

Kontrolstatistik – dokumentation Vandkemi

Version: 1

Sidst revideret: januar 2013

Emne:	vandkemi (vandløb, sø, marin)
Dato:	Jan. 2013
Filer:	(Navn + placering)
Periode:	Programmet anvendes til kontrol af data fra 2012 og frem
Kørsel af program:	Dagligt
Input-data:	Vandkemi i perioden 1995 til året før data til kontrol
Aggregeringsniveau:	vandløb: kvartal Sø: halvårlig (2. og 3. kvartal samt 1. og 4. kvartal) Marin: hydrologiske kvartaler
Anvendte grænser:	1 og 99 percentil, hvor muligt (dvs. antal obs>100) ellers min og max-værdier Værdier fra litteratur, for nye parametre, hvor antal obs<10
Kontrol-niveauer:	5 1: per observationssted/målested 2: per stations-type (sø: søtype, marin: havområde) 3: per sæson: tidsaggregering (vandløb: kvartal, søger: sommer/vinter, marin: hydrologisk kvartal) 4: global (alle værdier for parameteren indgår i statistik) 5: hardkodet opslagstabel (min/max-værdier)

Beskrivelse

Procedure og metode for kontrol af nye indkomne målte vandkemiske værdier i ODA-databasen fra 2012 og fremadrettet. Før 2006 er data fagligt godkendt af de daværende amter. Fra 2006-2011 er næringsstoffer kontrolleret på basis af tidligere data men på et lidt grovere niveau end nærværende kontrolprocedure - og alle data er fagligt godkendt af naturstyrelsen. På nuværende tidspunkt figurerer derfor anseeligt antal godkendte data for de mest anvendte parametre i ODA. Disse data anvendes til at fastsætte et realistisk niveau for værdier for hver måleparameter (outlierstatistik), som nye værdier kontrolleres op imod, for at hindre fejlagtige og usandsynlige værdier at komme ind i databasen. Falder værdien inden for de beregnede grænser accepteres værdien, mens en fejlmelding gives og værdien sættes under kontrol, hvis den falder uden for de beregnede grænser.

Der anvendes 3 metoder til at fastsætte grænser.

1: 1 og 99 percentil-grænser: Er værdien højere end den værdi, som repræsenterer 1 % af de højeste værdier i fordelingen af alle værdier eller tilsvarende lavere end den værdi som repræsenterer 1 % af de laveste værdier, markeres værdien som potentiel 'outlier' og placeres 'under kontrol', hvor værdien efterfølgende skal godkendes af en fagperson..

2: min-max: Grænserne består af den laveste værdi i kontroldatasættet som minimum og den højeste værdi i kontroldatasættet som maksimumværdi. Falder værdien, der kontrolleres hhv. under eller over disse værdier, markeres værdien som potentiel 'outlier' og placeres 'under kontrol'.

3: faste grænser baseret på generelle retningslinjer eller ekspertvurdering. Falder værdien, der kontrolleres hhv. uden for disse værdier, markeres værdien som potentiel 'outlier' og placeres 'under kontrol'.

Metoden, der anvendes er bestemt af hvilket niveau af kontrol der foretages (se diagram s. 4)

Hyppighed for kørsel af program til generering af outlier-statistik

Dagligt.

Fordele: Der opbygges hurtigt grundlag for kontrolstatistik for nye parametre, herved reduceres fejlmeldinger betydeligt for nye parametre.

Øvrige betragtninger. Data kontrolleres på forskelligt niveau over året, i modsætning til en årlig kørsel hvor alle data det enkelte år kontrolleres med samme kontrolgrænser.

Input-data

Data fra 1995 og frem vælges at kunne indgå i kontrolstatistikken. Dels er data mere homogene inden for målestedet, da der i første del af overvågningsperioden (1989-1995) skete betydelige ændringer. Dels er kvaliteten af data af varierende karakter i data fra den første periode af overvågningen.

