

Ekstern kvalitetskontrol

**Naturstyrelsens
Referencelaboratorium
for Kemiske Miljømålinger**

Titel: Ekstern QA

Resume: Danske laboratorier skal deltage i præstationsprøvninger eller andre former for ekstern kvalitetskontrol for at efterleve kravene til analysekvalitet i Bekendtgørelse om kvalitetskrav til miljømålinger. Deltagelse i laboratoriesammenligninger eller præstationsprøvninger er ligeledes et led i den kvalitetssikring, som laboratorier skal udføre i henhold til ISO 17025. Nærværende rapport beskriver, hvilke kriterier der bør lægges til grund ved laboratoriernes valg af udbyder af præstationsprøvninger. Desuden beskrives retningslinjer for omfang af sammenlignende målinger mellem to eller flere laboratorier som alternativ, hvor der ikke er adgang til relevant præstationsprøvning.

Emneord: Ekstern kvalitetskontrol, præstationsprøvning, sammenlignende målinger

Forfatter: Stine K. Ottsen

Godkendt af: Ulla Lund

Rådgiver: Eurofins Miljø A/S

Udgiver: Naturstyrelsens Referencelaboratorium for Kemiske Miljømålinger

Udgiverkategori: Statslig

År: 2012

Sprog: Dansk

Copyright: Må citeres med kildeangivelse.
Naturstyrelsens Referencelaboratorium for Kemiske Miljømålinger

	Rapport (endelig)	SJN	UOL	UOL	18022013
	Rapport (Udkast)	SJN	UOL	UOL	20112012
Revision	Beskrivelse	Udført	Kontrolleret	Godkendt	Dato

Indhold

1	Indledning	4
2	Deltagelse i præstationsprøvninger	5
2.1	Strategi for valg af præstationsprøvninger	5
2.2	Krav til udbyder	6
3	Sammenlignende målinger	8
3.1	Udførelse	8
3.2	Prøver	8
3.3	Bestemmelse af nominel værdi	9
3.4	Stabilitet og homogenitet	10
3.5	Rapportering	10
4	Referencer	11
	Bilag A - Litteraturliste	12

1 Indledning

Danske laboratorier skal deltage i præstationsprøvninger eller andre former for ekstern kvalitetskontrol for at efterleve kravene til analysekvalitet i Bekendtgørelse om kvalitetskrav til miljømålinger (herefter benævnt "bekendtgørelsen") /1/. Deltagelse i laboratoriesammenligninger eller præstationsprøvninger er ligeledes et led i den kvalitetssikring, som laboratorier skal udføre i henhold til ISO 17025 /2/.

Referencelaboratoriet arbejdede i 2007 med opfølgning i 2008 med sammenstilling af muligheder for fortsat ekstern kvalitetssikring efter udbuddet af danske præstationsprøvninger er kraftigt reduceret /8/. Formålet med nærværende rapport er at beskrive, hvilke kriterier der bør lægges til grund ved laboratoriernes valg af udbyder af præstationsprøvninger, samt retningslinjer for omfang af sammenlignende målinger mellem to eller flere laboratorier.

Ordet præstationsprøvning anvendes her om kommercielt udbudte prøvninger som har til formål at give en vurdering af deltagernes resultater. Andre prøvninger, f.eks. metodeundersøgelser eller certificering af referencematerialer kan ligeledes anvendes, men kræver en indsats fra deltagerens side for vurdering af resultaterne.

Sammenlignende prøvninger anvendes her som begreb, der dækker en mindre sammenligning mellem få laboratorier, der typisk udføres på kollegial basis eller ved at det laboratorium, der udfører sammenligning, køber analyser hos et andet (akkrediteret) laboratorium.

