



**DNMI**

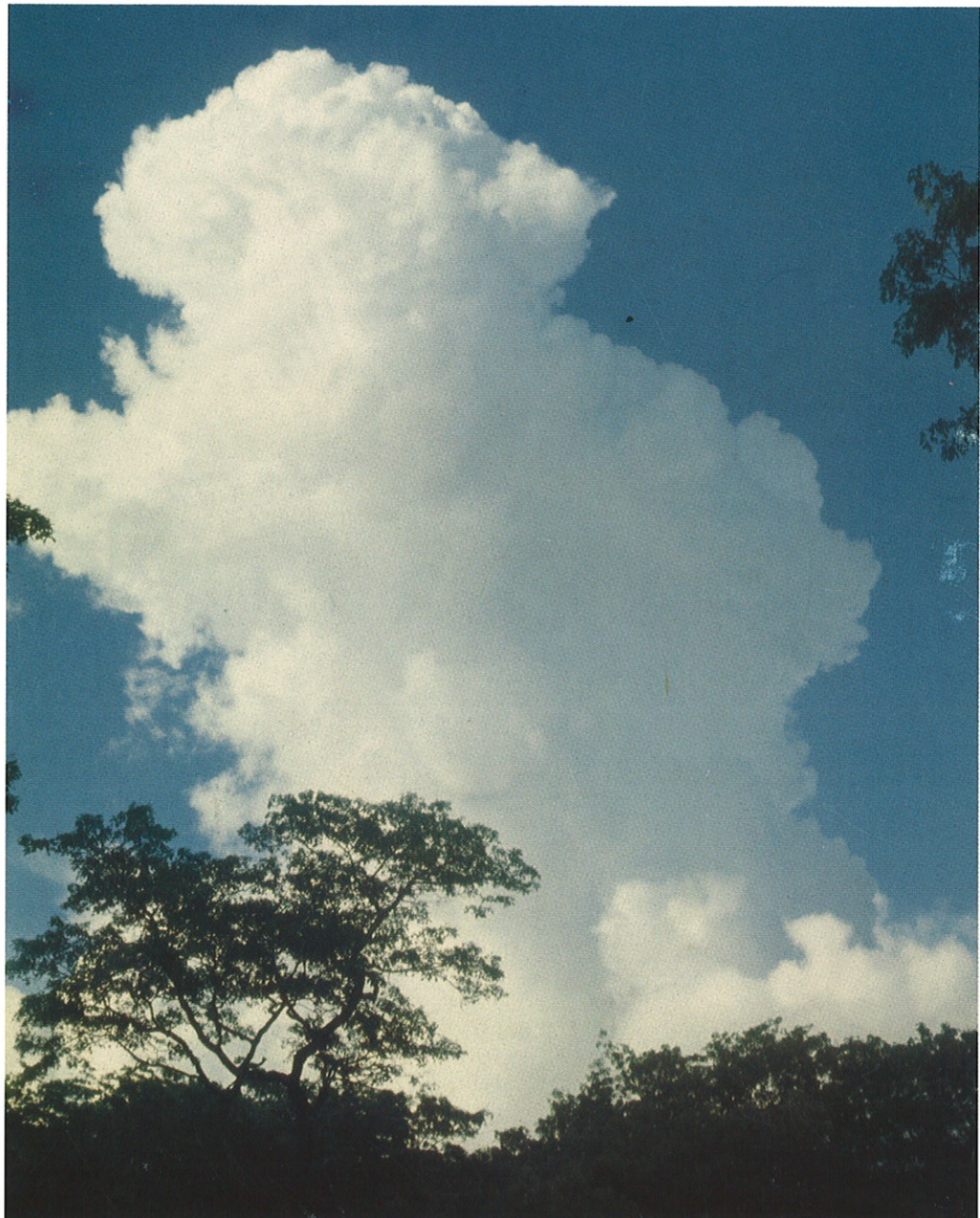
Det norske meteorologiske institutt

RAPPORT NR. 25/99

**KLIMA**

**DATRUT-prosjektet:  
Skisse av dataflyt, datakontroll og  
databasestruktur i en fremtidig helhetlig  
behandlingsrutine ved klimaavdelingen**

Lars Andresen, Per Ove Kjensli,  
Zbigniew Toporowski, Åse Moen Vidal, Petter Øgland



# DNMI-RAPPORT

DET NORSKE METEOROLOGISKE INSTITUTT  
POSTBOKS 43 BLINDERN 0313 OSLO 3  
TELEFON: 22 96 30 00

ISBN 0805-9918

RAPPORT NR.

**25/99 KLIMA**

DATO

28.09.99

TITTEL

**DATRUT-PROSJEKTET:  
SKISSE AV DATAFLYT, DATAKONTROLL OG  
DATABASESTRUKTUR I EN FREMTIDIG HELHETLIG  
BEHANDLINGSRUTINE VED KLIMAAVDELINGEN**

UTARBEIDET AV

**Lars Andresen, Per Ove Kjensli,  
Zbigniew Toporowski, Åse Moen Vidal, Petter Øgland**

OPPDRAGSGIVER

**DNMI - Klimaavdelingen**

OPPDRAGSNR.

SAMMENDRAG

Skissen viser flyt av observasjonsdata fra værstasjonene til DNMI, i henhold til mandatet for DATRUT-prosjektet. Underveis foregår en kontroll av data på forskjellig nivå i organisasjonen. Første kontroll skjer på stasjonen, dernest ved mottak på DNMI og i forbindelse med sann tids bruk av dataene og til sist i rutinekontrollene i ettertid ved Klimaavdelingen. Det legges stor innsats i tilbakemeldinger og avviksmeldinger.

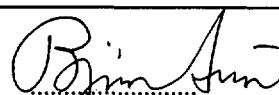
Forskjellige metoder i kvalitetskontrollen av observasjonsdata er skissert. Slike kontroller vil være basert både på stasjonens egne data og data fra andre stasjoner. I forbindelse med kontrollen produseres feillister, feilstatistikk og rettestatistikk. Korreksjoner og interpolerte data flagges. Algoritmer for automatisk kontroll og GIS for romkontroll blir viktige hjelpemidler i en fremtidig databehandlingsrutine.

For at det skal være mulig å garantere et konsistent historisk datalager og å gjennomføre den skisserte kvalitetskontroll av dataene, må databasen ha en hensiktsmessig struktur. Den vil bestå av arbeidslagre som kan håndtere alle aktuelle observasjonsdata og temporære lagre som er tilrettelagt for romkontroller og andre typer kontroller. I fremtiden vil sannsynligvis fly-, radar- og satellittdata bli brukt i kontrollarbeidet.

