Data assimilation for lake models applied in numerical weather prediction – some initial naive ideas

Nils Gustafsson, SMHI

Overview of talk

1. Related work in the past – what did we learn?
2. Model variables to be assimilated, example
3. Observations
4. Tasks for lake data assimilation
5. Coupling with the atmospheric assimilation?
6. Assimilation algorithms?
7. Initial lake assimilation for HIRLAM and HARMONIE

Related work in the past (1)

Ljungemyr P, Gustafsson N, Omstedt A, 1996:

Parameterization of Lake thermodynamics in a high resolution weather-forecasting model. Tellus 48A

- **Simple lake model, well-mixed for shallow lakes, PROBE with turbulence parameterization for deep lakes**
- **Model spin-up is important**
- **Some errors remain – potential for data assimilation**
- **Atmospheric model not yet developed enough to profit from lake model (1996)**
- **The computer code was too complex.**

Related work in the past (2)

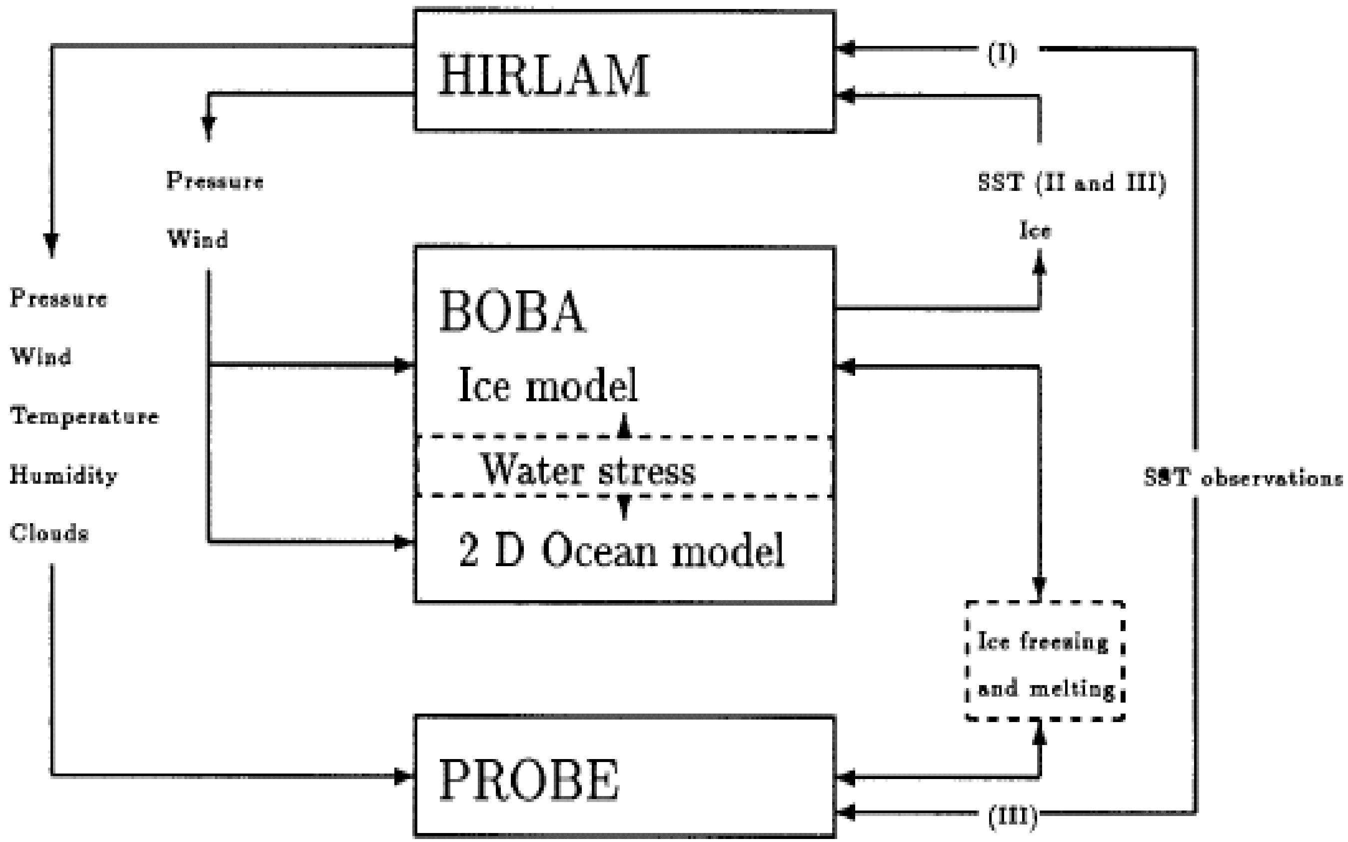
Gustafsson, N., Nyberg, L. and A. Omstedt (1998):

