Recent progress in developments of data assimilation for HIRLAM and HARMONIE

Nils Gustafsson
SMHI
With help from Jelena Bojarova, Magnus Lindskog and many other colleagues.
ALADIN/HIRLAM close collaboration started in 2005
Powerful and flexible research tool designed for synoptic scale systems.

- **HIRLAM** (HIgh Resolution Limited Area Model)
- **HARMONIE** (Hirlam Aladin Regional/Meso-scale Operational NWP In Europe)
- **ALADIN** (Aire Limitée Adaptation Dynamique Développement InterNational)
- **ECMWF IFS** (Integrated Forecasting System)

The goal: move operational activities here but is still under development.
Model domains in HIRLAM consortia

- **HIRLAM 7.3 RCR**
  - (15 km hor res, 60 vert lev)

- **HIRLAM 7.4 RCR**
  - (7 km hor res, 65 vert lev)

- **SMHI HARMONIE (ALARO)**
  - (5.5 km hor res, 60 vert lev)

- **DMI HARMONIE (AROME)**
  - (2.5 km hor res, 65 vert lev)

- **HIRLAM ref DA: 4D-Var**
- **HARMONIE ref DA: 3D-Var**

Focus is moving towards frequently updated short-range km-scale forecasts.
HIRLAM research:

A number of scientific research topics has reached a mature stage!!! (are published, in review or will soon be published).

Emphasis:
✓ Flow-dependent data assimilation methods for upper air;

✓ Novel approaches toward initialisation of small-scale processes (RUC; Radar, GPS, IASI, improved precipitation forecast; towards assimilation of cloud observations )

✓ Surface modelling and data assimilation
HIRLAM 4D-Var components (HARMONIE/ALADIN is similar):

- 3D-Var
- Tangent linear and adjoint of the semi-Lagrangian (SETTLS) spectral HIRLAM.
- Simplified physics packages: Buizza vertical diffusion and Meteo France (Janiskova) package (vertical diffusion, large-scale condensation and convection).
- Multi-incremental minimization (spectral or gridpoint HIRLAM in outer loops).
- Weak digital filter constraint.
- Control of lateral boundary conditions.
- Large scale error constraint (Jk)
- New moisture control variable (improving Gaussianity)
- Control of model error (weak constraint 4D-Var)

(Gustafsson et al., 2012, inpress)
Multi-incremental minimization with \( N \) steps in an outer loop

\[
\delta x_0 = \sum_{\tau=1}^{N_\tau} \delta x_0^\tau.
\]

For each step in the outer loop, the TL/AD models and the observation operators are re-linearized around a non-linear full resolution model solution and a quadratic minimization problem is solved.

\[
J_b(\delta x_0^\tau) = \frac{1}{2} \left( \sum_{l=1}^{\tau} \delta x_0^l \right)^T B^{-1} \left( \sum_{l=1}^{\tau} \delta x_0^l \right)
\]

\[
J_o(\delta x_0^\tau) = \frac{1}{2} \sum_{k=0}^{K} \left[ H \left( M_k (x_0^b + \sum_{l=1}^{\tau-1} \delta x_0^l) \right) + H^\tau M_k^{-1} \delta x_0^\tau - y_k \right]^T R^{-1} \left[ H \left( M_k (x_0^b + \sum_{l=1}^{\tau-1} \delta x_0^l) \right) + H^\tau M_k^{-1} \delta x_0^\tau - y_k \right].
\]
Effects of multi-incremental minimization including re-linearizations
Comparison of HIRLAM 3D-Var and 4D-Var for the stormy month of December 1999
Improving forecasting of mesoscale storms:

**Danish storm**

+30 h 3D-Var  
+30 h 4D-Var  
4D-Var analysis

**French storm**

+6h 3D-Var  
+6h 4D-Var  
4D-Var analysis
Why does 4D-Var make a better job than 3D-Var in storm cases?

