per-Oskar Westin, Ida Bodén



Detta material är framtaget med stöd från *EU:s regionala strukturfondsprogram för stärkt konkurrenskraft och sysselsättning* och i samarbete med Länsstyrelsen Västernorrlands län, Örnsköldsviks kommun & Bosch Rexroth Mellansel AB.

Oförberedda driftstopp hos maskiner är inte bara besvärligt och dyrt att åtgärda utan medför även merkostnader i form av en stillastående produktion. Genom att få en effektiv övervakning av maskinen och i realtid få en uppfattning om något är på väg att hända som kan orsaka haveri och därmed driftstopp får vi inte bara en tryggare och stabilare produktion utan även en möjlighet att åtgärda fel innan stora kostnader uppstår.

Vid ett haveri är det sällan bara en sak som orsakat haveriet och det är framför allt sällan man på förhand vet hur haveriet kommer att manifestera sig. I praktiken kan det betyda att man behöver flera parallella system som övervakar och larmar för eventuella haverier vilket i sin tur kan göra tolkningen komplicerad. Ett bättre sätt att göra detta vore att samla data från flera mätningar och analyser till en gemensam bedömning av systemstatus. ITH har som ett av de första institut i världen, tillsammans med Bosch Rexroth och iViDA AB, genomfört en studie där hydraulolja har analyserats online så väl som offline och kombinerat den informationen med loggade körvariabler. Den sammantagna informationen har analyserats med *Multivariat Dataanalys* (MVA). Genom att använda MVA kan flera tester analyseras tillsammans och resultaten från flera analysmetoder kan samköras på ett kraftfullt sätt för att se ifall det går att förutse om ett haveri är närliggande eller inte. I denna artikel presenterar vi våra lärdomar och resultat från denna jämförelse. Resultaten visar att man efter ett fåtal tester kan förutse ett haveri med 93 % precision genom detta förfarande. Det finns även ett mervärde att kombinera data från online partikelräkning med ICP-resultat (grundämnesanalys) då vi ser en tydligare gräns mellan de mätningar som utgör steady-state-körning och de som indikerar på haveri hos maskinen. Vi visar också att MVA kan användas tillsammans med FT-IR mätningar för att bedöma systemhälsa och identifiera problemkällor.

Insamling av testdata

Den analyserade datamängden är insamlad från testuppställningen som används inom ITHs EU-projekt för *Tillståndsovervakning av hydraulsystem*. Projektet och tidigare resultat har rapporterats i referenserna [1-3]. Testuppställningen är fokuserad på onlinemätningar med partikelräknare på en rigg där accelererad nöting av rullager provoceras fram och där ett haveri motsvaras av allvarliga lagerskador. I fyra olika testomgångar har riggens belastning och varvtal varierats (*körvariabler*) och en mängd mätningar utförts. Utöver partikelnivåer i systemet mättes ett antal andra variabler såsom temperaturer och rotationsmotstånd (*systemtillstånd*). Utöver dessa kontinuerliga mätningar har flaskprover från systemet analyserats med hjälp av grundämnesanalys (ICP) och FT-IR för att kunna göra en helhetsbedömning av systemhälsan.

Den insamlade datamängden har sedan manuellt klassats in i två tillstånd: "steady-state" och "haveri", där haveri-klassen är datapunkter som bedömts uppvisa tecken på annalkande haveri. Dessa två klasser har sedan använts för att utvärdera möjligheten att genom MVA klassificera nya mätpunkter till steady-state eller haveri.