Coupling high-resolution atmosphere and ocean models for the Baltic Sea. Mon. Wea. Rev. ,126

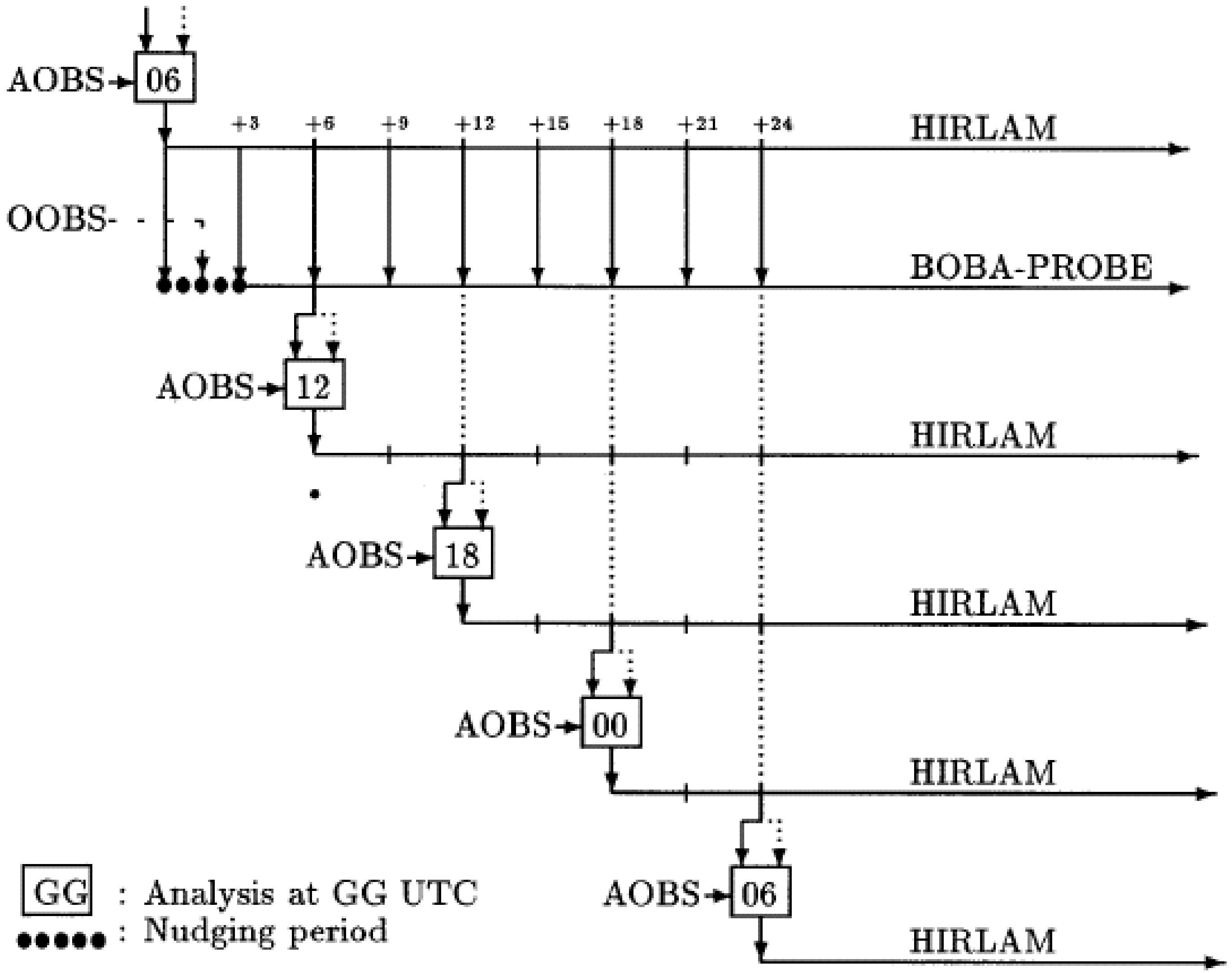
Coupling between HIRLAM, BOBA (2D ice forecasting system) and PROBE (applied to different thermodynamic regions) in the Baltic Sea area.