UNDERSKRIFT



Lars Andresen  
SAKSBEHANDLER



Bjørn Aune  
FAGSJEF

# INNHALDSFORTEGNELSE

<b>Innholdsfortegnelse.....</b>	<b>1</b>
<b>1. Innledning.....</b>	<b>3</b>
<b>2. Fremtidig kvalitetskontroll av meteorologiske observasjonsdata.....</b>	<b>3</b>
2.1. Forutsetninger.....	3
2.2. Kontroller på forskjellig nivå.....	4
2.3. Metoder i kvalitetskontrollen.....	6
2.4. Databasestruktur.....	10
2.4.1. Klimadatabasen.....	10
2.4.2. Klimadatavarehus.....	11
2.5. Manglende data.....	13
2.6. Flagging.....	13
2.7. Sikring av originaldata.....	13
<b>Appendiks.....</b>	<b>15</b>
<b>Appendiks A. Mandatet.....</b>	<b>16</b>
<b>Appendiks B. Definisjoner.....</b>	<b>17</b>
B1. Observasjonsstasjoner og data.....	17
B2. Observasjoner, værsymboler og værkoder.....	18

## 1. Innledning

I notat KLIMA 14/99 er mandatet for DATRUT-prosjektet beskrevet (se Appendiks A). Første punkt lyder slik:

Skissere en fremtidig, helhetlig behandlingsrutine for alle typer observasjonsdata som Klimaavdelingen har ansvaret for lagring av. Skissen skal definere sammenhengen mellom dataflyt, datakontroll og databasestruktur.

Denne rapporten inneholder en slik skisse og viser hva vi i DATRUT-prosjektet tenker om en fremtidig databehandlingsrutine. Med fremtidig mener vi alle meget klart "så snart som mulig". Men skal vi komme videre fra der vi er nå, må det settes inn tilstrekkelige ressurser for den praktiske delen av arbeidet. DATRUT-prosjektet er kun et første skritt i en slik prosess.

I neste rapport fra DATRUT-prosjektet vil vi komme inn på de øvrige punktene i mandatet og komme med forslag til hvordan vi bør gjennomføre den praktiske overgangen mellom nå-situasjonen og den situasjonen vi ønsker for fremtiden.

## 2. Fremtidig kvalitetskontroll av meteorologiske observasjonsdata

### 2.1. Forutsetninger

Kvalitetskontrollen skal håndtere data fra alle eksisterende stasjonstyper, dvs. manuelle værstasjoner (synop, maritime, klima), delvis automatiserte værstasjoner, flyplasstasjoner (metar), nedbørstasjoner, fordampningsstasjoner, solstrålingsstasjoner (bl.a. stasjoner med solskinnautograf), automatstasjoner av ulike slag (synop, maritime, landbruk, oppdrag, nedbørintensitet) og radiosondestasjoner. Stasjonstypebetegnelse er definert i Appendiks B1.

Alle manuelle værstasjoner, delvis automatiserte værstasjoner og øvrige manuelle stasjoner skal være utstyrt slik at det er mulig å utføre en førstekontroll på stasjonen, med rapportering til observatøren. På automatstasjoner skal det foretas en tilsvarende førstekontroll, f.eks. ved en ev. dublering av sensorer, og programvare, som gjør det mulig å produsere "feilmeldingsflagg", som kan følge datastrømmen fra stasjonen. Dermed blir det mulig for Obs- og Metdivisjonen å ta rask aksjon når det oppstår feil på stasjonene. Kontrollene på de ulike nivåer er beskrevet i kap. 2.2.

Det skal lagres data fra nabolandenes stasjoner i den grad det er nødvendig av hensyn til vår egen kvalitetskontroll og av hensyn til klimabeskrivelser, klimatologiske utredninger og værdokumentasjon i nærområdene til disse landene.

Det bør legges til rette for en sanntids forsendelse av data fra alle typer værstasjoner.

Med noen få unntak bør største observasjonshyppighet for klimatologiske formål ikke overstige 1 per time. For øvrig kan observasjonshyppigheten være 1-8 per døgn (fortrinnsvis ved standardtidspunktene definert av WMO). Observasjonsbetegnelser, værsymboler og værkoder er definert i Appendiks B2.

Det bør være en oppgave for DNMI å tilrettelegge fly-, radar- og satellitt-observasjoner for forsknings- og klimatologiske formål. (I en viss utstrekning benyttes slike data i dag allerede

for værvarslingsformål.) Det må minst legges til rette for å benytte slike data i romkontroll og interpolering av klimatologiske data. Numeriske analysedata (korttidsprognose 6 timer frem) bør være tilgjengelige i sann tid for alle kontroller som gjøres på DNMI.

## 2.2. Kontroller på forskjellig nivå

Kvalitetskontrollen vil starte allerede ute på stasjonen og fortsette straks dataene er ankommet DNMI. Vi tenker oss derfor kontroller i flere trinn. Som overskrifter benyttes terminologien i NORDKLIM-prosjektet (med unntak av kontrollene på selve stasjonen, som der ikke er beskrevet). AQC står for "automatic quality control" og er ment å være en helautomatisk kontroll. HQC står for "human quality control" og tilsvarer vår manuelle døgn-/uke-/månedskontroll.

I dag sendes en vesentlig del av meteorologiske observasjonsdata via post til DNMI / Klimaavdelingen. For fremtiden tenker vi oss at samtlige manuelle værstasjoner (både VS- og VK-stasjoner) blir utstyrt med PiO-programvare, slik at fullstendige observasjoner sendes elektronisk etter at dataene har gjennomgått en objektiv kvalitetskontroll på stasjonen. Dermed reduseres også registreringsarbeidet ved DNMI. De fleste manuelle maritime stasjoner (skip) har allerede i noen år hatt tilsvarende utstyr (programvare Turbo, utviklet av KNMI, Nederland), og har erfart en bedring i datakvaliteten.

Tilsvarende forventes det at nedbørstasjoner vil kunne overføre sine data elektronisk.

Også tradisjonelle analoge registreringer vil forsvinne når "instrumentelle" værparametre registreres på timebasis med aktuell verdi og min./maks.-verdier. I en overgangsperiode vil man likevel benytte analoge registreringer på referansestasjoner, inntil datakvalitet/-regularitet er tilfredsstillende. Elektronisk datalagring på diskett tenkes erstattet av direkte overføring eller overføring en gang per døgn.