Multivariat dataanalys

Multivariat dataanalys (MVA) är ett samlingsnamn för flera dataanalysmetoder där man kan analysera många olika variabler samtidigt. De metoder som används här är i första hand *Principalkomponentanalys* (PCA) samt *Partial least square-diskriminant analys* (PLS-DA). En beskrivning av PCA ges i Bild 1 och i ekvation (1) nedan.

$$\begin{matrix} & m & & & m \\ \boxed{X} & = & \boxed{T} & \boxed{P^T} & + & \boxed{E} \\ n & & n & lv & & n \end{matrix}$$

Bild 1. Matriser i PCA

$$X = TP^T + E \quad (1)$$

där

T innehåller de nya koordinaterna, scores

P innehåller koefficienterna som transformerar de gamla koordinaterna till de nya, loadings

E innehåller den information som inte förklaras av TP^T

lv är antal principalkomponenter

m är antal X-variabler

n är antal observationer

För en mer ingående beskrivning av PCA och PLS-DA se ref [4]. Kortfattat går de ut på att man har ett stort och komplext dataset (**X**) där varje objekt (**n**) utgörs av ett oljeprov och varje variabel (**m**) utgörs av de mätresultat som man får vid analysen av själva oljeprovet (se Bild 1. Matriser i **PCA**). I en mer traditionell statistisk analys analyserar man två, eller högst fem variabler åt gången och jämför dessa med varandra. En av styrkorna med MVA är att hur många variabler som helst kan analyseras på en och samma gång. Vid själva analysen fungerar varje variabel som en koordinat i en multidimensionell rymd. I den rymden beräknas nya koordinater, scores (**T**), fram utifrån en ny axel, eller principalkomponent, med hjälp av koefficienter/loadings (**P**). Dessa scores ger objekten/oljeprovorna specifika platser i det nya koordinatsystemet utifrån deras egenskaper som mäts upp online och offline initialt. Detta betyder att i en graf över det nya koordinatsystemet, en s.k. *Score plot*, kommer oljeproven att gruppera sig utifrån sina specifika egenskaper och man kan på så vis grafiskt rita upp hur olika oljeprov förhåller sig till varandra. Utifrån denna kunskap kan man med hjälp av en PLS-DA sedan fastställa grupptillhörighet hos en ny, okänd mätning och kan på så vis förutse ett kommande haveri. Enkelt uttryckt ritas proverna in i en tvådimensionell graf där mätpunkter från steady-state och haveri kommer att skilja sig i olika grupper.

Genom att studera loadings i en s.k. *Loading plot* eller weights i en s.k. *Weights plot* kan man identifiera vilka variabler som *orsakar* grupperingarna av prover i en *Score plot*. Man kan alltså ur en stor mängd mätresultat identifiera vilka mätmetoder som är relevanta för att identifiera en viss förändring som skett med systemet.

Resultat

Nedan kommer vi att visa hur samkörningen av systemtillstånd med partikelnivåer i en online-mätning resulterar i en mer preciserad bedömning av systemhälsa. Därefter kommer vi att ge ett exempel på hur möjligheten att samköra online-analyser med offline-analyser leder till en tydligare gruppering av resultaten och därmed i framtiden kommer att för-

bättra möjligheten att bedöma systemhälsa och förutse kommande haverier. Slutligen vill vi illustrera hur MVA på resultat från FT-IR kan användas för att lyfta fram även små skillnader i systemhälsa och ligga till grund för att identifiera systemförändringar.

Online mätning. MVA utfördes på data i två omgångar. Dels MVA enbart på data från partikelräkning och dels på all data insamlad om systemtillstånd, dvs. inklusive temperaturer och rotationsmotstånd. I båda fallen användes körvariabler för att särskilja olika resultatgrupperingar. Vi kunde konstatera att när systemet närmar sig haveri sker en tydlig förskjutning av mätpunkterna bort från steady-state-grupperingen (se Bild 2) mot haveri-grupperingen. Dessutom grupperar sig punkterna beroende på körvariabler. Detta visar att bästa möjliga tillståndsovervakning får man om man tar hänsyn till körvariabler lika väl som mätresultat. Det var också tydligt att tolkningen av resultaten underlättades av att inkludera mer data om systemtillstånd.