- **OI data assimilation in HIRLAM – atmosphere only**
- **Assimilation of SST:s in PROBE (Nudging of Sensible Heat Flux) - no assimilation of ice information**
- **Coupling in 6h assimilation cycles**

Coupling HIRLAM – BOBA - PROBE



Coupling in assimilation cycles



Thermodyn. regions for PROBE

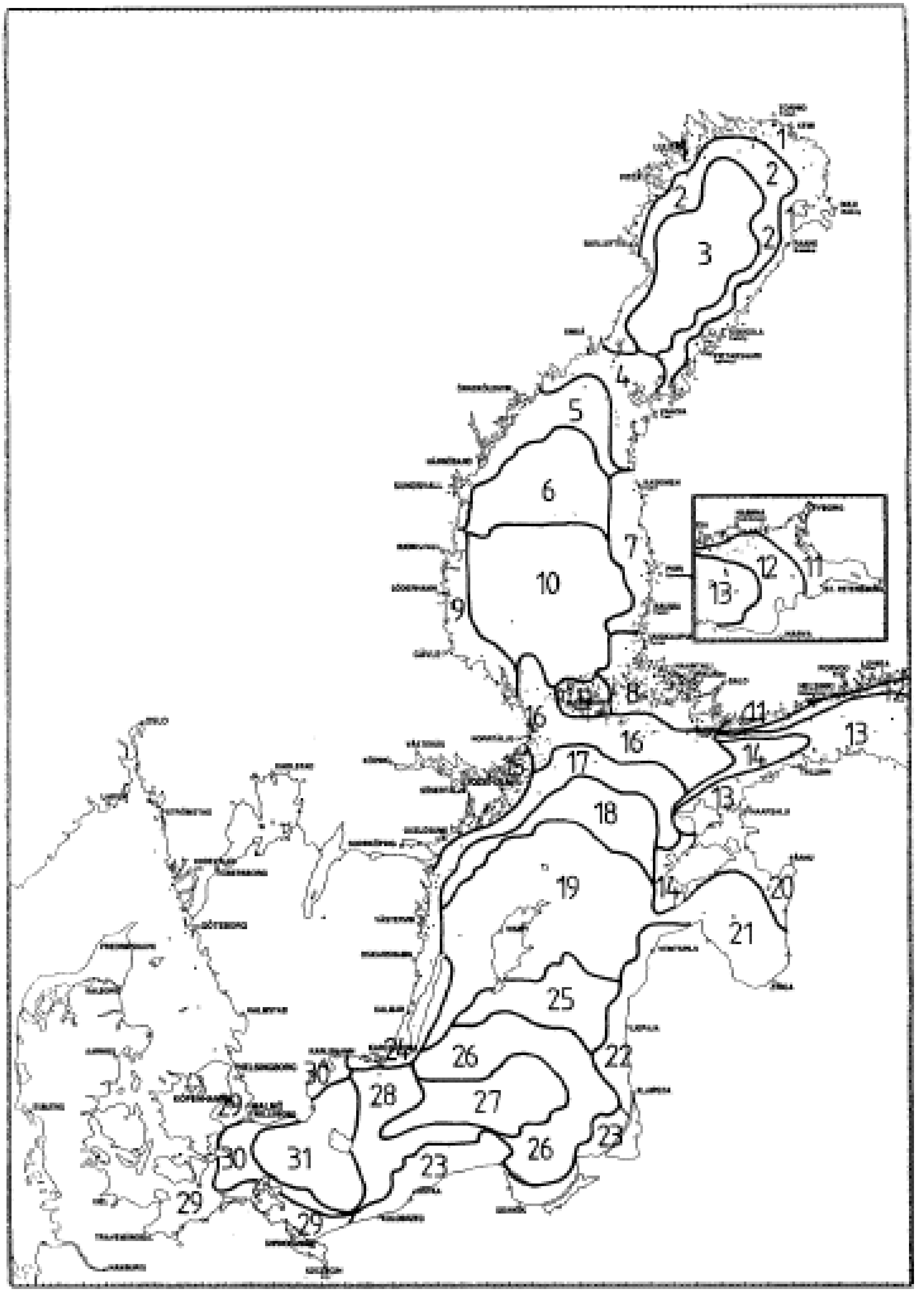


FIG. 7. The 31 thermodynamic regions of the BOBA–PROBE ocean model system.