- Improved implicit flow-dependent structure functions.
- Better quality control decisions due to a better background and better implicit structure functions.
- Use of more observation?
- In general, it is difficult to point to a specific reason for improvements, the improvements generally occur gradually through the assimilation cycles.
Assimilation increments

3D-Var
06UTC

4D-Var
06UTC

3D-Var
12UTC

4D-Var
12UTC
Large scale error constraint in HIRLAM 4D-Var (Dahlgren et al., in press)

\[ J = J_b + J_o + J_k \]

\[ J_k(\hat{\zeta}) = \frac{1}{2}(\hat{\zeta} - \hat{\zeta}_{ls})^T B_{ls}^{-1}(\hat{\zeta} - \hat{\zeta}_{ls}) \]

\( \zeta_{ls} \) (large scale vorticity) is constrained by a short ECMWF forecast at the start of the assimilation window.
In case an object that we want to assimilate is outside the lateral boundaries at the start of the assimilation window, but described by observations inside the lateral boundaries later during the assimilation window, we must control it by lateral boundary conditions. Let it *come in*! Similarly observations may tell us to let objects *go out* even if this is not described by the host model LBCs.
Control of lateral boundary conditions
(Gustafsson, in review)

Two formulations were compared:

1. Control LBC at the end of the assimilation window
2. Control tendency of the LBCs over the assimilation window (pre-conditioning)

Similar results were obtained!
Single simulated observation experiment

Simulated observation SW 9 m/s at 32N 12W 3Dec 1999 11 UTC

3 December 1999 Assimilation window 06UTC - 11UTC

Strong SW inflow in the background field

With control of LBC

Without control of LBC

Relaxation towards 0 in LBC zone
Real observation assimilation experiment for December 1999

EXPERIMENTS

(1) nils99_e11_nolbc2 : without control of lateral boundary conditions and with the lateral boundary conditions at the start of the assimilation window equal to zero.

(2) nils99_e11_nolbc3 : without control of lateral boundary conditions and the lateral boundary conditions at the start of the assimilation window equal to the initial condition increments.

(3) nils99_e11_lbc4 : with control of the lateral boundary conditions at the end of the assimilation window.

(4) nils99_e11_lbc3 : with control of the tendency of the lateral boundary conditions over the assimilation window.

SMHI E11 domain
2 control LBC experiments better than no-control HIRLAM; No control NCAR in between!
CORE CONCLUSION (with regard to control of LBCs)

The control of lateral boundary conditions only requires a modest increase in computing time ($\approx 8\%$). Although the impact of the control of the lateral boundaries, as verified by forecast verification scores, turned out to be quite modest, the qualitative evidence provided by the single simulated observation experiments motivates us to recommend the control of lateral boundary conditions in 4D-Var data assimilation for limited area models.
New moisture control variable (QJRMS)

\[
\delta R H^* = \frac{\delta q}{q_s(T_b)} \frac{(\delta R H^*)^{ub}}{\sigma (R H_b + 0.5\delta R H)}
\]
Mesoscale data assimilation and the use of ensembles

- Will assumptions on weak non-linearities at 10 km resolution break down at the km scale? Most likely yes!
- Can ensembles provide information on uncertainties and balances at the mesoscale? Most likely Yes!
- Is it too risky to put all our efforts on 4D-Var, that works on the 10 km scale but may fail at the km scale? Yes!
- Should we drop our investments in variational data assimilation and start with ensemble Kalman Filters or even particle filters? No, EnsKF and particle filters still have basic weaknesses!
- Can we develop hybrid data assimilation schemes that combine the best of the two worlds? We believe so!
Different approaches for using ensembles in variational data assimilation

- Covariance modeling with parameters of the covariance model determined from an ensemble. Use for example a wavelet-based covariance model (Alex Deckmyn; Loik Berre et al. Meteo-France)
- Use the ensemble-based covariances in a hybrid variational ensemble data assimilation (Barker et al. WRF, UK Met.Office, HIRLAM)
- Ensembles can also be used to determine static background error statistics
HIRLAM first approach to use ensembles in 3D-Var and 4D-Var

- Use the ETKF algorithm for re-scaling of a 6h forecast ensemble to an analysis ensemble (estimation of the analysis error covariance). **Status:** A first version seems to provide reasonable results.
- Use ensemble of 6h forecasts to estimate the background error covariance and blend it with the static background error covariance. **Status:** Coded and tested with promising results in one winter experiment and one summer experiment.
The HIRLAM ETKF re-scaling scheme 
(Bojarova et al. 2011) 
a generalized error breeding)