2 Deltagelse i præstationsprøvnings

Det primære formål med ekstern kvalitetssikring er at give laboratorierne et værktøj til kontrol af målingernes korrekthed. Præstationsprøvnings giver en objektiv vurdering af de deltagende laboratorier. Denne vurdering kan benyttes til at forbedre laboratoriets præstationer og/eller give tillid til, at laboratoriet evner at udføre en specifik måling på fornuftig vis.

Valget af præstationsprøvning har to ben: Laboratoriet skal først have overblik over sine analyseområder, og planlægge hvordan de kan dækkes ind ved deltagelse i præstationsprøvnings. Dernæst skal udbyderen af præstationsprøvningserne udvælges.

2.1 Strategi for valg af præstationsprøvnings

Ideelt skulle et laboratorium deltage i en præstationsprøvning for hver eneste målemetode, hver eneste parameter og hver eneste matrix. Dette er dog ikke muligt, hverken logistisk eller økonomisk. Derfor kan laboratorier identificere **undergrupper** bestående af måleteknikker, egenskaber og matricer, der er sammenlignelige. Undergruppen kan indeholde mere end en måleteknik, egenskab og prøvetype, så længe laboratoriet kan demonstrere deres ækvivalens og sammenlignelighed. En gruppering bør generelt ikke indeholde forskellige tekniske kompetencer, dvs. forskellige kvalifikationer, træning, udstyr, viden eller erfaring. Når en gruppering skal fastlægges kan det være brugbart at tage udgangspunkt i denne rækkefølge som anført i EA-4/18 /3/:

1. Målemetoden: Det er muligt, men ikke almindeligt, at inkludere forskellige målemetoder i den samme gruppering.
2. Parameteren: Det er muligt at inkludere mere end en parameter i den samme gruppering.
3. Prøvetype: Det er muligt at inkludere forskellige prøvetyper i den samme gruppering.

Når laboratoriet har identificeret de grupperinger som er dækkende for deres analyseområde, skal **frekvensen** af præstationsprøvnings fastlægges.

Ifølge bekendtgørelsen skal "Laboratorier (...) godtgøre deres kompetence ved så vidt muligt én gang årligt at deltage i præstationsprøvnings for alle anvendte målemetoder, men ikke nødvendigvis for alle matricer.". Deltagelse i præstationsprøvnings for hver målemetode skal altså foregå mindst én gang årligt med mindre, at der er vægtige modargumenter. Mangel på udbydere af præstationsprøvnings er ikke alene et argument for manglende deltagelse. I tilfælde af at der ikke findes præstationsprøvningsprogrammer, skal laboratorierne forsøge at arrangere sammenlignende prøvnings.

Fastlæggelsen af gruppering og frekvens af præstationsprøvnings vurderes ud fra en omhyggelig analyse af laboratoriets øvrige kvalitetskontrol. Især metoder til at beskrive, kvantificere og følge udviklingen i bias er væsentlige. Hvis den interne kvalitetskontrol kun omfatter syntetiske prøver, er præstationsprøvnings-

ningsdeltagelse påkrævet for at undersøge matricens indvirkning på analysen, og om estimerne for bias og standardafvigelse er for optimistiske. Metodeafprøvning og certificering af referencematerialer tæller som præstationsprøvnin-
ger.

Det kan være relevant at deltage i flere præstationsprøvnin-
ger for den samme gruppering hvert år, hvis laboratoriet vurderer, at det er nødvendigt for, at forskellige aspekter som målemetode, måleusikkerhed, detektionsgrænse og ma-
trixeffekt er dækket ind. Hvis en gruppering indeholder flere prøvetyper, er det ikke nødvendigt, at alle prøvetyper testes ved præstationsprøvnin-
ger hvert år, men om muligt er det ønskeligt, at der foregår en rotation af matricerne.

Når laboratoriet har klarlagt sine behov for deltagelse i præstationsprøvnin-
ger, skal dette behov omsættes til valg af udbydere.