Filtering of observation increments:

$$T_{\mathrm{sea}}^{\mathrm{AN}}(k) = T_{\mathrm{sea}}^{\mathrm{FC}}(k) + \sum_{l=1}^{N} W_{kl}\left(T_{\mathrm{sea}}^{\mathrm{OBS}}(l) - T_{\mathrm{sea}}^{\mathrm{FC}}(l)\right), \quad (1)$$

Calculation of an efficient sensible heat flux used for nudging:

$$F_s^{\mathrm{AN}} = \frac{\Delta d}{\Delta t} C_{pw} \rho_w \left(T_{\mathrm{sea}}^{\mathrm{AN}} - T_{\mathrm{sea}}^{\mathrm{FC}}\right), \quad (2)$$

January 1987

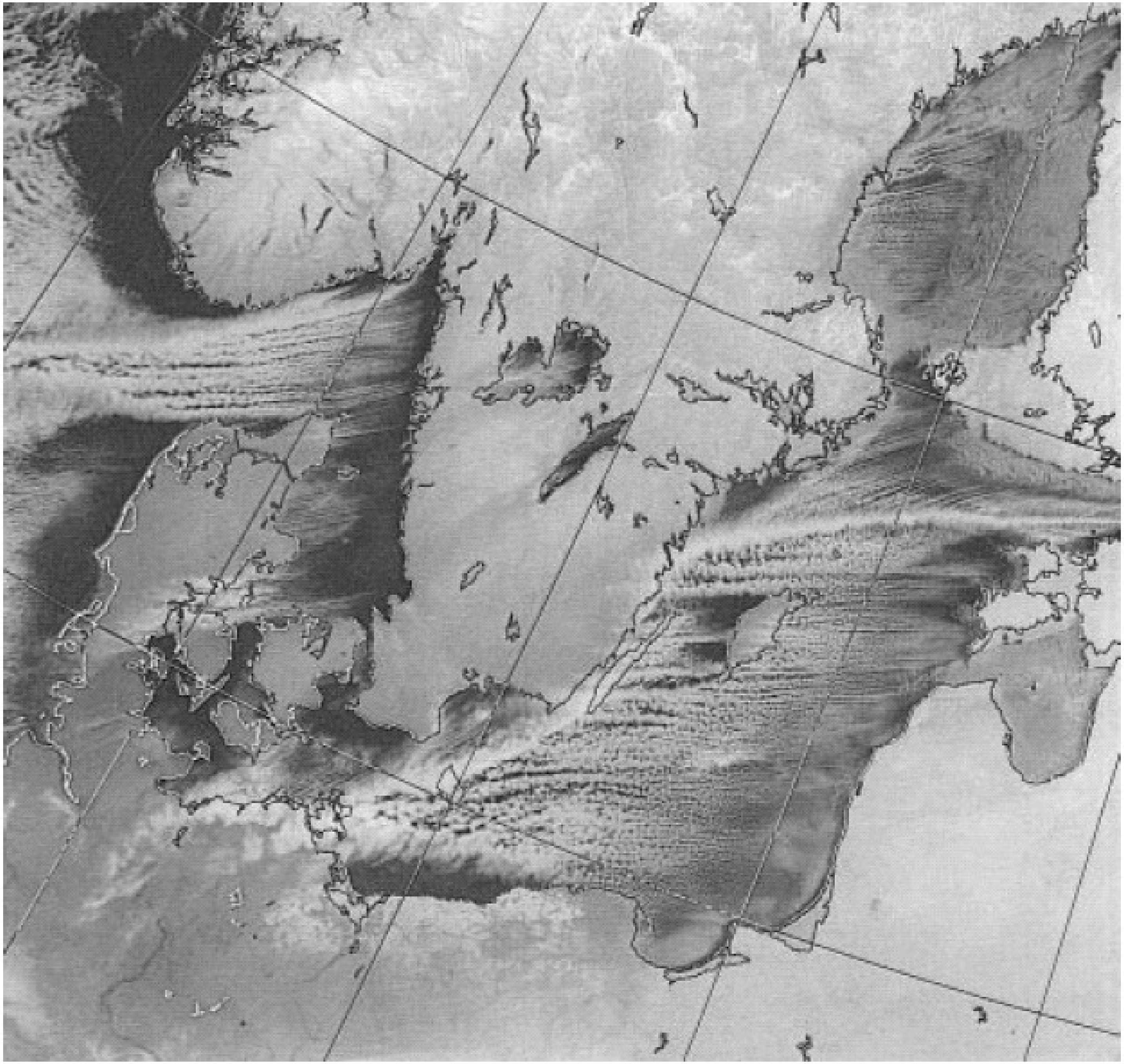
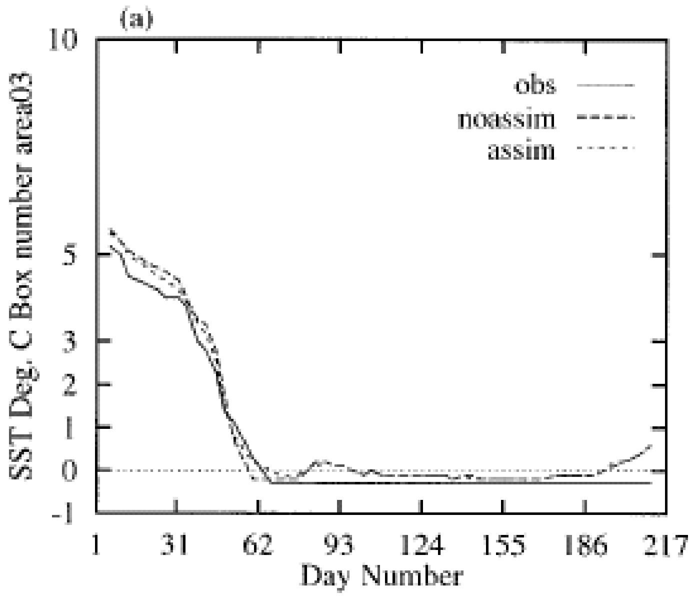
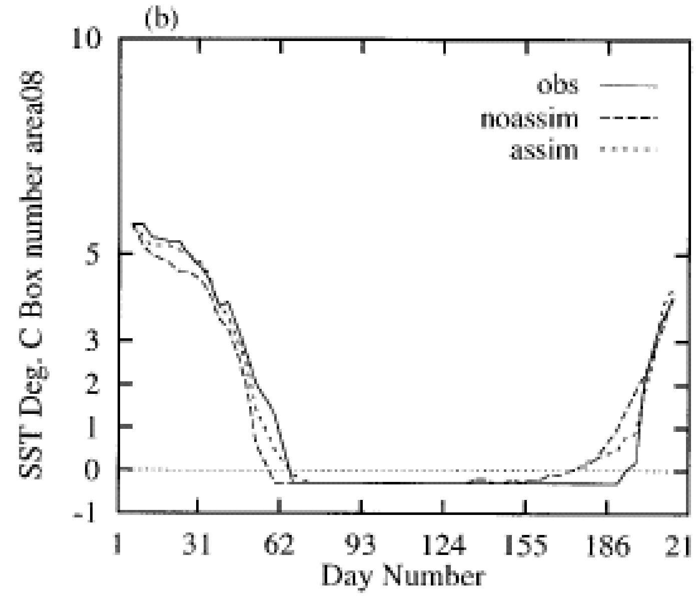