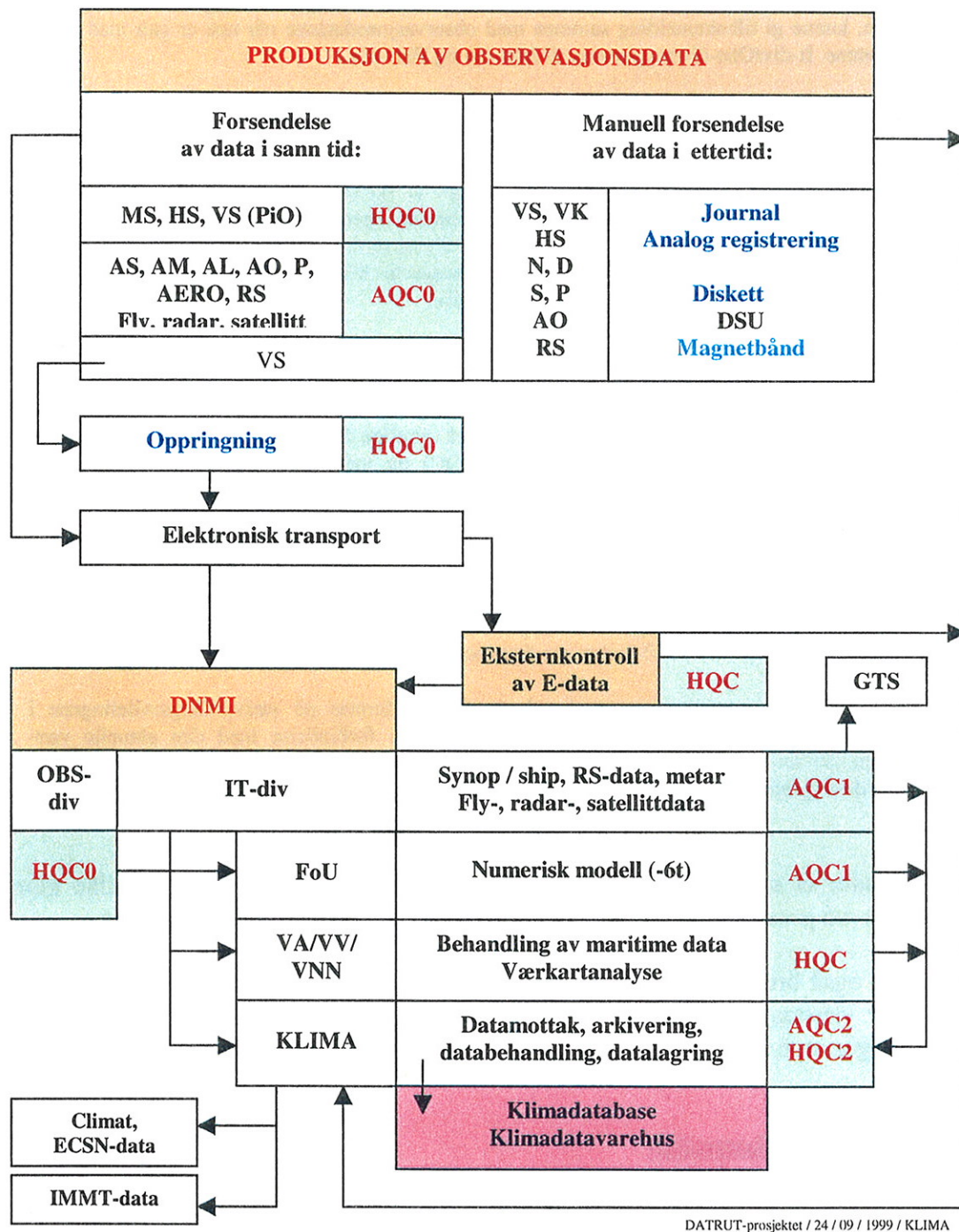
På DNMI vil det være IT-divisjonen som har hovedansvaret for distribusjonen av observasjonsdata inn i organisasjonen. OBS-divisjonen bør imidlertid ha mulighet for å ringe opp automatstasjonene direkte og dermed overvåke stasjonene, med tanke på service, vedlikehold og rask aksjon ved avvik fra normaltilstanden.

Dataene vil gjennomgå en kontroll i IT-divisjonen. Alle synopdata blir tilegnet et flagg (konfidensverdi, kvalitetsverdi), som er en slags samlet pålitelighetsstatus for den enkelte observasjon, før observasjonene distribueres via GTS. I dag benyttes disse flaggene i FoU-kontrollen (se nedenfor), men ikke i MET-divisjonen.

Det vil også foregå en kvalitetskontroll ved FoU og værvarslingsavdelingene med tanke på å utelukke feilaktige observasjoner i den videre bruk av dataene. Denne kontrollen kan utvides til å foreslå korreksjoner i det videre kontrollarbeidet ved Klimaavdelingen.

Det synes fornuftig å utvikle et system for intern distribusjon av datakvalitetsinformasjon til brukere av observasjonsdata ved DNMI.

De såkalte E-data fra plattformene i Nordsjøen og Norskehavet går som rådata og prosesserte data direkte inn til DNMI, men også til eksterne institusjoner, som står for kvalitetskontrollen av dataene. Disse overføres så i ettertid elektronisk til DNMI eller via diskett og post til DNMI. Miljødatasenteret ved Klimaavdelingen har ansvar for å sende kvartalsvise IMMT-data til de meteorologiske værcentralene i Hamburg og Bracknell.



Figur 1.

Forenklet diagram for flyten av observasjonsdata fra observasjonsnettet via elektronisk dataforsendelse og postforsendelser til DNMI. De forskjellige trinn i kvalitetskontrollen er markert med AQC og HQC, betegnelser som er benyttet i NORDKLIM-prosjektet. Det som er merket blått figuren, tenkes redusert til et minimum i fremtiden. Stasjonsbetegnelser er definert i Appendiks B1.

### AQC0, HQC0

Med førstekontroll menes den kontroll som foregår på stasjonen i sann tid. På manuelle værstasjoner illustreres denne kontrollen best ved hjelp av PiO-modellen (PC i observasjonsarbeidet). På observatørens PC er lagt inn et program som gir tilbakemelding når observatøren er i ferd med å gjøre formelle feil eller logiske feil i observasjonsarbeidet. På AVS-stasjoner vil stasjoner som er utstyrt med dobbelt sett av viktige sensorer, og som fungerer uavhengig av hverandre, kunne gi tilbakemelding sammen med observasjonsdataene når noe er galt med måleinstrumentene. It-div/Obs-div rapporterer til Met-div og FoU-div.

### AQC1

Med andrekontroll menes den kontroll som foregår på DNMI ved mottak av observasjonene. Dersom data mangler foretas en sjekk på om dataene fortsatt finnes på stasjonen. Det kan skje overføringsfeil mellom værstasjonen og DNMI. IT-div vil da kunne ta kontakt med observatøren for ny overføring, ev. rapportere feil til Met-div. Forutsatt ingen formelle feil, så kan grove feil i observasjonene oppdages på dette trinnet (se f.eks. Grenseverdikontroll, kap. 2.3) og rapporteres til Met-div, FoU-div, ev. også tilbake til observatøren. Met-div retter opp feil. Feil ved instrumenter og utstyr rapporteres til Obs-div ev. IT-div.

### AQC2 og HQC2

Med tredjekontroll menes den kontroll som foregår på Klimaavdelingen i ettertid, etter at alle relevante observasjoner fra alle typer observasjonsstasjoner er mottatt og komplettert. En hel-automatisk kontroll (AQC) bør helst gå fortløpende på samtlige data fra dagen før. Men fortsatt må det gjennomføres daglige manuelle kontroller (HQC) og månedskontroller (AQC og HQC). Misforståelser angående observasjonsarbeidet meldes direkte tilbake til observatør og/eller ansvarlige for oppfølging av observatørene. Feil ved instrumenter og utstyr rapporteres til Obs-div, ev. IT-div. Feil i observasjoner rapporteres dette til Met-div og FoU-div.