Sæson (tidsaggregationsniveau)

Vandløb: en stor del af parametrene følger variation i vandafstrømningen, som varierer betydeligt kvartalsvist. Alle parametre kontrolleres kvartalsvist, hvis muligt.

Søer: mange vandkemiske parametre reagerer på den produktive periode (forår-sommer), hvor også hyppigheden af prøver er størst. Fra 2004 kræves kun 1 vinterprøve i NOVANA-programmet, hvorfor kvartalsvis inddeling af prøvetagninger ikke er hensigtsmæssig. Derfor er inddelingen halvårlig (sommer (2. og 3. kvartal)/vinter (1. og 4. kvartal)). Alle parametre kontrolleres på sommer/vinter-niveau hvis muligt.

Marin: en del af parametrene følger sæsonvariation. Disse er opdelt i hydrologiske kvartaler (dec+ jan+ feb= kvartal 1, marts+april+maj=kvartal2 osv). Disse parametre kontrolleres kvartalsvist, hvis muligt.

Anvendte grænser for outliers

Metode	Antal observationer		Grundlag
1 og 99 percentiler ¹⁾	>100	per parameter, prøvefraktion, enhed	Baseret på indeværende data fra 1995 til året før kontrol-år
Min og max værdier	<100	per parameter, prøvefraktion, enhed	Baseret på indeværende data fra 1995 til året før kontrol-år
Faste pre-definerede grænser	<10 totalt i kontroldatasættet	per parameter, enhed	Baseret på generelle retningslinjer eller ekspertvurdering

1) 1 og 99 percentiler er for hver parameter (per prøvefraktion, enhed) anvendt som hhv. nedre og øvre grænse for outliers. For normalfordelte data ligger disse niveauer omtrent mellem middelværdien +/- 2*standardafvigelsen og middelværdien +/- 3*standardafvigelsen. Dette er tjekket grafisk for udvalgte parametre og ser fornuftigt ud for en stor del af parametrene (for eksempel, se Fig. 1).