FIG. 3. Infrared *NOAA-9* satellite image 11 January 1987 1235 UTC.

Bay of Botnia



Åland Archepelago



Effect of assimilation for selected regions

Eastern Baltic Proper

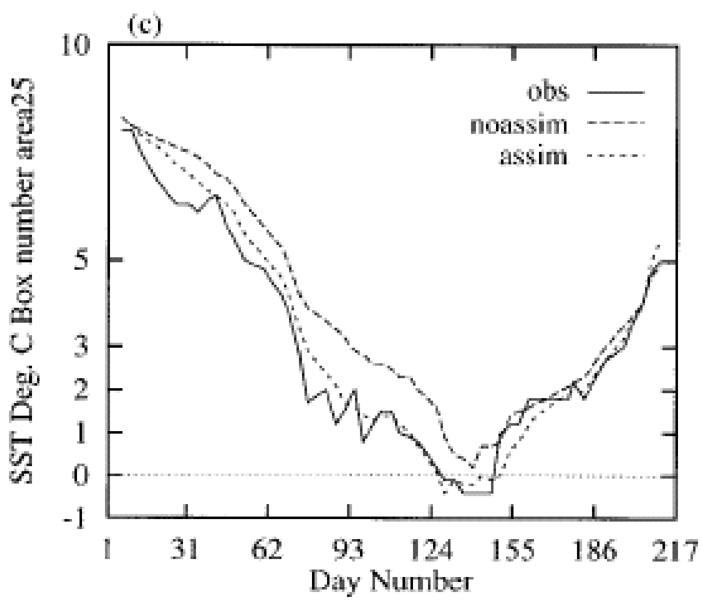


FIG. 9. SST variations in three PROBE subareas during the period November 1986–May 1987. Observed values (full line), simulated values without data assimilation (dashed line), and simulated values with data assimilation (dotted line). (a) Central Bothnian Bay (region 3), (b) the Åland archipelago (region 8), and (c) eastern Baltic proper (region 25).

Conclusions, regarding data assimilation, from these 2 works:

- **Data assimilation is needed!**
- **We must be prepared to use assimilation techniques that modifies model variables that are not directly observed!**

The three tasks of data assimilation (Tony Hollingsworth):

- Filtering
- Interpolation
- Balancing

Lake model variables to be assimilated, Flake as an example:

- **Mixed layer temperature**
- **Bottom temperature**
- **Mixed layer depth**
- **Ice cover (depth?)**
- **Snow on the ice?**

We would need forecast error characteristics (the B matrix):

- **Error standard deviations**
- **Correlations between errors of different variables**
- **Biases?**

Can these error characteristics be estimated by model simulations?

Observations related to lakes

- **In situ Lake surface water temperatures (sparse)**
- **Lake surface (skin?) water temperatures derived from satellite radiance data**

We need observation error characteristics:

- **Error variances**
- **Error biases**
- **Error correlations (in the horizontal)**
- **Representativity (with regard to lake depth, for example)**

Tasks of the Lake model data assimilation

- Quality Control of the Lake observations (reject poor and non-representative data)
- Bias reduction
- Horizontal spread of the observed information, for example taking the lake depth into account
- Spread of the observed information to other model variables

Most difficult – particular problems

- **Long, narrow and deep lakes (Norway)**
- **Lakes controlled by fresh-water inflow**
- **Small lakes, not seen by satellites**

Complete coupling with the atmospheric data assimilation?

- Desirable to avoid initial adjustment oscillations
- Difficult at the present stage of development – balance and smoothness constraints in the atmospheric data assimilation are not suitable for the inhomogenious and anisotropic surface conditions
- Future possibility: Externalized observation operator for the surface/soil/lake assimilation schemes but joint cost function to be minimized
- At present: Separate soil/surface/lake assimilation scheme, possibly applied iteratively within outer 4D-Var loops. Possibility: Minimize shock by using the same set of observations in all schemes?

Lake data assimilation to be applied in NWP as soon as possible (within one year)

- **Apply spatial interpolation of surface (mixed layer) water temperatures taking lake quaracteristics into account (lake elevation, lake depth, …)**
- **Apply a 1D adjustment procedure to estimate model variables that are not observed, taking (in some way) the model dynamics into account) – keep the atmosphere fixed during this adjustment.**

Data assimilation algorithms, suitable for lake assimilation

The spatial spread of information ("2D" problem), mainly surface (mixed layer) temperature:

- **Successive Corrections (SC)**
- **Statistical (Optimal) Interpolation (OI)**
- **Variational techniques**
- **Ensemble Kalman Filters**

Adjustment of other model variables (1D problem):

- **Nudging a la PROBE ("Hope model will do the job")**
- **Optimal interpolation (pre-calculated weights)**
- **Variational techniques (linearization problems?)**
- **Simplified Extended Kalman Filter**
- **Ensemble Kalman Filters**

Spatial spread - SC

$$T^A(\tau, x, y, z, D) = T^A(\tau - 1, x, y, z, D) + \sum_{obs=1}^{NOBS} w(\delta x, \delta y, \delta z, \delta D, \tau) \times (T^{obs}(xobs, yobs, zobs, Dobs) - T^A(\tau - 1, xobs, yobs, zobs, Dobs))$$

where

τ is iteration number

x, y, z horizontal position and altitude

D lake depth

$T^A(\tau, ...)$ analysis at iteration τ

T^{obs} observations

w *ad hoc* interpolation weights

Spatial spread – SC - comments

- Ad hoc
- Easy to implement (+)
- A version of SC exists in the HIRLAM SPAN surface analysis package
- Not allways so good results (-); HIRLAM is moving from SC to OI for SST
- May amplify noise if not combined with appropriate smoothing