Med FoU-kontrollen menes den kontroll som foretas av værstasjonsdata ved kjøring av numeriske modeller. Når data forkastes pga. feil, rapporteres til Met-div, som følger opp i forhold til observatør, Obs-div og IT-div. Vi tenker oss dette som en QC1-kontroll.

Med VA/VV/VNN-kontrollen menes den kontroll som foretas på værvarslingsavdelingene i forbindelse med kartanalysen og andre analysemetoder i forbindelse med den aktuelle vær-situasjonen. Vi tenker oss dette som en HQC-kontroll, med rapportering som nevnt ovenfor. Maritime data gjennomgår en HQC-kontroll på VV, før de går inn i klimahistorisk datalager.

I figur 1 ovenfor er skissert dataflyten fra observasjonsnettet til DNMI og hvilke kontrollposter dataene må passere underveis. Alle tilbakemeldinger gis så raskt som mulig.

I HQC ligger også bruk av værdata og statistikk, der man kan oppdage feil i data i ettertid. Inspeksjon av værstasjoner kan føre til at systematiske feil i instrumenter og observasjonspraksis oppdages, noe som igjen vil endre datakvaliteten.

## 2.3. Metoder i kvalitetskontrollen

Med kvalitetskontroll av observasjonsdata mener vi i praksis å benytte metoder for å oppdage og rapportere feil eller mistenkelige verdier og metoder for å korrigere eventuelle feil. Vi skisserer her hva slags type kontroller vi tenker oss i fremtiden og hvilke verktøy som kan være aktuelle å bruke (se figur 2). Vi kan ikke nå ta stilling til i hvilken grad kontrollene bør utføres manuelt eller automatisk. Det er imidlertid et overordnet mål å automatisere kontrollene så langt det er praktisk mulig og faglig forsvarlig.

Det er hensiktsmessig å benytte forskjellige former for kvalitetskontroll. Vi skisserer her noen av de som er aktuelle å bruke.

#### *Format-kontroll. Formell kontroll.*

Det kontrolleres at Synop og Metar har den nødvendige identifikasjon (stasjonsnummer, tidsstempling, osv), at de forskjellige gruppene har riktig antall karakterer, at det er samsvar mellom dagbok og met, m.m. Det kontrolleres også mot et informasjonsarkiv, slik at det er samsvar mellom den faktiske observasjonen og observasjonsprogrammet for stasjonen. Ved format-feil er det en stor risiko for at observasjonen forkastes. Feil rapporteres tilbake til avsender. Se også "Logisk kontroll".

#### *Datadekningskontroll.*

Det foretas en kontinuerlig registrering av datadekningen for stasjonen. Denne gjør det mulig å avgjøre om beregninger av døgn-, uke- eller månedsverdier skal foretas på vanlig eller alternativ måte. Når mangler oppdages, tas kontakt med stasjonen for å sjekke om data fortsatt er tilgjengelige.

#### *Grenseverdikontroll.*

For hver parameter settes grenseverdier (definert i egen tabell), som parameteren må holde seg innenfor. I motsatt fall er parameterverdien feil og må enten forkastes eller korrigeres.

#### *Ekstremverdikontroll.*

Minimums- og maksimumsverdier av lufttrykk, temperatur-, vindhastighet-, nedbørhøyde og snødybde sjekkes mot tidligere registrerte minimums-/maksimums-verdier (lagret i egen tabell). Falske verdier lukes ut på et så tidlig tidspunkt som mulig, om nødvendig etter kontakt med observatøren. Mistenkelige verdier flagges.

#### *Tidsutviklingskontroll. Sprangkontroll, konstantverdi-kontroll.*

For relevante parametre settes maksimalverdier for endring av parameterverdien mellom forrige og inneværende observasjon, ev. trekkes også flere foregående observasjoner, ev. også flere påfølgende observasjoner av samme parameter inn i vurderingen. Det settes også en grense for hvor mange påfølgende observasjoner de enkelte parametre kan ha samme verdi. Når kontrollen slår ut vil noen parametre bli rettet automatisk etter bestemte kriterier, ev. forkastes, eller det må foretas en faglig vurdering om parameterverdien skal godkjennes eller korrigeres manuelt.

#### *Logisk kontroll.*

En type logisk kontroll er rent formell. FF kan ikke være større enn FX, som igjen ikke kan være større enn FG. Tilsvarende forhold finnes også for en del andre parametre, f. eks. når det gjelder: PO/P, TN/TT/TX, N/Nh, værsymboler/ww/W<sub>1</sub>W<sub>2</sub>/W<sub>d</sub>, N<sub>h</sub>/h/C<sub>L</sub>C<sub>M</sub>C<sub>H</sub>. En annen type logisk kontroll går mer på konsistens og helhet i observasjonen. Hvis det er nedbørhøyde, må det også være nedbørsymboler og passende værkoder, og omvendt. Hvis det er bygevær, bør det være skytyper som indikerer bygevær. Hvis det er tåke, må det være sikt under 1000 m, eller luftfuktigheten kan indikere tåkedis, ev. tåke på avstand, osv. Alle feil som oppdages, blir rettet. Utvikling av hensiktsmessige algoritmer kan føre til en hel eller delvis automatisering av denne type kontroll. F.eks. kan signalforstyrrelser på automatstasjoner være årsak til slike feil, jf. vindmålingene fra Slettnes fyr.

#### *Romkontroll.*

Den aktuelle stasjonens observasjoner eller måleverdier ved et observasjonstidspunkt sammenliknes med tilsvarende observasjoner fra nærliggende representative stasjoner eller med en annen forventet verdi. Eller man kan sammenlikne tidsutviklingen i måleverdiene, ev. også hensiktsmessige gjennomsnittsverdier/-summer. Hjelpemidler som kan bli viktige ved retting av feil i romkontrollen, er numerisk analyse, grafiske verktøy (f.eks. GIS) og forskjellige metoder for interpolasjon.

#### *Statistisk kontroll. Klimakontroll.*

Når måneden er omme lages statistikk over utvalgte parametre, som vurderes mht. statistisk fordeling (sammenlikning med forventet fordeling ut fra historiske data). For AVS er det særlig viktig å sjekke at sensorene måler på riktig nivå (f.eks. at vindretningssensor er rett orientert). Men også en vurdering av den statistiske fordeling av parameterverdier sammenliknet med representative nabostasjoners, er interessante i kvalitetskontrollen (klimakontroll). En slik kvalitetskontroll vil også være et godt hjelpemiddel i veiledningen av observatørene og i arbeidet med å følge opp instrumentkvaliteten på stasjonen.



### *Produktkontroll.*

Når data benyttes til statistikk, oppdrag eller klimautredninger foretas en kritisk vurdering av datakvaliteten, for å kvalitetssikre klimaproduktene. For enkelte produkter kan denne gjøre automatisk. Tilbakemelding gis til de ansvarlige for datakvaliteten.