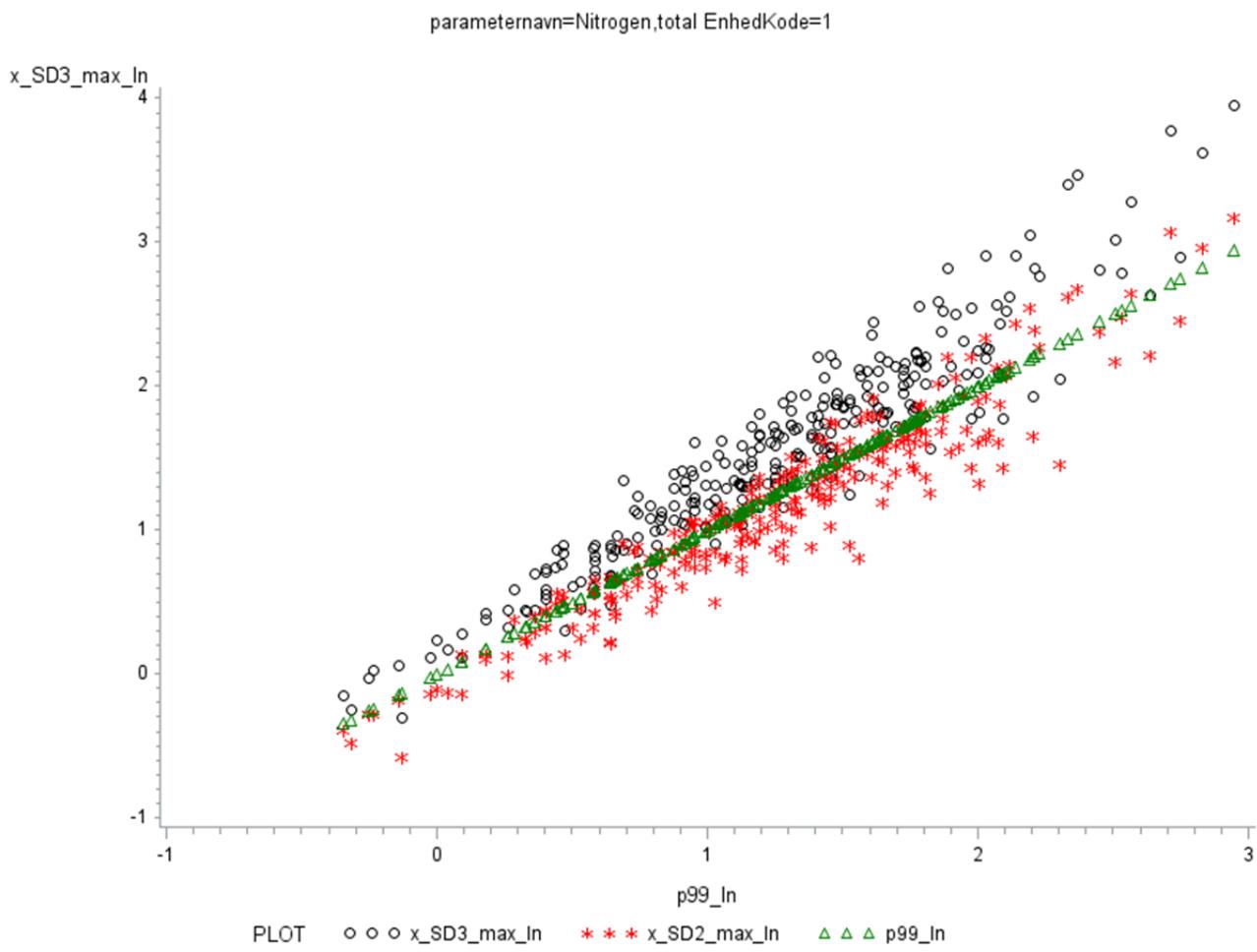
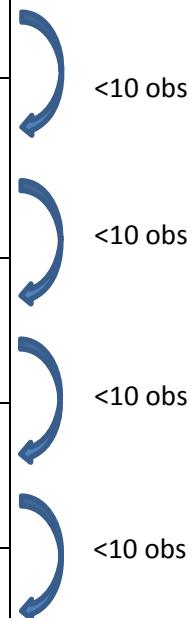


Fig. 1 99 percentilen (grønne trekantede) plottet mod hhv. 2*standardafvigelsen (røde stjerner) og 3*standardafvigelsen (sorte cirkler) for ln-transformerede sø-koncentrationer af total P

kontrolniveauer

Niveau	Kontrolniveau	Anvendes ved:	Anvendt grænse	Anvendes for emne
1	Målesteds	≥ 10 obs per parameter per målestation per sæson	1-99 % *	Sø Vandløb
2	Stationstype/område	≥ 10 obs per parameter per stationstype per sæson	1-99 % *	Sø Marin (udvalgte parametre)
3	Sæson	≥ 10 obs per parameter per sæson	1-99 % *	Sø Vandløb
4	Global	≥ 10 obs per parameter globalt	1-99 % *	Sø Vandløb Marin
5	Hardkodet opslagstabel	< 10 obs per parameter globalt	Fastsatte pre-definerede grænsler	Sø Vandløb Marin



* reelt vil percentiler være lig med min/max-værdier ved mindre end 100 observationer

1: per observationssted/målesteds

Parameteren (per prøvefraktion, enhed) på observationsstedet kontrolleres per observationssted/målesteds hvis der er 10 eller flere observationer for parameteren per aggregeringsniveau (vandløb: kvartal, sø: sommer/vinter) per observationssted/målesteds.

Hvis der er færre end 10 observationer kontrolleres værdien på niveauet lavere (stationstype)

Dvs. det kan variere fra parameter til parameter på det enkelte observationssted på hvilket niveau observationen kontrolleres.