• Follows Bishop et al. (2001)
• Uses observation positions and observation errors in 
HIRLAM 3-4D-Var (no re-coding of obs. operators)
• No localization of covariances (maybe too brave?)
• Simple multiplicative variance inflation based on 
observation innovations
• Ad hoc down-scaling of non-leading eigen-vectors of 
the ensemble space analysis error covariance matrix
• Additive variance inflation using perturbations with 
the structure of the static background error 
covariance (“fresh blood”)
• Mixing with global TEPS perturbations along lateral 
boundaries and in the stratosphere
Diagnosis of ETKF perturbations-horizontal spectra

Figure 10: The horizontal spectral density of the variance for the forecast error of vorticity at 00h (black), 06h (blue), 12h (red) and 24h (magenta), estimated from the ETKF perturbations (a) and from the TEPS perturbations (b)
Lorenc (2003) augmentation of the control vector:

\[
delta x = \delta x_{3d} + \sum_{k=1}^{K} (a_k \circ \delta x_{k}^{ens})
\]  

(3)

\[
J(\delta x_{3D}, a) = \beta_{3D} J_{3D}(\delta x_{3D}) + \beta_{ens} J_{ens}(a) + J_0
\]  

(4)

\[
\frac{1}{\beta_{3D}} + \frac{1}{\beta_{ens}} = 1
\]  

(5)

\[
J_{ens} = \frac{1}{2} a^T A^{-1} a
\]  

(6)

\(a_k \circ \delta x_{k}^{ens}\) means element-by-element multiplication ("localization" similar to Schur product)

The localization and ensemble weights \(a\) can be assumed to have a certain length scale.
First version of HIRLAM implementation

- a has horizontal variations only (horizontal localization) and is controlled in spectral space with the assumption of isotropy – this is equivalent to a horizontal localization of covariances with a Schur product based on a horizontal correlation function.
- The localization is applied for vorticity, divergence, temperature, surface pressure and specific humidity – for better balancing.
- The implementation of the Lorenc (2003) algorithm was quite simple (a few days of coding).
Figure 4. 300 hPa wind and geopotential (left) and 850 hPa temperature (right) taken from the background model state at 21 August 2007 06UTC + 6h; experiment based on equal static and ensemble contributions to the background error variance
Example of ensemble variance (spread) fields (12 members).

Figure 5. Examples of estimated background error variances based on the ensemble of 4th forecast valid at 21 August 2007 12UTC from the hybrid data assimilation experiment hybrid6bag. Wind components at model level 10 (upper left), wind components at model level 20 (upper right), temperature at model level 30 (lower left) and specific humidity at model level 30 (lower right).
Single observation experiments

Figure 6. Assimilation increments from single simulated observation experiments. Left figure: wind and temperature increments at 300 hPa (the wind observation was inserted on 300 hPa at (65N,25W)); Right figure temperature and specific humidity increments at 850 hPa (the temperature observation was inserted on 850 hPa at (40N,30W))
Verification of temperature profiles; 50\% 3D-Var, 50\% ens, 100\% 3DVar; 10\% 3D-Var 90\% Ens

138 stations Area: ALL
Temperature Period: 20070816-20070822
At 00, 12 + 12 24 36 48

No cases

hPa

deg C

RMSE ETKF
RMSE hybrid6, iag
RMSE hybrid3, beta 1, 2
BIAS ETKF
BIAS hybrid6, iag
BIAS hybrid3, beta 1, 2
CASES
Verification of 700 hPa temperature; 50% 3D-Var, 50% Ens, 100% 3DVAR; 10% 3D-Var, 90% Ens.