En stor del av kontrollene vil basere seg på stasjonens egne observasjoner og noteringer, som det fremgår av oversikten ovenfor. Romkontrollen, klimakontrollen og delvis annen statistisk kontroll er imidlertid avhengig av observasjoner fra nabostasjonene og ev. andre stasjoner, som har representative observasjoner for det aktuelle vørelementet som kontrolleres. Vi ser allerede at det i områder med dårlig datadekning er nødvendig å beregne forventede verdier (f.eks. modelldata), som observasjonene kan sammenliknes mot.

Det er viktig at alle stasjonstyper, som observerer det aktuelle vørelementet, trekkes inn i romkontrollen. F.eks. bør forskjellige typer automatstasjoner og manuelle stasjoner benyttes i temperaturkontrollen. I kontroll av nedbørsummer bør, i tillegg til N-stasjonene, også D-stasjoner og P-stasjoner trekkes inn, sammen med de andre stasjonene nevnt ovenfor.

En grunnleggende type romkontroll baseres på observasjonene kl. 00, 06/07, 12 og 18 UTC. Dette er de observasjonstidspunkter (med unntak for nattobsen) som alle manuelle vær-stasjoner har, og det er tidspunkter for modellkjøringene, basert på radiosondedata. Mellomliggende observasjoner, per time eller 3. time, kan neppe kontrolleres effektivt i denne type romkontroll og bør harmoniseres på annen måte. Romkontroll fire ganger per døgn burde være tilstrekkelig til å oppdage systematiske feil i observasjonene.

I romkontrollen er det viktig å utnytte alle tilgjengelige data fra radar, satellitt og fly. Det må derfor legges til rette for at slike data kan benyttes sammen med modelldata og tradisjonelle bakkedata.

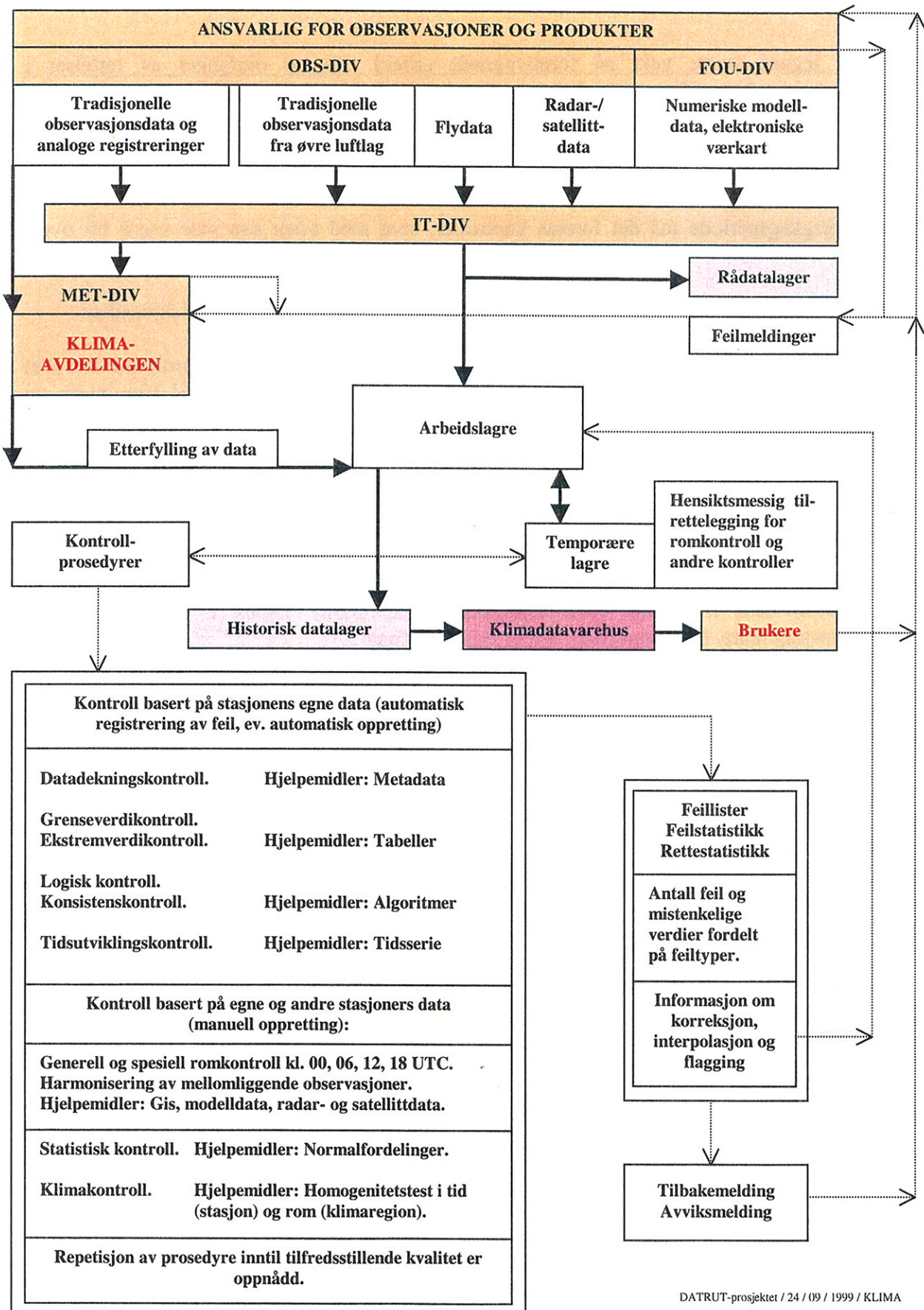
I ettertid vil statistiske kontroller av hver enkeltstasjons månedsdata og en sammenlikning av siste måneds data med tilsvarende data bakover i tid (homogenitetstest) og med tilsvarende data for representative nabostasjoner, være et hjelpemiddel for å oppdage systematiske feil i dataene.

I all romkontroll vil GIS være et naturlig hjelpemiddel.

I arbeidet med å videreutvikle kvalitetskontrollen av vørelementene er det viktig å finne frem til de mest effektive kontrollene på et så tidlig stadium som mulig. Det må vurderes å benytte en kombinasjon av forskjellige typer kontroller og det må legges mye arbeid i utprøving av automatiserte kontrollsystemer for å redusere det manuelle arbeidet til et minimum. I dette arbeidet må det vurderes å ta i bruk moderne grafiske verktøy og interpolasjon i numeriske finskala-modeller.

Man må også utnytte maksimalt den sanntidskontroll som utføres i IT-, FoU- og MET-div (VA/VV/VNN).

Når feil oppdages må tilbakemelding gis til de som er ansvarlig for observasjonene. Alvorlige feil bør så langt det er mulig, meldes til brukere av observasjonsdata.



Figur 2. Forenklet diagram for dataflyt inn mot klimadatabasen og generelle kontrollprosedyrer. Dataflyt er markert med heltrukken pil.