2: per stations-type (søtype, område)

Dette kontrol-niveau er tilføjet med henblik på at kunneindsnævre range af outliers ved at reducere variationen indenfor parameteren ved opdeling i 'stations-typer' eller områder. Herved vil en høj værdi inden for en type med generelt lave værdier fanges (og omvendt), mens dette ikke vil være tilfældet, hvis der ses globalt på parameter-range. Der defineres kun én stationstype per observationssted/målesteds. Dvs. alle parametre kontrolleres i forhold til denne ene type-opdeling, og opdelingen vil derfor være mere hensigtsmæssig for nogle parametre end for andre.

Hvis der er færre end 10 observationer per sæson per stationstype kontrolleres værdien på niveauet lavere (sæson)

For Marine parametre foretages kontrol på områdeniveau for udvalgte parametre, de øvrige kontrolleres på globalt niveau.

Sø:

Type-kontrol forudsætter for *sø-data* at stationen (søen) tidligere er blevet undersøgt for de 3 parametre, der definerer sø-typer ud fra medianværdier (areal, alkalinitet, total-P) for søen. I tabellen for kontrolstatistik er hvert observationsstednr/målested linket til en sø (soenr) med en søtype, hvor dette er muligt

Søtype-definition:

- baseret på inddelinger i den danske tilgang til type-opdeling i forhold til vandrammedirektivet

Søareal	<0.1	$\geq 0.1-0.5$	≥ 5 ha
Total alkalinitet	<0.2meq/l		
Total fosfor	<0.025	$\geq 0.025-0.050$	$\geq 0.050-0.100$ $\geq 0.100-0.200$ ≥ 0.200 mgTP/l

Øvrige vandrammedirektiv-typologi faktorer, der ikke er medtaget i søtype-inddeling med henblik på validering af vandkemiske parametre:

Salinitet: afspejles til dels også i TP-niveau (saline søer er ofte mere næringsrige)

Dybde: afspejles til dels også i TP-niveau og sø-areal (dybe søer er ofte større søer og mindre næringsrige)

Farve: stort range indenfor både TP-niveau og alkalinitet, ydermere er der ikke så mange data for denne parameter (først efter 2004)

Der er potentielt 30 søtyper.

MARIN

De marine observationssteder er inddelt i over 100 områder baseret på 2. ordens hydrologiske referencer med underopdelinger f. eks. for fjorde. For CTD-parametrene salt, temperatur og ilt samt de vandkemiske næringsstoffer er der beregnet outlier-grænser for hvert område og kvartal hvor der er over 10 observationer. Hvis der er under 10 observationer inddrages nærtliggende områder i beregningerne. Derudover bruges globale grænser.

3: sæson (alle værdier for parameteren per tidsaggregering (sæson) indgår i statistik)

Observationen for parameteren (per prøvefraktion, enhed) kontrolleres per sæson-aggregeringsniveau (vandløb: kvartal, sø: sommer/vinter) med grænserne 1-99 centiler hvis der 10 observationer eller mere.

Hvis der er færre end 10 observationer per aggregationsniveau (vandløb: kvartal, sø: sommer/vinter) kontrolleres værdien på niveauet lavere (global)

4: global (alle værdier for parameteren indgår i statistik)

Observationen for parameteren (per prøvefraktion, enhed) kontrolleres med grænserne 1-99 percentiler hvor det er muligt (>100 observationer) og ellers med grænserne min/max.

5: Hardkodet opslagstabel

Er der færre end 10 observationer globalt for parameteren kontrolleres værdien i en tabel med fastsatte grænser baseret på generelle retningslinjer eller ekspertvurdering.

Findes parameteren ikke i denne tabel returneres værdien med fejlmelding 'ingen kontrolgrænse for denne værdi' – og opslagstabellen skal opdateres med parameteren.

For miljøfremmede stoffer i vand genereres hardkodede grænser baseret på detektionsgrænser (Det Nationale Overvågningsprogramfor Vand og Natur. NOVANA 2011-2015. Programbeskrivelse, 2011) samt generelle nationale miljøkvalitetskrav (Bekendtgørelse om miljøkvalitetskrav for vandområder og krav til udledning af forurenende stoffer til vandløb, søer eller havet (nr. 1022, 2010) som hhv. min og max-grænse.