

Rutiner för överföring av data till Datavårdskapet för sjöar och vattendrag vid Institutionen för vatten och miljö vid SLU

Institutionen för vatten och miljö vid SLU driver på Naturvårdsverkets uppdrag Datavårdskapet för sjöar och vattendrag. I uppdraget ingår att samla in, kvalitetsgranska, lagra och tillhandahålla vattenkemiska och biologiska data av nationell och regional karaktär.

För att underlätta insamlingen och minimera risken för felinmatningar är det önskvärt att data överförs på ett så standardiserat sätt som möjligt. För vattenkemiska data återfinns en rapporteringsmall på datavårdskapets registersida (<http://www.slu.se/vatten-miljo/DV-registersida>). För biologiska data håller leveransmallar på att utvecklas. Om du inte finner en färdig mall för den typ av data du skall leverera var vänlig och kontakta datavårdskapets kontaktperson före leverans (Anders.Stenstrom@vatten.slu.se eller per tel. 018-673132).

Kvalitetssäkrade data

De data som levereras från olika tillsyningsmyndigheter, konsulter och/eller vatten(vårds)förbund förutsätts redan vara kvalitetssäkrade, först och främst av analyserande laboratorium, men även av berörd tillsynsmyndighet. Vid överföringen till Institutionen för vatten och miljö kommer ytterliggare en kvalitetsgranskning att ske innan materialet görs tillgängligt på Internet. Denna granskning kommer att omfatta bl.a. en rimlighetsanalys och jämförelser med tidigare undersökningsresultat, säkerställande av provtagningsplats mm. Inga förändringar av materialet kommer dock att ske utan samråd med berörd tillsynsmyndighet och/eller analyserande laboratorium.

Vid granskningen kan eventuellt kvarvarande oklarheter som av någon anledning inte kan redas ut, markeras i det slutgiltiga datamaterialet som observationer inom parentes. Detta gör att vid en ev. nedladdning av materialet för användning i exempelvis MS Excel, kommer dessa observationer att klassificeras som text. Detta innebär att dessa avvikande poster inte kan användas vid t.ex. statistisk bearbetning eller figurframställning utan att parenteserna tas bort. Det blir följaktligen den som använder materialet som slutligen kommer att tvingas ta ställning till om den/de avvikande observationerna bör ingå eller ej.

Dataleverantörerna uppmanas att inte byta ut observationer som är lägre (eller högre) än eventuella detektionsgränser mot något subjektivt valt värde. För att kunna utföra t.ex. statistiska bearbetningar byts i vissa fall "<-värden" ut mot exempelvis hela, halva eller en tredjedel av värdet för den aktuella detektionsgränsen. Detta utgör ett stort bekymmer när detektionsgränser förändras med tiden och man till sist har ett datamaterial där man inte kan avgöra om enskilda observationer är utanför detektionsområdet eller korrekta värden. Även i detta fall skall < eller > tecknen kvarstå vid leveransen, vilket innebär att vid användandet av

ett material med dylika tecken, måste användare besluta om dessa observationer skall användas vid en ev. bearbetning och i så fall med vilket värde dessa skall bytas ut mot.

Vad skall överföras?

Data från vattenkemiska och biologiska undersökningar (ej fisk, se nedan) i sjöar och vattendrag. För mer information rörande vilka parametrar och enheter som behandlas se respektive leveransmall. OBS! Fler parametrar kan i vissa fall läggas till. Detta skall då anges i följebrevet till leveransen.

Vad skall **inte** överföras?

Uppgifter rörande badvattenkvalitet (Datavärd: Smittskyddsinstitutet), miljögifter i biota (IVL) och provfisken (Fiskeriverkets sötvattenslaboratorium) skall inte tillsändas Institutionen för vatten och miljö, utan skall sändas till respektive datavärd.

Lämpliga filformat

Data överlämnas helst i MS Excel-format, men även textformat går bra. Om textformat används bör data vara antingen vara separerade med tabb-tecken eller semikolon (SKV-format).

Dataindelning

Om både vattenkemiska och biologiska data levereras samtidigt bör dessa vara åtskilda i två separata filer eller kalkylblad (MS Excel

Temperatur- och syrgasprofiler i sjöar levereras som en separat fil/kalkylblad (en för samtliga sjöar).

Även vattendrags- och sjöstationer skall hållas åtskilda i separata filer eller kalkylblad (MS Excel), dvs. en fil/kalkylblad för vattendragsstationer och en för sjöstationer.

Rubriker/namngivning

Variabler bör anges med både variabelnamn (t.ex. Alkalinitet) och enhet (mekv/l) i samma textruta och **inte** uppdelat på flera rader (BILAGA 1). Om någon av matriserna (t.ex. vattenkemi) är uppdelad på flera filer/tabeller måste rubrikerna vara de samma i samtliga fall, annars finns det risk att data försvinner vid överföringen till databasen.