2.2 Krav til udbydere

Laboratoriets valg af udbydere af præstationsprøvnin-
ger er kritisk for udbyttet af deltagelsen i en præstationsprøvning. Ifølge bekendtgørelsen skal præstations-
prøvnin-
ger "være repræsentative for og egnede til at dokumentere den analy-
sekvalitet, der er anført i bilag 1.3 - 1.17."

Ved vurdering af om en præstationsprøvning er egnet til formålet bør laboratori-
et som minimum vurdere følgende seks punkter (Eurachem, 5.3 /4/) [se også
2.2 i /8/]:

1. Svarer antallet af prøver, matricer og/eller niveauer i præstationsprøv-
ningen til det som opleves i laboratoriet?
 - a. Ifølge bekendtgørelsen skal præstationsprøvnin-
ger, "så vidt det er relevant for de pågældende målinger, foretages på prøvemate-
riale, der så vidt muligt repræsenterer laboratoriefremstillede,
naturlige prøver med kendt indhold og naturlige miljøprøver."
2. Er det statistiske design (planlægning, indsamling, analyse og rapporte-
ring af data) beskrevet og tager det hensyn til forskellige målemetoder?
 - a. Metodeafhængige parametre får i mange tilfælde ikke tillagt
nominel værdi i forhold til den metode, der kræves i bekendtgø-
relsen. Ifølge bekendtgørelsen skal laboratoriet for metoderela-
terede parametre "så vidt muligt vælge præstationsprøvnin-
ger, hvor den nominelle værdi er fastlagt ved anvendelse af de i bi-
lag 1.3 - 1.17 specificerede metoder."
3. Er antallet og oprindelsen af deltagere i præstationsprøvnin-
gen passende?
4. Er frekvensen af runder tilstrækkelig?
5. Overholder udbyderen ISO/IEC 17043? Akkreditering af præstations-
prøvnin-
gsudbydere af et nationalt akkrediteringsorgan giver yderligere til-
lid til udbyderes kompetence.

- a. Ifølge bekendtgørelsen skal præstationsprøvningsraterne "være tilrettelagt af akkrediterede organisationer eller nationalt eller internationalt anerkendt organisationer, som opfylder kravene i ISO/IEC 17043 eller andre tilsvarende internationalt accepterede standarder."
6. Præstationsprøvningsrater har også en vigtig uddannelsesmæssig rolle. Hvis præstationsprøvningsudbyderen behandler data på basis af forskellige analysemetoder og også behandler usikkerhedsresultaterne rapporteret af deltagerne, kan det give vigtige informationer til yderligere evaluering af analysemetoder.

Vurderingen af præstationsprøvningsudbydere kan gøres systematisk ved at benytte en ensartet fremgangsmåde, som fx skemaet i bilag i A i Eurachem /4/. Dette skema giver laboratorierne lejlighed til at dokumentere udvælgelsesprocessen og udarbejde en plan for deltagelse i præstationsprøvningsplaner, som det kræves i bekendtgørelsen: "Laboratoriet skal udarbejde planer for præstationsprøvning og opretholde registrering af deltagelsen. Planerne skal udformes på en sådan måde, at det er muligt at vurdere, om omfanget af præstationsprøvningsplaner dækker analyseområdet på relevant vis." Planen skal omfatte grupperinger, frekvens og udbydere, og hvordan opfølgning på præstationsprøvningsresultater skal foregå, så laboratoriet har en fremgangsmåde ved utilfredsstillende resultater. Hvis et resultat fra tidligere præstationsprøvning på området ikke har været tilfredsstillende, bør frekvensen af præstationsprøvningsplaner øges. Før frekvensen øges skal der selvfølgelig foretages en korrektion.

Planen vurderes en gang årligt /2/, f.eks. i forbindelse med den årlige granskning.

Inden for visse områder kan deltagelse i præstationsprøvninger være vanskelig grundet mangel på præstationsprøvninger med den relevante kombination af parametre, prøvetype og koncentrationsområde. For at sikre den nødvendige information fra eksterne prøvninger er det en mulighed at arrangere sammenlignende målinger i fællesskab med andre laboratorier.