Spatial spread – OI - comments

- **Similar to SC but only one iteration, where the weights are determined via formulation of covariance models for the background (forecast) errors and the observation errors. Matrix problem in observation space.**
- **The covariance models need to be formulated carefully in order to avoid ill-conditioning**
- **OI is used for surface assimilation in HIRLAM (SPAN) with a box data selection scheme and in HARMONIE (CANARI) with a local data selection – applied to 2 meter temperature and relative humidity, to SST and fraction of sea ice and to snow depth**

Spatial spread - variational techniques - comments

- Minimization of a cost function that measues the distance to a background state and the distance to the observations => Minimization in model space (3D-Var or 4D-Var) or in observation space (PSAS)
- Many 3D-Var and 4D-Var schemes apply spectral transforms to handle the background error covariance <=> Implicit assumptions of homogeneity and isotropy <=> Not so well suited for surface variable assimilation
- Wavelet transforms may be used to introduce anisotropy and inhomogeneity; Research is ongoing.

EnsKF - comments

- Ensemble Kalman Filter data assimilation schemes provide ensemble perturbations also for surface variables and these can, at least in principle, be used for spatial distribution of observation increments as well as for adjustment of model variables that are not observed.
- Small ensemble size (that we may afford) => small correlations are poorly estimated => Covariance localization or local filters are applied.
- NILU in Norway will try to implement EnsKF for surface assimilation with SURFEX and HARMONIE (A research fund application is being evaluated).
- Met.no will try to develope a hybride 3-4D-Var/EnsKF, in the first instance for the atmosphere (Also the COSMO consortium for the global model).

Adjustment of model variables - nudging

The nudging approach applied to PROBE for the Baltic Sea, and applied also in the SMHI ocean model HIROMB for assimilations of SST:s, could possibly also be applied to a lake model like FLake. But there are open issues:

- Should nudging be applied only to sensible heat flux?
- Will it work for large nudging forcing?

- Maybe it is enough to just replace the mixed layer temperatures with the observed (analysed) ones?

Adjustment of model variables - OI

The present soil assimilation scheme in HIRLAM and HARMONIE is used to modify soil moisture and soil temperature from 2 meter temperature and relative humidity analysis increments. It is based on pre-calculated weights obtained by Monte Carlo simulation runs with the ISBA surface scheme. Can this technique be used also for Lake data assimilation? Yes, but

- One main drawback of the technique is that the weights need to be re-calculated when the surface schemes changes. We have experienced this problem with the new HIRLAM surface scheme.

Adjustment of model variables - variational techniques

One may imagine a range of variational techniques for the adjustment of the model variables in lake data assimilation, for example:

- **Use of lake model variables as control variables in 4D-Var**
- **2D-Var (a single lake position and time), keeping the atmosphere fixed**

BUT

- **Are TL and AD versions of the lake model adequate?**
- **How to treat interpolation/filtering in the heterogeneous environment?**

Adjustment of model variables - Simplified Extended Kalman Filter (EKF)

With the EKF we project the forecast error covariance for the model variables forward in time with the Tangent-Linear version of the forecast model. This covariance is then used to estimate un-observed model variables from what is observed.

A simplified EKF is being developed within HIRLAM and ALADIN for soil assimilation. The simplifications are:

- **The atmosphere development (model background trajectory) is not changed during the assimilation window**
- **The TL model (the Jacobian) is derived by perturbation of one model variable at a time and model sensitivity runs**
- **Advantages: Model changes are automatically taken care of! Flow dependency!**
- **Applicable also to lake models? Needed?**

Lake data assimilation in HIRLAM/HARMONIE ?

Start simple, introduce refinements gradually !

Initially:

- **Put emphasis on collection of observations and satellite retrievals.**
- **Quality control to reject poor observations (cloud contamination etc.)**
- **Use availble OI schemes in SPAN and CANARI for spatial quality control and spread.**
- **Replace mixed layer temperatures with analysed values**

Later refinements:

- **Take lake depth and altitude differences into account in the spatial spread**
- **Use Simplified EKF for the assimilation of other model variables**