Angivning av provtagningsobjekt och provtagningsplats

Provtagningsplatsens koordinater skall anges enligt Rikets nät (X-Y) eller SWEREF TM (tydligt utmärkt vilket format som använts). Om provtagningsobjektet är en sjö skall dessutom

utloppets koordinater anges enligt SMHI:s identifikationssystem (SVAR). Samtliga koordinater anges i separata kolumner namngivna enligt bilaga 1. Utöver dessa generella identifikationsangivelser bör även en kolumn med stationsnamn, samt en kolumn med ett ev. lokalt id anges. Dessa uppgifter ges för samtliga observationer (poster/rader i filen).

Provtagningsdatum

Datum skall anges som år, månad och dag i tre **separata** kolumner för att inte riskera fel vid överföringar mellan olika typer av datorer. Tips! Om datum finns i en Excel-kolumn går det lätt att dela upp denna till tre kolumner m.h.a. kommandot ”text till kolumner” (data-menyn).

Provtagningsdjup

Aktuellt provtagningsdjup skall alltid anges, inte vilket djup som borde ha provtagits. För biologiska parametrar kan antingen enskilda provtagningsdjup (t.ex. 2 m) eller provtagningsintervall vid samlingsprov (t.ex. 0–20 m) anges. För vattenkemiska parametrar kan för närvarande endast enskilda provtagningsdjup tas emot. Om samlings-/integrerade prov har tagits bör **samråd** med datavärden ske innan leverans sker.

Djupprofiler

Omfattande undersökningar av temperatur och/eller syrgashalter i djupprofiler kan med fördel levereras som separata tabeller. Är det däremot endast något enstaka djup som har undersökts extra, så är det bättre att det ingår i huvudtabellen.

Datakvalitet

En rimlighetsanalys av data bör ske i alla led innan dessa levereras vidare, dvs. av analyserande laboratorium, eventuellt vattenvårdsförbund och ansvarig tillsynsmyndighet. Analysen kan t.ex. innebära att man jämför med gamla undersökningar och kontrollerar avvikande observationer. Avvikande observationer kan man t.ex. definiera som ett långtidsmedelvärde ± 2 standardavvikelser (95%-igt konfidensintervall) eller annat lämpligt intervall. Rimlighetsanalysen kan även inkludera mer manuella tester genom att man gör figurer där halter avsätts mot tiden och/eller avsätter olika relevanta parametrar mot varandra (t.ex. Ca/Mg, Alk/Ca, Alk/pH, $\sum(\text{Ca, Mg, Na, K})/\text{konduktivitet}$, klorofyll/växtplankton-biomassa. Rimlighetsanalys utförs dessutom som rutin av den mottagande datavärden.

Avrundningar

Data bör bestå av ett rimligt antal värdesiffror. Omotiverad avrundning bör däremot inte ske, t.ex. att kvävehalter i vissa fall avrundas till hela tiotal eller femtiotal. Denna typ av avrundning gör det svårt att på ett rättvist sätt kunna bearbeta materialet vidare (t.ex. genom att beräkna årsmedelvärden e.dyl.), då avrundningsfelet kan bli avsevärt. Denna typ av

avrundningar hör hemma i samband med bearbetade data i t.ex. årsrapporter och inte i en databas. Minst tre värdesiffror bör om möjligt anges.

Anmärkningar

Relevanta kommentarer eller anmärkningar för enstaka provtagningstillfällen och/eller enstaka analyserade parametrar kan anges i speciell kolumn som namnges ”Anmärkningar”. Här bör bland annat ingå anmärkningar som har att göra med t.ex. den specifika provtagningssituationen (t.ex. mycket kraftig vind, ihållande regn, sjön var täckt med 0,5 m snö). Denna typ av information kan vara värdefull vid en utvärdering av materialet.

Kringuppgifter/metadata

För att erhålla en så heltäckande bild som möjligt är det viktigt att relevant kringinformation bifogas. Denna typ av information är bl.a. viktig vid kvalitetsgranskningen. Exempel på relevant kringinformation är: provtagare, ansvariga för kvalitetssäkring och leverans. Denna typ av uppgifter lämnas separat, antingen i den checklista som bifogas leveransen eller som annan separat fil.

Checklista för leverans av data till Institutionen för vatten och miljö vid SLU

Undersökningsprogram och tidsperiod

.....

Typ av data Vattenkemi Biologi Annat.....

Uppgiftslämnare (inkl e-postadress)

.....

Kvalitetsansvarig (inkl e-postadress)

.....

Filformat? MS Excel Text (SKV/tab) Annat.....

Totalt antal filer/kalkylblad i leveransen

.....st filer.....st kalkylblad

Har datamaterialet kvalitetssäkrats? Ja Nej Vet ej

Om ja, av vem och när.....

Finns osäkra värden markerade (satta inom parentes)? Ja Nej

Hur är värden utanför resp. variablers detektionsgräns behandlade?

< eller > tecken Annat sätt..... Ej behandlat Vet ej

Har programmet förändrats under senaste året?

Nya stationer Nedlagda stationer Nya mätvariabler Annat

Övrig information

.....

.....

.....

.....

.....

.....

.....