ALADIN-HARMONIE/Norway and its assimilation system - the implementation phase

Roger Randriamampianina and Andrea Storto

1 Introduction

In the framework of the HIRLAM-ALADIN cooperation agreement, a new assimilation and forecast system is currently under development, with special focus on high resolution assimilation and forecast. The joint efforts of HIRLAM and ALADIN community aim to create a common basic system suitable for use in operational mode. The system has been called HARMONIE (Hirlam Aladin Regional/Meso-scale Operational NWP In Europe).

The Norwegian Meteorological Institute has been the first HIRLAM Meteorological Service to put efforts in building the assimilation part of the HARMONIE system, which recently joint the forecast part in the reference system; ECMWF HPCE supercomputer has been chosen as platform for implementation.

This document provides an overview of the assimilation system with special focus on practical aspects that have been followed during the implementation: the main goal is to address questions about operational and research configurations rather than verification results, with the idea of supplying guidance for the practical implementation to those ones that are interested in marking first steps towards the ALADIN/HARMONIE world.

Basic configuration of the system, reported schematically in Figure 1, consists of using a 6 hours forecast from previous cycle as background, i) updating the sea surface temperature (SST) through the ECMWF SST analysis; ii) extracting and pre-processing all the available and supported observations; iii) performing a surface assimilation based on the ALADIN community Optimal Interpolation software (CANARI) to analyse surface parameters (skin temperature, soil water content) over land; iv) performing the spectral upper-air analysis for vorticity, divergence, temperature, specific humidity and surface pressure, v) running the forecast model after proper downscaling of lateral boundary conditions from ECMWF global model. The assimilation and forecast steps are run 4 times per day (00, 06, 12 and 18 UTC): at 00 and 12 UTC the forecast model produces forecasts up to 48 hours, while at 06 and 18 UTC it produces 6-hour forecast, which is needed as background for the next assimilation cycle. The following sections describe in detail all the steps in the order as they are run in the system.

![Figure 1: Flowchart of HARMONIE assimilation and forecast systems](image)
Practical notes
As mentioned in the Introduction, the ECMWF HPCE server was chosen as the reference high-performance computer system for HARMONIE.
Within the system, screening of observation, variational analysis (3D-Var) and forecast model run in multi-processor configuration, while pre-processing of observations, surface assimilation, pre- and post-processing of forecast model run in single-processor mode.
The compiling tool used for building ALADIN/HARMONIE executables is the GMKPACK (El Khatib, 2005), that allows relatively easy code portability supporting now a lot of different platforms.
For downscaling global models fields and some basic post-processing, and for performing conversion between model formats (e.g. GRIB to FA and vice versa), the reference tool is the GL package by Ulf Andrae (see e.g. Andrae, 2007).

2 System configuration

Domain geometry and climatological fields generation
Two domains have been installed at Met.no. The bigger one (ALADIN16 hereafter) has been set-up in the framework of the THORPEX-IPY project and comprises the North Pole; the smaller one (ALADIN11) has a geometry that matches as close as possible the operational HIRLAM8 (was HIRLAM10) in order to have similar domains for comparison purposes, and to be able to perform higher resolution data assimilation. For both the domains, the vertical discretization consists of 60 eta levels, using the former L60 level definitions of the ECMWF/IFS global model. Geometrical parameters for the two domains are summarized in Table 1, while the two domains are shown in Figure 2.

<table>
<thead>
<tr>
<th>Domain</th>
<th>ALADIN16</th>
<th>ALADIN11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>16 Km</td>
<td>11 Km</td>
</tr>
<tr>
<td>Vertical Levels</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Projection</td>
<td>Polar Stereographic</td>
<td>Lambert Conformal</td>
</tr>
<tr>
<td>Nodes in x</td>
<td>540</td>
<td>270</td>
</tr>
<tr>
<td>Nodes in y</td>
<td>450</td>
<td>405</td>
</tr>
<tr>
<td>Centre coordinates (lon; lat)</td>
<td>(-5.0; 70.0)</td>
<td>(14.0; 66.2)</td>
</tr>
<tr>
<td>Pole coordinates (lon; lat)</td>
<td>-</td>
<td>(-40.0; 66.2)</td>
</tr>
<tr>
<td>Truncation</td>
<td>269</td>
<td>202</td>
</tr>
<tr>
<td>Negative truncation</td>
<td>224</td>
<td>134</td>
</tr>
</tbody>
</table>

Table 1: Domain parameters for ALADIN16 and ALADIN11

Physiographic and climatological data (orography, vegetation index, albedo, aerosol optical thickness, etc.) are interpolated from the global Météo-France dataset.

Pre-processing of observations
Software packages for the data pre-processing (the so-called OULAN and BATOR) are available in each export cycle of the ARPEGE/ALADIN/AROME/HARMONIE (AAAAH) assimilation system. These packages are originally created to extract and pre-process observational data from the French local database. Thus, they need local modification before being used elsewhere. Two local versions of the OULAN package have been recently developed: the Hungarian and the Norwegian one. The Hungarian package was created to extract data taking into account their local formats.
(ASCII, netCDF, binary, GRIB and BUFR), while the Norwegian version was created to extract local or archived data exclusively in BUFR format. For more details about the Hungarian version of the OULAN program, see Randriamampianina (2005 and 2006). Figure 3 shows the Norwegian concept for observation pre-processing. Basic idea is to extract data from one large BUFR data file, containing all the observations. The extraction/reading of the data is done after splitting the large BUFR file into different observation (BUFR) files. Most of conventional data are extracted using the OULAN package, which creates an ASCII observational file (OBSOUL). To learn more about the OBSOUL file generation and its format rules please refer to Randriamampianina (2005 and 2006) and Kertész, (2007). The ODB (Observational Data Base), developed originally at ECMWF, is used for all the observation manipulation routines, and is created using the BATOR program. BATOR can read simultaneously different data sources: conventional data in OBSOUL format, satellite data in BUFR (ATOVS radiances and Atmospheric Motion Vectors) and MSG/SEVIRI radiances in GRIB or binary formats. Note that we use level 1-d radiances of the SEVIRI data. For the pre-treatment (calibration and cloud flagging) of SEVIRI observations, the satellite application facility for nowcasting (SAF/NWC) software package is used. This package is not a part of the ARPEGE/ALADIN/AROME/HARMONIE export packs, but can be purchased from EUMETSAT. To read more about the ODB data structure and format, see Kertész (2005) and Saarinen (2006). The local version of OULAN is not part of the export pack any more. Since the format of the OBSOUL file does not change so often, OULAN is compatible with almost all the cycles. Critical routines are the ones intended for initializing flags related to instrument characteristics (mostly OULAN CAROBS). Changes in these routines may affect specification and quality check of the different observations.