För att se i vilken omfattning det var möjligt att förutse ett maskinhaveri gjordes två olika PLS-DA på data från alla fyra testerna. Vid den första analysen togs endast partikelnivåer med och en precision på 91 % erhöles vid klassificering av okända data. D.v.s. att man kunde förutse ett haveri med en korrekthet på 91 %. Vid den andra analysen tog man, förutom partikelnivåer, hänsyn till övriga data om systemtillstånd vilket ökade precisionen till 93 %. Detta understryker värdet av att inkludera fler mätvärden i analysen.

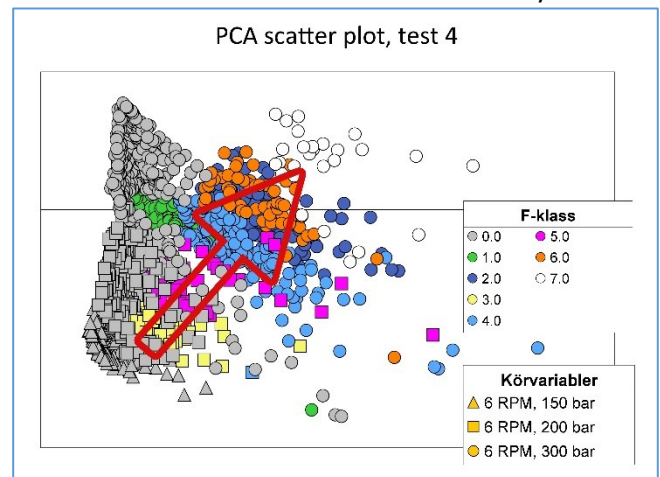


Bild 2. PCA Score plot, Gruppering av resultat från partikelnivå- och systemtillståndsvariabler, från en testomgång, visar en generell trend uppåt, åt höger ju närmare haveri man hamnar. Notera dock att vid ändrade körvariabler (F-klass 3 och 5) skiftar grupperingen av F-klass. F-klass är en subjektiv klassificering av hur nära haveri man befinner sig. "0.0" motsvarar steady-state och 7.0 är strax innan haveri.

Samkörning av data. Från den första testomgången kombinerades resultat från online-mätningar och ICP-resultat från flaskproverna. Genom att tillföra ICP-resultaten till online-resultaten får vi ett minskat

överlapp av punkter som indikerar på haveri (gröna punkter Bild 4) och steady-state (blå punkter i Bild 4) jämfört med om endast Online-resultaten används (jämför Bild "4" a och b). ICP-resultaten bidrar alltså med att stärka tolkningen av individuella mätpunkter som steady-state eller haveri.