Det må legges større vekt på forebyggende arbeid, slik at omfanget av rettelser i kontrollarbeidet kan reduseres maksimalt. For å oppnå dette bør så mange kontroller som mulig, plasseres så nær produksjonsnivå som mulig. Manuelle stasjoner kontrolleres gjennom PiO-systemet, automatstasjoner ved dublering av sensorer, osv. Flest mulige kontroller foretas i sann tid eller nær sann tid, slik at nødvendige tilbakemeldinger kan gis så raskt som mulig.

I en overgangsperiode må det foretas kontroller, som med tiden kan vise seg å bli overflødige.

I figur 2 ovenfor er skissert dataflyt og generelle kontrollprosedyrer for observasjonsdata.

Målet for databehandlingen av observasjonsdataene er å lagre kvalitetskontrollerte data i et permanent historisk lager, HL (for intern bruk), og i et datavarehus, DVH (for intern og ekstern bruk), som er tilpasset en hurtig og effektiv uthenting av data.

## 2.4. Databasestruktur

Vi tenker oss minst to databaser, en klimadatabase (KLIBAS) og et klimadatavarehus (DVH). Disse vil inneholde tradisjonelle observasjonsdata (bakkedata og øvre atmosfæredata). Andre datatyper, som f.eks. radar- og satellittbilder kan lagres i databaser for romlig informasjon, men rutenett-data fra slike observasjonskilder må kunne lagres i KLIBAS på en hensiktsmessig måte. Databasestrukturen er skissert i figur 3.

### *2.4.1. Klimadatabasen*

KLIBAS vil bestå av metadata-tabeller, arbeidslager-tabeller, temporære tabeller, statistikk-tabeller og hovedlagertabeller, ev. også permanente hjelpetabeller for bearbeiding av data.

Metadata-tabellene vil inneholde historisk stasjonsinformasjon fra stasjonen ble opprettet. Det kan være opplysninger om stasjonsnummer, stasjonsnavn, stasjonstype, geografisk beliggenhet, opprettelse og ev. flyttinger, navn på observatør, observasjonsprogram og annen historikk.

Arbeidslager-tabellene (AL) inneholder til å begynne med ukontrollerte eller ikke-ferdigkontrollerte data, som forblir i tabellene inntil de er ferdigkontrollert.

Det er nødvendig å ha arbeidslagre som kan håndtere alle typer observasjonsdata. I dag har vi stasjonstyper som gir observasjoner med følgende hyppigheter per døgn: 1, 2, 3, 4, 5, 6, 7, 8, 12, 24, 48, 72, 144 og 1440.

Temporære tabeller bygges opp med tanke på romkontroller, tidsseriekontroller eller andre kontroller som ikke kan benytte arbeidslagrene direkte. Tabellene vil eksistere så lenge kontrolloppgavene blir utført. De må kunne inneholde både direkte observerte parametre og sammensatte eller avledete parametre for et utvalg observasjonstidspunkter. Hovedtyngden av data vil gjelde tidspunktene 00, 06, 12 og 18 UTC, men også timesdata (f.eks. med tanke på kontroll av P-data), døgn- og månedsdata kan være aktuelle å bruke.

Statistikktabeller vil inneholde statistikk som ikke kan reberegnes digitalt. Det betyr f.eks. at historiske månedsverdier, som tidligere ble manuelt beregnet, kan legges inn direkte i månedstabeller (blåbøker, sekeldata, m.m.).

Hovedlagertabellene, eller historiske datalager (HL), vil inneholde den endelige lagringen av data. Dataene blir lagret etter stasjonstype. For å kunne garantere et konsistent datalager og et effektivt databasearbeid må værstasjonsdata lagres og rettes kun ett sted i HL. Det betyr at det ikke skal forekomme duplisering av ferdig bearbeidede data. For å kunne identifisere en observasjon entydig, må en bruke en indeks sammensatt av sted (stasjonsnummer), tidspunkt og stasjonstype (nøkkelverdier). En fullstendig observasjon av én datatype blir lagret som én rad i tabellene, med entydig nøkkel.

Fordi det vil være behov for å benytte helt ferske data i forskjellige former for statistikk, må vi tillate at ferske data ikke er ferdigkontrollert, med de følger dette kan få for datakvaliteten. Det vil således hele tiden foregå en retting og justering av data i HL. Dette vil også få som konsekvens at DVH daglig ajourføres, med nye enkeltdata og bearbeidede data (fortløpende statistikk).

Fly-data bør lagres i KLIBAS i et veldefinert rutenett, eller for utvalgte geografiske punkter, i definerte høydenivå eller trykkflater oppover i atmosfæren. Slike data blir ikke lagret på Klimaavdelingen i dag.

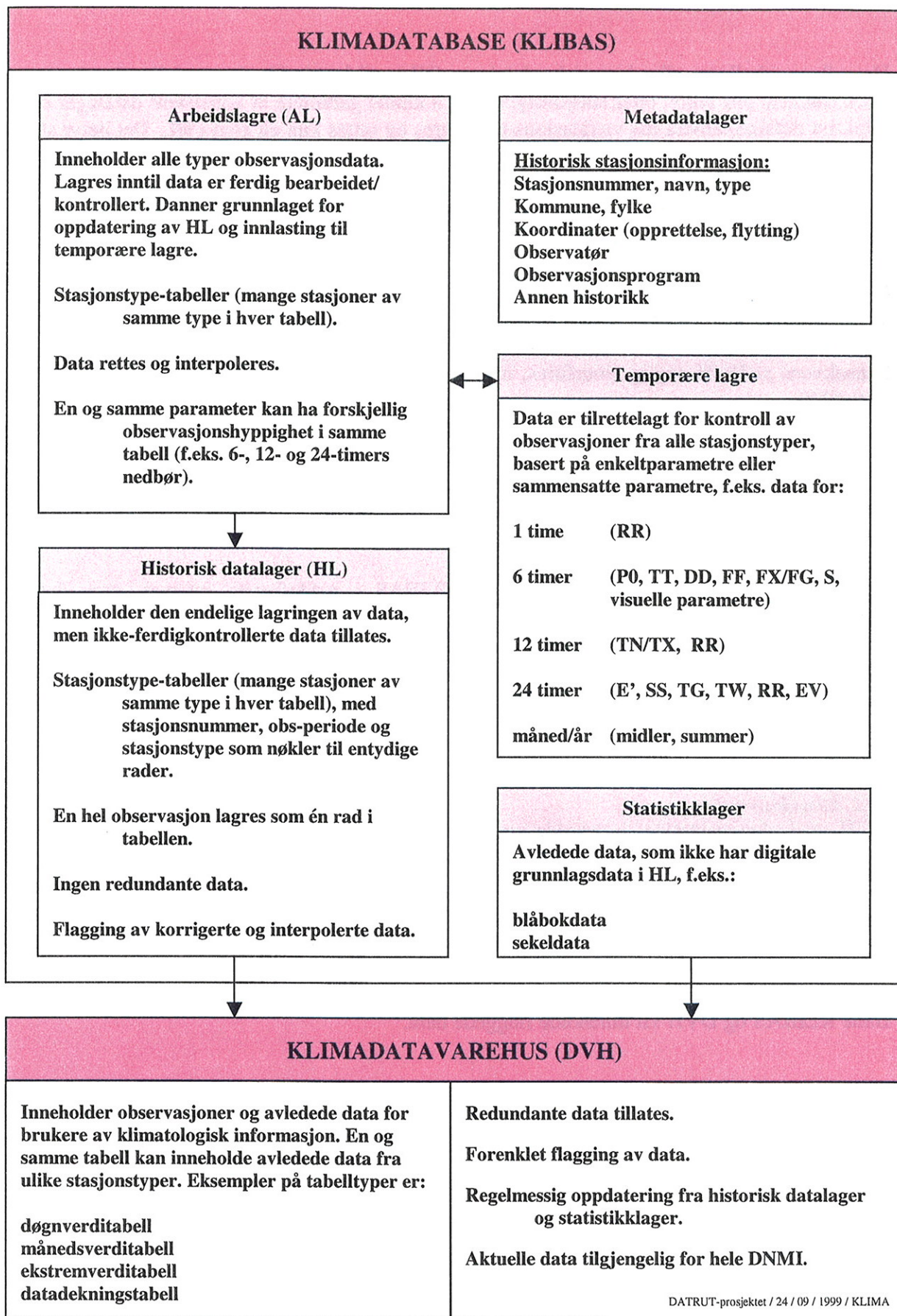
Radar- og satellittdata bør også lagres i KLIBAS i veldefinerte rutenett, som dekker hovedlandet, Svalbard og Jan Mayen og kystnære deler av havområdene utenfor Norge. Slike data blir ikke lagret på Klimaavdelingen i dag.