Area: ALL using 138 stations
Period: 20070816-20070823
Temperature 700 hPa Hours: 00, 12

Graph showing error metrics (RMSE) for different forecast lengths and methods.
Verification scores: “winter case” 17Jan2008-27Jan2008 (too optimistic due to bugs in the reference)

3D-Var versus hybrid approach (ETKF, EnsDA, TEPS)

3D-Var versus 4DVAR approach

Temp. Wind speed Rel. humid.
800 hPa
Hybrid

Notice increments along a front
Outlook - the ETKF and the hybrid VAR/Ens data assimilation

- Test ETKF for perturbation rescaling in GLAMEPS
- Evaluate the impact of the hybrid data assimilation in comparison with 3-4-Var in HIRLAM (the hypothesis is that the hybrid should beat 4D-Var as well as EnsKF)
- Compare ETKF-based forecast perturbations with forecast perturbations based on observation perturbations as input to the hybrid
- Possibly consider the hybrid also for HARMONIE data assimilation
- Investigate whether ideas of particle filters can be applied within the hybrid framework
- 4DEnsVar
Non-additive errors
(phase-/displacement-/alignment-/timing errors)

Handling – two step method

- Estimate the phase error (displacement field) and warp the background state.

\[ Hx_b \rightarrow \text{Estimate} \rightarrow T \rightarrow \text{Warp} \rightarrow x_b(s + T) \]

- Minimize the additive error using standard VAR-method.

Example
HIRLAM model RUC parallel experiments

Parallel experiments over H11 and U11 domains

Summer period:
1 May 2010 - 5 September 2010

Winter period:
13 January 2011 - 28 February 2011

Domains D11/H11/U11:
11 km hor res., 60 vert lev.

Experimental Design

<table>
<thead>
<tr>
<th>Settings</th>
<th>H11</th>
<th>H11+GNSS</th>
<th>U11</th>
<th>U11+GNSS</th>
<th>U11+GNSS+RAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>boundaries</td>
<td>D11</td>
<td>D11</td>
<td>H11</td>
<td>H11</td>
<td>H11</td>
</tr>
<tr>
<td>Cycle (hours)</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Observation cut-off time (min)</td>
<td>70</td>
<td>70</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Observations used in assimilation

- Radiosonde: Yes, Yes, No, No, No
- surface pressure: Yes, Yes, Yes, Yes, Yes
- AMDAR: Yes, Yes, Yes, Yes, Yes
- Mode-S: No, Yes, Yes, Yes, Yes
- GNSS ZTD: No, Yes, No, Yes, Yes
- Radar radial velocities: No, No, No, No, Yes
Parallel experiments
U11 RUC with and without cloud initialization

siebren.de.haan@knmi.nl

1. Transfer of MSG cloud cover to 3D cloud cover in HIRLAM model:
   - cloud cover N from NWC SAF
   - cloud base from (interpolated) synoptic observations
   - cloud top from MSG (10.8 micron channel)

2. Translate N to humidity

Cloud forecast Verification scores

Verification results by comparison of Hirlam cloudiness to synoptic observations (bias and standard deviation of errors) (large verification area over Europe)
REF: Hirlam reference run
MSG: Hirlam run with MSG cloud initialisation
HARMONIE Forecasting system

DOMAINS

AEMET
DMI
FMI
KNMI
Met Eirann
met.no
SMHI
Veðurstofa
<table>
<thead>
<tr>
<th>Domain</th>
<th>Cycle</th>
<th>Size</th>
<th>DX</th>
<th>MODEL</th>
<th>DA</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEMET</td>
<td>36h1.3</td>
<td>384 x 400 x 60</td>
<td>2.5km</td>
<td>AROME</td>
<td>Downscaling or 3DVAR (two suites)</td>
<td>3h HIRLAM 16km</td>
</tr>
<tr>
<td>DMI</td>
<td>36h1.3</td>
<td>384 x 400 x 65</td>
<td>2.5km</td>
<td>AROME</td>
<td>3DVAR CANARI OI_MAIN</td>
<td>3h ECMWF LBC</td>
</tr>
<tr>
<td>FMI</td>
<td>35h1</td>
<td>300 x 600 x 60</td>
<td>2.5km</td>
<td>AROME</td>
<td>Downscaling</td>
<td>1h HIRLAM 7.5km</td>
</tr>
<tr>
<td>KNMI</td>
<td>36h1.2</td>
<td>300 x 300 x 60</td>
<td>2.5km</td>
<td>AROME</td>
<td>3DVAR CANARI OI_MAIN</td>
<td>12h cycling</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Runs at ECMWF</td>
<td></td>
</tr>
<tr>
<td>Met Eirann</td>
<td>36h1.3</td>
<td>540 x 500 x 60</td>
<td>2.5km</td>
<td>AROME</td>
<td>BLENDING CANARI OI_MAIN</td>
<td>HIRLAM 10km LBC</td>
</tr>
<tr>
<td>Met.no</td>
<td>36h1.1</td>
<td>300 x 500 x 40</td>
<td>4km</td>
<td>ALARO NH SURFEX</td>
<td>BLENDING CANARI OI_MAIN</td>
<td>HIRLAM 8km LBC</td>
</tr>
<tr>
<td>SMHI</td>
<td>36h1.3</td>
<td>506 x 574 x 60</td>
<td>5.5km</td>
<td>ALARO SURFEX</td>
<td>3DVAR CANARI OI_MAIN</td>
<td>3h ECMWF LBC</td>
</tr>
</tbody>
</table>
Research activities on data assimilation within the HARMONIE forecasting system:

1) Build up the necessary expertise (Learn the source code and gain the necessary knowledge in OOPS to be able to use this tool efficiently)

2) Invest in the design of the common preprocessing tools and local data exchange (radar data assimilation efforts)

3) Implementation of high-frequency high-resolution observations and the impact assessment through the coordinated extended impact studies

4) Further development of data assimilation algorithms (4D-VAR, ETKF/EDA, Hybrid ensemble variational data assimilation scheme, modelling of background error covariance)
The latest progress with 4D-Var of the HARMONIE

- **3D-Var vs 4D-Var** (2 months)
- 1 Jan 2010-28 Feb 2010
- Swedish ALARO + “old surface”
- Conventional + AMSU-A
- model res. 5.5
- 4D-Var minimization 11 km

**T profile**

**q profile**

**wind speed profile**

**Ps bias + rms**

**Ps time series**
Investigation on Seasonal/Diurnal Dependency of Structure function (Shiyu Zhuang (DMI))

**Dataset:**
- Sampling: 4 Harmonie ensemble 6h forecasts with ECMWF ensemble forecasts as initial and boundary
- Statistic periods
  - Spring: 20100301-20100531 00(night), 12(day) UTC
  - Summer: 20100601-20100831 00(night), 12(day) UTC
  - Autumn: 20100901-20101130 00(night), 12(day) UTC
  - Winter: 20101201-20110228 00(night), 12(day) UTC

**Model:**
- Harmonie 36H1 trunk; AROME; Denmark domain
Handling of Balances

(seasonal variation of coupling of humidity background errors with errors of other variables as derived for km-scale model over Danish domain)

SUMMER (12 UTC)

Fraction of explained humidity variance

WINTER (12UTC)

Fraction of explained humidity variance

VORTICITY DIVERG. T and Ps

Air-mass/flow dependence to be represented
Seasonal dependency

- **Horizontal spectra**  Winter - more energy in synoptical scales;  Summer - more energy in meso scale
- **Vertical correlation**  Slightly wider in summer than in winter
- **Humidity standard deviation**  Larger in summer than in winter
- **Moisture balances**  Winter – coupling between vorticity and humidity is comparable to coupling between unbalanced temperature and humidity; summer – coupling between unbalanced temperature and humidity is dominate. (lower lever)
Diurnal dependency

• **Horizontal spectra**  more energy in meso-scale at day than at night during spring/summer/winter; less energy day/night variation during autumn.

• **Vertical correlation**  Slightly wider at daytime than at night for all season.

• **Humidity standard deviation**  not much diurnal change.

• **Moisture balances**  especially during summer (lower level) – coupling between unbalanced temperature and humidity at daytime is larger than at night.
Concluding remarks

- HIRLAM is still alive and is used as a research tool
- HARMONIE is hopefully becoming operational within a year
- Flow-dependency is important, for storm developments as well as for mesoscale forecasting in general
- The variational ensemble assimilation is very promising – its exact formulation is still open