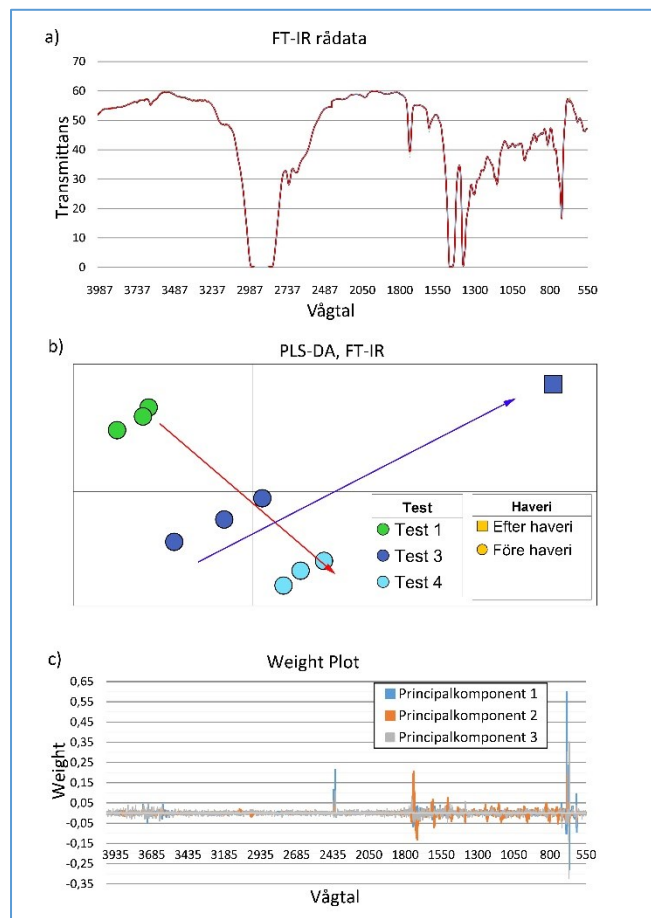


Bild 3. (a) FT-IR rådata, överlappande spektra från alla mätningar. (b) PLS-DA score plot med Haveri som diskrep- tant. Konsumtion av ZDDP (Anti-wear tillsats) ökar med blå pil medan oxidering ökar med röd pil. (c) Weights för de tre första principal-komponenterna. Våglängd vars stapel är högre än 0.05 eller lägre än -0.05 bedöms som relevant för modellen.

FT-IR. FT-IR mättes på flaskprover från de tre första testomgångarna. MVA ger möjligheten att jämföra stora mängder spektral data även om FT-IR spektra i en första anblick ser identiskt ut oavsett mättillfälle (Bild 3a). FT-IR på provet som tagits efter haveri skiljer sig tydligt från de prov som är tagna vid steady-state (Bild 3b). Vi kan se genom PLS-DA analysen att vissa våglängder i FT-IR-spektra framträder som karaktäristiska vid jämförelsen av prover (Bild 3c). Det är dessa våglängder som står för de främsta skillnaderna mellan grupperna som identifieras i Bild 3b. Ett visst våglängdsområde kan eventuellt korrespondera med förändrat innehåll av anti-wear tillsatser (ca 650 cm^{-1}). Över tid ser vi dessutom en generell förändring vid våglängder som associeras med oxidation (ca 1720 cm^{-1}).

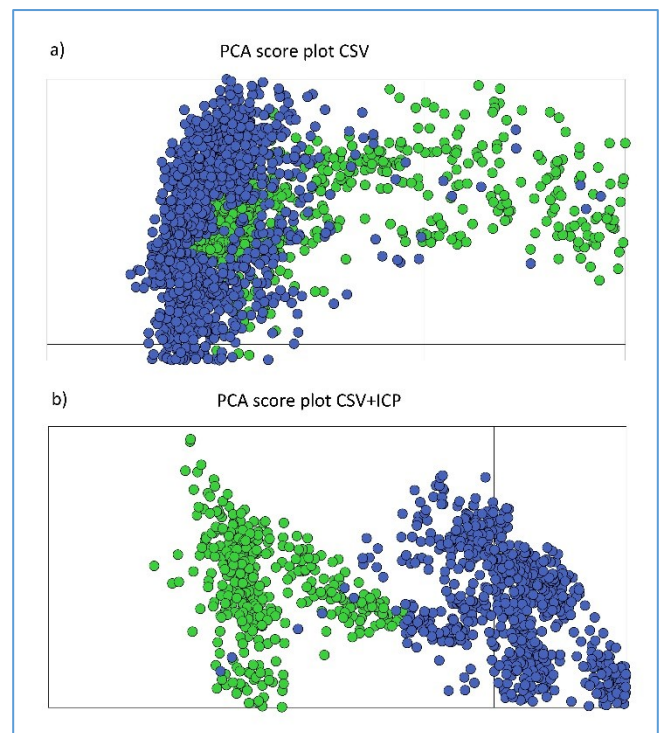


Bild 4. (a) PCA score plot med ett större överlapp mellan oljeprover som är steady-state (blå punkter) och haveri (gröna punkter) jämfört med (b) där data från både partikelräkning (CSV) och ICP är sammanslaget.