Modelldata bør lagres i en egen database. Denne kan inneholde både være verdier i rutenett og verdier i definerte stasjonspunkter.

#### **2.4.2. Klimadatavarehus**

DVH vil i utgangspunktet inneholde godkjente observerte og avledede data. Men det er et krav fra brukerne at også ferske data skal være tilgjengelige uten nevneverdig forsinkelse. Det må derfor tillates at ikke-ferdigkontrollerte data blir liggende i DVH inntil de kan erstattes av ev. korrigerte data fra HL. I DVH kan de samme data være lagret flere steder, som avledede data. Det foretas en regelmessig oppdatering fra historisk lager.

Både KLIBAS og DVH vil inneholde flaggede data.



Figur 3.  
Skisse av Klimaavdelingens fremtidige databasestruktur.

## 2.5. Manglende data

Klimaavdelingen må ta stilling til i hvilket omfang manglende data skal interpoleres. Dette innebærer problemstillinger av typen: Skal hele observasjonen på en stasjon interpoleres eller bare de viktigste parametrene? Skal hovedobser og mellomobser interpoleres i samme omfang? Skal det foretas interpolasjoner når stasjonen er ute av drift flere dager i trekk? I hvilken grad skal interpolasjon foregå automatisk?

DATRUT-prosjektet er foreløpig av den oppfatning at vi ikke kommer utenom interpolasjoner. I alle fall må vi fortsatt ha rutiner som gjør det mulig å beregne månedsverdier av de viktigste meteorologiske parametre, uten at det nødvendigvis må foretas interpolasjoner for en lang rekke sammenhengende mangler.

Vi håper å kunne komme nærmere tilbake til dette i neste rapport.

## 2.6. Flagging

Når kvalitetskontroll fører til endring eller justering av parameterverdi ut fra en skjønnsmessig vurdering eller ut fra en objektiv metodikk, må data merkes (flagges), slik at brukeren av dataene er kjent med dette. Brukere av data må også vite når data ikke har gjennomgått noen kvalitetskontroll.

Når meldinger om systematiske feil i dataene blir rapportert i ettertid, må data merkes på en slik måte at brukeren blir kjent med dette, inntil dataene blir korrigert. En kan f.eks. tenke seg at programmer som henter ut data, har mulighet for å sjekke status for aktuell stasjon, slik at generell melding kan gis.

Klimaavdelingen må også ta stilling til omfanget av dataflagging og i hvilken grad data skal korrigeres. Hvor detaljert skal flaggingen være? Skal observasjoner som er foretatt utenfor godkjent observasjonstid, bare flagges - eller korrigeres også?

Vi håper å kunne komme nærmere tilbake til dette i neste rapport.

## 2.7. Sikring av originaldata

Frem til nå har dagbøker, ukekort og annet rådatamateriale blitt arkivert direkte eller ev. mikrofotografert og arkivert. Når etter hvert manuell bearbeiding av data erstattes av automatisk elektronisk databehandling, så reiser spørsmålet seg om dagbøker, ukekort, etc. skal erstattes av et elektronisk rådatalager av originale observasjonsdata. Et slikt lager må ha et tilstrekkelig sikkerhetsnivå for historisk arkivering.

Dagens praksis er slik at på alle manuelle værstasjoner føres de manuelle observasjonene inn i dagbok for innsending til DNMI, og på nedbørstasjoner føres ukekort. På HS-stasjonene føres alle manuelle/visuelle observasjoner i dagboka, instrumentelle målinger fra automatstasjonen føres ikke. Denne praksis må revurderes i tråd med de anbefalinger DATRUT-prosjektet vil komme med senere.

## Appendiks



**NOTAT  
til  
Klimaavdelingen**

**Prosjekt Databehandlingsrutiner (DATRUT)**

Jeg viser til notat datert fra Lars Andresen.

Det opprettes et avdelingsinternt prosjekt med følgende mandat:

1. Skissere en fremtidig, helhetlig behandlingsrutine for alle typer observasjonsdata som Klimaavdelingen har ansvaret for lagring av. Skissen skal definere sammenhengen mellom dataflyt, datakontroll og databasestruktur.
2. Iverksette "straks"-tiltak for å effektivisere og heve kvaliteten på nåværende rutinebehandling av data fra manuelle værstasjoner.
3. Gjennomgå dagens databehandlingsrutine for de bemannede værstasjonene med tanke på både kvalitetshevning og effektivisering. Gjennomgangen skal gi:
  - a) En vurdering av hvordan moderne hjelpemidler/verktøy kan forenkle kvalitetskontrollen.
  - b) En oversikt over samtlige algoritmer i nåværende kvalitetskontroll og forslag til nye algoritmer/kontroller.
  - c) En vurdering av hvilke deler av rutinen som kan automatiseres og hvilke deler som fortsatt må være manuelle.
  - d) Forslag om en gradvis omlegging av nåværende rutine og en beskrivelse av alle ledd i ny en rutine, inkludert ressursbehov og ansvarsforhold.

Forslag til endringer skal tilpasses eventuelle beslutninger som er tatt på grunnlag av skissen som er utarbeidet under punkt 1.