3 Sammenlignende målinger

Sammenlignende prøvninger kan ses som præstationsprøvninger i meget lille skala, hvor de eneste deltagere er miljølaboratorier i Danmark. Ved sammenlignende prøvninger med kun danske laboratorier opnås, at prøver, herunder parametre, metoder og koncentrationsområder, er relevante for danske forhold /8/.

Denne tekst vedrører kun kvantitative målinger (numeriske resultater).

En sammenlignende måling består i sin enkleste udførelse af at et laboratorium deler en prøve i to dele, analyserer den ene del selv og sender den anden til analyse hos et akkrediteret laboratorium. Forskellen mellem resultaterne fra de to laboratorier vurderes i forhold til et defineret kriterium, som f.eks. kan fastlægges som beskrevet nedenfor.

Sammenligning mellem kun to laboratorier indebærer nogen usikkerhed. Nedenfor er beskrevet en fremgangsmåde, som giver lidt bedre basis for vurderingen.

3.1 Udførelse

Sammenlignende målinger kan udføres ved, at to eller flere laboratorier enes om prøvetype, parametre, metoder og koncentrationsområder. Rent praktisk kan det foregå ved, at en samling af laboratorier skiftevis arrangerer prøvefremstilling, transport til øvrige laboratorier, indsamling af resultater og databehandling. På denne måde fordeles omkostningerne ved organiseringen af prøvningen. Samtidig får et laboratorium ikke indsigt i de øvrige deltageres præstationer, og det vil ikke være muligt for arrangøren at drage fordel af kendskab til de øvrige måleresultater og herved forbedre dets egne resultater. En fast arrangør vil desuden ikke selv kunne deltage i en sammenlignende prøvning, da det ikke er uvildigt /8/.

3.2 Prøver

Certificerede referencematerialer, hvis sammensætning er ukendt for deltagerne, er at foretrække som prøver. Det er dog langt fra sikkert at der findes certificerede referencematerialer med den kombination af prøvetype, parameter og koncentrationsområde, der er relevant for den planlagte sammenlignende måling. Hvis sådanne referencematerialer findes, er udvalget lille, hvilket ikke er gavnligt for variationen fra sammenlignende måling til sammenlignende måling. Samtidig er der en vis sandsynlighed for, at et eller flere af de deltagende laboratorier på forhånd har kendskab til materialet og dermed dets certificerede værdi.

En anden og billigere løsning til fremskaffelse af prøver er split-samples. Ved split-sample opdeles en naturlig prøve i flere dele, og hver deltager analyserer en delprøve. Brugen af naturlige prøver fremfor certificerede referencematerialer giver større variation på sammensætningen af prøverne og matricen vil have indflydelse på analyserne. Ulempen ved brug af naturlige prøver er, at de er ustabile og det er derfor nødvendigt, at der er aftalt en fast analysedato, at prø-

vematerialet udsendes på samme tidspunkt til deltagerne, så de kan analysere samtidigt og herved mindske indflydelse fra ustabilitet.

3.3 Bestemmelse af nominel værdi

For at kunne bedømme laboratoriernes præstationer må der ved de sammenlignende målinger fastlægges en nominel værdi for de medtagne parametre. Metoden til bestemmelse af den nominelle værdi skal være fastlagt af de delta-gende laboratorier i fællesskab før den sammenlignende prøvning igangsættes.

Ved syntetiske prøver kan den nominelle værdi findes ud fra de afvejede mængder ved prøvefremstillingen. Begrænsningerne ved syntetisk prøve er, at basismaterialet ikke må indeholde den tilsatte parameter, eller at andelen af parameteren i basismaterialet skal være kendt /5/. Ved naturlige prøver findes den nominelle værdi oftest som en konsensusværdi mellem deltagerne, enten som middelværdien eller medianværdien af deltagernes resultater. Ved konsensusværdi ved et lavt antal deltagende laboratorier øges risikoen for, at den nominelle værdi ikke er korrekt på grund af fejlbehæftede målinger /8/.