Slutsats och diskussion

I detta arbete har data från fyra olika testomgångar utvärderats. Vi har kunnat konstatera att multivariat analys (MVA) ger möjlighet till tolkning av insamlad data på ett sätt som kan användas för tillståndsovervakning. Detta genom att resultaten från MVA skapar grupperingar där mätningar från annalkande haveri skiljer ut sig från mätningar av steady-state. Vi har visat att man, genom att komplettera data från partikelräkning med andra oberoende systemvariabler, kan förstärka denna tolkning och få en tydligare bild av status i systemet. Ju mer data som samlas in och används för att bygga en databas/modell, desto tydligare kan man förvänta sig att analys och tolkning blir. I detta arbete når vi en förbättring från 91 % till 93 % precision för att förutse ett haveri. Förbättringen sker alltså med hjälp av mätningar som man i en normal analys skulle ha bortsett ifrån. Nivån i sig bör anses som hög med tanke på att det är en relativt liten datamängd som ligger till grund för tolkningen trots att de fyra olika testernas data samkördes. Dessutom belastades de olika testerna på olika sätt i försöksupställningen vilket ytterligare försvårar tolkningen. En prediktionsmodell som är byggd på större mängd data från flera maskiner i fält kommer att ge en ännu högre precision.

Genom att analysera resultaten har vi också kunnat se hur kördata, dvs. de individuella mätningarnas belastning och varvtal, påverkar grupperingen. För att minimera riskerna för feltolkning och få en så kraftfull tolkningsmöjlighet som möjligt bör alltså kördata log-

gas och användas för att klassificera tillståndsmätningarna. Enkelt uttryckt: *Det som bedöms som normalt tillstånd vid hög belastning kan vara en indikation på nära förestående haveri vid lägre belastning.*

När vi samkör data från online-mätningar med ICP-resultat från flaskprover så ser vi att vi får en ökad separation mellan de mätningar som är klassade som haveri och de som klassats som steady-state (Bild 4a och b). Detta visar i praktiken på att det blir lättare att förutse om en okänd mätning tillhör haveri eller steady-state vilket också skulle öka precisionen i att förutse ett haveri och framför allt att få reda orsaken till det kommande haveriet.

Den multivariata dataanalysapproachen som används här ger också möjlighet till s.k. Machine-learning system där en växande datamängd tränar systemet till att ge bättre och bättre prediktioner. Med andra ord, systemet lär sig självt till att bli bättre på att förutse haverier. Det är rimligt att anta att en större datamängd därför skulle ge en precision på 100 % vilket innebär att det går att bygga larmsystem som aldrig larmar i onödan och som alltid larmar innan haveriet är ett faktum.

Eftersom FT-IR är kapabelt till att identifiera förändringar i oxidation, nitrering, sotbildning, tillsatsförbrukning och föroreningar (vatten, glykol, oljor etc.) kommer gruppering genom MVA av FT-IR data att utgöra en ytterst användbar grund för att bedöma systemhälsa och identifiera problemkällor. PLS-DA kan användas för att identifiera karaktäristiska våglängder i en stor mängd oljeprover. I bästa fall kan man knyta dessa våglängder till specifika komponenter i hydraulvätskan (basolja, tillsatser eller föroreningar) och därigenom kartlägga olika haveritillstånd baserat på hur prover grupperar sig genom multivariat analys.

- [4] Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006) *Multi- and Megavariate Data Analysis : Part I: Basic Principles and Applications*. Umetrics Inc; 2006.

Referenser

- [1] Sjödin, E., & Westin, P-O. (2013). *Comparison of optical and magnetic particle detection systems for detection of pitting damage in low speed hydraulic motors*. Proceedings of The 13th Scandinavian International Conference on Fluid Power, SICFP2013, Juni 3-5, 2013, Linköping.
- [2] Westin, P-O., Marklund, P., & Sandström, J. (2013). *Wear detection by monitoring hydraulic oil contamination – an experimental comparison between on-line and off-line measurements*. Proceedings of World Tribology Congress, WTC'13, Torino, Italien.
- [3] Westin, P-O. En jämförelse av onlinemätning och flaskprover vid mätning av partikelinnehåll. Fluid Scandinavia 3/2014.