4. Rapportere skissen i punkt 1 innen 15. april og forslagene i punkt 3 innen 1. oktober 1999. Etter dette tidspunkt gjøres en vurdering av veien videre. "Straks"-tiltak settes i gang etter hvert.

Som deltagere i prosjektet oppnevnes:

Lars Andresen (leder), Per Ove Kjensli, Åse Moen Vidal, Zbigniew Toporowski

Prosjektet skal kunne innhente assistanse fra andre på avdelingen. Hvis assistansen krever vesentlig arbeid skal tillatelse innhentes fra fagsjef før det igangsettes. Tillatelse skal også innhentes før det eventuelt settes i gang "straks"-tiltak som medfører vesentlige endringer i arbeidsrutiner.

  
Bjørn Aune



## Appendiks B. Definisjoner

### B1. Observasjonsstasjoner og data

Med manuelle stasjoner menes stasjoner som er bemannet med personell som observerer visuelt eller foretar manuelle avlesninger av måleinstrumenter og fører disse inn i en dagbok eller et elektronisk medium.

Med automatiske værstasjoner menes stasjoner, der instrumenter foretar målinger og lagrer disse automatisk, uten medvirkning av personell. Radiosonde stasjoner regnes som automatiske stasjoner. De kan riktig nok være bemannet for at ballonger skal kunne bli klargjort for slipp, men målinger og datalagringsutstyr for øvrig er helt automatisert.

Delvis automatiske, eller hybride værstasjoner, er stasjoner, der observasjoner av skyer, synsvidde, generelt vær og sjøgang, og i et relativt stort omfang også av vind og nedbør, må utføres manuelt. Såkalte "instrumentelle" observasjoner foretas av det automatiske registreringsutstyret på stasjonen. På grunn av manglende bemanning av tradisjonelle værstasjoner om natten, er særlig viktige stasjoner delvis automatisert.

VS-stasjon:	Manuell værstasjon, der data videresendes som Synop.
MS-stasjon:	Manuell værstasjon til havs (på plattform, rigg eller skip), der data videresendes som Ship.
VK-stasjon:	Manuell værstasjon for klimatologiske formål på land.
HS-stasjon:	Delvis automatisert værstasjon, enten på land eller til havs. Manuelle observasjoner av "visuelle" parametre og automatiske observasjoner av instrumentelle parametre.
PiO-stasjon:	Manuelle værstasjoner med PC i observasjonstjenesten. Det er utarbeidet programmer som avslører logiske feil og peker på mistenkelige observasjonsverdier. Observatøren får mulighet til å rette opp feil i observasjonen før den blir sendt til DNMI.
AERO-stasjon:	Manuell aeronautisk stasjon eller meteorologisk stasjon for luftfartsformål, der data videresendes som Metar.
N-stasjon:	Manuell nedbørstasjon for klimatologiske formål.
D-stasjon:	Manuell fordampningsstasjon for klimatologiske formål.
S-stasjon:	Manuell eller automatisk solstrålingsstasjon. For noen stasjoner må solskinnstid registreres manuelt.
AVS-stasjon:	Automatisk værstasjon for værvarslingsformål og klimatologiske formål, på land eller til havs (fellesbetegnelse).
AS-stasjon:	Automatisk værstasjon, der data videresendes som Synop.
AM-stasjon:	Automatisk maritim stasjon, hovedsakelig på plattformer (tilsvarende AS-data. I tillegg E-data (Environmental data) for klimatologiske formål.
AL-stasjon:	Automatisk landbruksstasjon for klimatologiske formål.
AO-stasjoner:	Automatisk værstasjon, som benyttes for oppdrag til forskjellige klimatologiske formål.
P-stasjon:	Automatisk nedbørintensitetsstasjon (Plumatic) for klimatologiske formål.
RS-stasjon:	Radiosonde stasjon for værvarslingsformål og klimatologiske formål.
ECSN-stasjoner:	Stasjoner som er med i European Climate Support Network.
Climat:	Klimatologiske månedsdata for utvalgte bakkestasjoner.
IMMT-data:	Utvekslingsformat for kvalitetskontrollerte maritime data (International Maritime Meteorological Tape).
FLY-data:	Måleinstrumenter montert på flykroppen. Flydata er foreløpig ikke tilrettelagt for lagring og kvalitetskontroll.
Radar-observasjoner:	Radardata er foreløpig ikke tilrettelagt for lagring og kvalitetskontroll.
Satellitt-observasjoner:	Satellittdata er foreløpig ikke tilrettelagt for lagring og kvalitetskontroll.

## B2. Observasjoner, værsymboler og værkoder

For å forenkle skrivearbeidet kan det være nyttig å sette navn på de ulike observasjonene og forklare deler av en manuell observasjon.

Med hovedobs menes observasjonen som er gjort til en av hovedobservasjons-terminene: 00, 06, 12, 18 UTC. På manuelle stasjoner utføres observasjonen 5-20 minutter før hel time, men helst så nær opp til hel time som mulig. På automatstasjoner foretas observasjonen hver hele time.

Med mellomobs menes observasjonen som er gjort til en av mellomobservasjons-terminene: 03, 09, 15, 21 UTC. Utførelsen er som beskrevet under hovedobs.

Med timesobser menes fortløpende timevise observasjoner for hele eller deler av døgnet. Utførelsen er som beskrevet under hovedobs.

Med 12-timers obs menes en observasjon som utføres på samme klokkeslett hver dag, to ganger i døgnet. På MVS-stasjoner måles f.eks. nedbørhøyde, maksimums- og minimumstemperatur kl. 06 og 18 UTC.

Med døgnoobs menes en observasjon som utføres på samme klokkeslett hver dag. På N-stasjoner måles nedbørhøyden, snødybden og snødekket hver morgen kl. 08 hele året gjennom, enten det er normaltid (07 UTC) eller sommertid (06 UTC). På MVS-stasjoner måles f.eks. gressminimumstemperatur, snødybde og markas tilstand kl. 06 UTC, mens sjøtemperatur måles kl. 12 UTC.

Med værsymboler menes de symbolene som er beskrevet på dagbokas forside på MVS-stasjoner. Regn, yr, snø og sludd skal f.eks. angis med hhv. en prikk (•), et komma (°), en stjerne (\*) og en stjerne med prikk over. Det er i alt 29 forskjellige symboler (hvorav 13 forskjellige nedbørslag, dugg, rim og tåkerim ikke inkludert). På VK-stasjoner benyttes kun værsymboler, ikke værkoder (se nedenfor). På N-stasjoner brukes bare et utvalg av værsymbolene nevnt ovenfor.

Med værkoder menes definerte kodetall for ww (været ved observasjonstiden),  $W_1 W_2$  (været siden forrige hovedobs, maks 6 timer) og  $W_d$  (tilleggsinformasjon om været siden forrige hovedobs, inntil 12 timer). På VS-stasjoner benyttes både værsymboler (se ovenfor) og værkoder.