Standardusikkerheden på den nominelle værdi, u_x , er afhængig af, hvilken metode der er benyttet til at fastlægge den nominelle værdi samt af antallet af deltagende laboratorier. Ved syntetiske prøver estimeres standardusikkerheden på den nominelle værdi som kombinationen af usikkerheden på de gravimetriske og volumetriske målinger samt usikkerheden ved den praktiske udførelse af afmålinger og afvejninger. Usikkerhed på nominel værdi som konsensusværdi er usikkerheden på medianværdien (eller middelværdien), som fastlægges ud fra den robuste standardafvigelse og antallet af data, der indgår i medianværdien

Deltagernes resultater evalueres på baggrund af Z-score, som beregnes ved

$$z = \frac{x - X}{\hat{\sigma}}$$

Hvor x er deltagerens resultat, X er den nominelle værdi, og $\hat{\sigma}$ er standardafvigelsen til vurdering af præstationen. For parametre, der er dækket af bekendtgørelsen, er $\hat{\sigma}$ defineret som $1,3 \cdot s_{T \max}$ eller $1,3 \cdot CV_{T \max}$ (bekendtgørelsens bilag 1.1.7), hvor værdierne for $s_{T \max}$ og $CV_{T \max}$ er anført i bilag 1.3 til 1.17. For parametre uden for bekendtgørelsen kan der drages en analogi med sammenlignelige parametre, der findes i bekendtgørelsen og herved opstille de værdier, som skal benyttes for $s_{T \max}$ og $CV_{T \max}$.

$\hat{\sigma}$ benyttes også til at vurdere størrelsen af usikkerhed på den nominelle værdi ved prøvningen. Hvis u_x er for stor i forhold til $\hat{\sigma}$ er der risiko for, at nogle laboratorier vil få advarsler grundet usikkerhed i bestemmelsen af den nominelle værdi, men der vil ikke være en årsag hos laboratorierne. Hvis $u_x \leq 0,3 \cdot \hat{\sigma}$ er usikkerheden på den nominelle værdi ubetydelig /5/.

For at mindske indflydelsen fra variationer i gentagelse, kan antallet af bestemmelser øges, og alle laboratorier udfører det samme antal bestemmelser /5/.

3.4 Stabilitet og homogenitet

For at mindske udgifterne til de sammenlignende prøvninger udelades decidede tests for stabilitet og homogenitet og i stedet prioriteres et større antal prøvninger. Da de sammenlignende prøvninger foregår over et begrænset geografisk område, kan prøverne leveres til laboratorierne hurtigt efter fremstilling, og herved kan parametre, som er forholdsvis ustabile, muligvis også medtages /8/.

Homogenitet kan undersøges i forbindelse med, at prøverne analyseres ved brug af enten et staggered-nested eksperiment (ISO 5725-3, afsnit 9.5) eller et fully-nested eksperiment (ISO 5725-3, afsnit 9.4). Et fully-nested eksperiment kræver flere analyser end et staggered-nested eksperiment, hvorfor sidstnævnte metode anbefales. Homogeniteten kontrolleres ved en variansanalyse.

Et fully-nested experiment til bestemmelse af homogenitet udføres typisk på ét laboratorium. Et passende antal prøver analyseres ved dobbeltbestemmelse og alle analyserne udføres i samme analyseserie.

Et staggered-nested experiment kan f.eks. være, at alle deltagere analyserer mindst to beholdere af hver prøve og udfører analysen for den ene beholder som dobbelt- eller flerdobbelt bestemmelse. Alle analyserne udføres for hver deltager i samme analyseserie som et led i den sammenlignende prøvning.

Kriteriet for tilstrækkelig homogenitet knyttes til størrelsen af $\hat{\sigma}$ på samme måde som kriteriet for standardusikkerhed på den nominelle værdi, dvs. at standardafvigelsen mellem flasker højst må være $0,3 \cdot \hat{\sigma}$. Herved sikres at heterogenitet ikke får målelig indflydelse på laboratoriernes z-score.

Inddragelse af homogenitetstest giver et sikrere resultat af den sammenlignende måling, men øger omkostningerne. Afhængig af omkostningerne til homogenitetstests kan det overvejes, om en større frekvens af sammenlignende prøvning eller analyse af flere identiske beholdere hos hver deltager kan være en mere cost-effektiv løsning.

3.5 Rapportering

Rapportering af resultater til de øvrige deltagere skal indeholde beskrivelse af, hvordan den nominelle værdi er opnået, hvilke resultater der er brugt til bestemmelsen, standardusikkerheden på den nominelle værdi, og eventuelle tests for homogenitet.

4 Referencer

- /1/ Bekendtgørelse nr. 900 af 17. august 2011 om kvalitetskrav til miljømålinger.
- /2/ DS/EN ISO/IEC 17025: Generelle krav til prøvnings- og kalibreringslaboratoriers kompetence. 2005.
- /3/ EA-4/18: Guidance on the level and frequency of proficiency testing participation. 2010.
- /4/ Eurachem: Selection, Use and Interpretation of Proficiency Testing (PT) schemes. 2011.
- /5/ ISO 13528: Statistical methods for use in proficiency testing by interlaboratory comparison. 2002.
- /6/ ISO/IEC 17043: Conformity assessment – General requirements for proficiency testing, 2010.
- /7/ I. Kuselman and A. Fajgelj: IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants – chemical analytical laboratories (IUPAC Technical Report), *Pure Appl. Chem.*, Vol. 82, No. 5, pp. 1099-1135, 2010.
- /8/ By- og Landskabsstyrelsens Referencelaboratorium for Miljøkemiske Analyser: Metoder til ekstern kvalitetssikring, Rapport 2008.

Bilag A - Litteraturliste

Total litteraturoversigt

By- og Landskabsstyrelsens Referencelaboratorium for Miljøkemiske Analyser: Metoder til ekstern kvalitetssikring, Rapport 2008.

DHI: Udvikling af præstationsprøvninger med henblik på sikring af pålidelige miljødata, Rapport 2000.

Eurachem: Selection, Use and Interpretation of Proficiency Testing (PT) schemes, 2011.

ISO/IEC Guide 43-2: Proficiency testing by interlaboratory comparison – Part 2: Selection and use of proficiency testing schemes by laboratory accreditation bodies, 1997.

I. Kuselman and A. Fajgelj: IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants – chemical analytical laboratories (IUPAC Technical Report), *Pure Appl. Chem.*, Vol. 82, No. 5, pp. 1099-1135, 2010.

ISO/IEC 17043: Conformity assessment – General requirements for proficiency testing, 2010.

ISO Guide 35: Reference materials – General and statistical principles for certification, 2006.

ISO TS 16489: Water quality – Guidance for establishing equivalency of results, 2003.

ISO 13528: Statistical methods for use in proficiency testing by interlaboratory comparison, 2002.

M. Thompson et al.: The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report), *Pure and Applied Chemistry* 78, 145-196, 2006.

ISO TS 13530: Water quality – Guidance on analytical quality control for chemical and physicochemical water analysis, 2009.

ILAC-P9: Policy for Participation in National and International Proficiency Testing Activities, 2005.

ILAC-G13: Guidelines for the Requirements for the Competence of Providers of Proficiency Testing Schemes, 2007.

EA-4/18: Guidance on the level and frequency of proficiency testing participation, 2010.

DS/EN ISO/IEC 17025: Generelle krav til prøvnings- og kalibreringslaboratori-ers competence, 2005.