**SST update and surface Optimal Interpolation**

Sea surface temperature analysis comes from the ECMWF SST analysis, and is performed through a mere update of the field before surface assimilation. The CANARI software is the reference assimilation system for surface analysis. It consists of two main steps that are run at once before upper-air analysis: observations screening (background quality check and redundancy check) and optimal interpolation. In the current Met.no configuration, CANARI provides univariate analysis of temperature and relative humidity at 2 meters, from which sensitive parameters for the forecast model (skin temperature, soil water content) are diagnosed.
the future, the possibility to extend CANARI to sea surface temperature, and snow analysis will be investigated. The reason behind the steps sequence (CANARI followed by 3D-Var) relies on the use of surface parameters, such as skin temperature, in 3D-Var radiance observation operators. The CANARI software runs now in single-processor configuration, and an ad hoc observational database is created for the surface analysis, to avoid uneconomical allocation of computational resources (i.e. the CANARI database contains surface observations only) and to preserve different geographical selection of observations before feeding up the database itself: the gridpoint optimal interpolation takes in fact advantage also of observations outside the domain for analysing points close to the boundaries, while for the spectral upper-air analysis, the observations close to the boundaries have to be filtered out in order to avoid undesired effects in the domain extension zone.

**Assessment of background error covariances**

Specification of background error covariances (matrix B) for use within the variational algorithm has been performed applying both the “NMC method” and the ensemble method. NMC method has supplied two errors statistics for winter and summer seasons (called hereafter “winter” and “summer” B). In particular, the “winter B” error statistics are derived from differences between the 48 and 24 hours forecasts for the period from 20061201 to 20070228 with two daily runs, while the “summer B” error statistics are derived from forecasts differences corresponding to the period from 20060601 to 20060830. In both cases, forecasts have been initialized by dynamical adaptation from global ECMWF/IFS analysis and by using three-hourly ECMWF/IFS forecasts as lateral boundary conditions for the limited area model forecasts.

Alternative background error covariances have been computed by downscaling ensemble analysis performed with the ECMWF/IFS global model, and running subsequently 6 hours LAM forecasts for all the 10 ensemble members. Ensemble analysis were obtained from the experiment carried out by Lars Isaksen and his colleagues (Lars Isaksen et al. 2007), which consisted of one month and a half ensemble analysis with perturbed observations and spectral backscatter scheme to simulate analysis and short-range forecast errors. Covariances are then calculated from a one month dataset of differences between ensemble mean and members forecasts relative to 6 hours forecasts, namely 620 differences (31 days, 2 daily runs, 10 couples of ensemble mean minus members differences).

Formulation of B follows the one suggested by Berre (2001): statistics are computed in spectral space assuming covariances being homogeneous and isotropic; horizontal correlations are vertically varying and cross-covariances are calculated using a multiple linear regression scheme.

Main findings about differences in the error structures computed via different methods are summarized as follows:

- Vertical correlations are bigger in the ensemble B for small scales and for NMC B for large scales;
- NMC method yields larger correlation length scale excepted for a few very high stratospheric levels;
- Variances and cross-covariances show the same structure for both methods, despite larger values found via the NMC method;
- Analysis increments for single-observation experiments extend over a much bigger region for NMC-derived background errors, especially for satellite radiance observations.

The impact of the two error simulation methods on the assimilation and forecast system is currently under evaluation. A major theoretical problem is the correct estimation of scaling factor to multiply the covariances,
in order to scale 48 minus 24 forecasts errors to 6 hours background errors (NMC method) or to take into account insufficient data sampling (ensemble methods). These scaling factors have been set-up empirically so far, following suggestions that the application of error diagnostics in observation space (Desroziers et al., 2005) supplied.

**Radiance bias correction**

Direct assimilation of satellite measurements requires correction of biases caused by both measurement problems and errors in the radiative transfer model. The bias correction is based on finding the difference between the observed radiances and those simulated from the model states. Since the bias correction methods were originally developed for global models, their adaptation to limited area models (LAMs) raises further questions. The quality of the bias correction coefficients – scan-angle biases and coefficients for air-mass predictors – depends on the sample of the observation-minus-model-first-guess, obtained at each satellite scan position. The amount of satellite measurements along the scan line is much smaller in case of a limited domain compared to global models. This can cause problems when evaluating the scan-angle biases for a limited area model. Randriamampianina (2005) investigated the implementation of the method published by Harris and Kelly (2001) to correct the biases of the ATOVS/AMSU-A radiances in ALADIN Hungary. Bias correction issues were recognized to be very critical for direct assimilation of radiances. SAF for numerical weather prediction (SAF/NWP) organized a special workshop, where data assimilation experts discussed how to handle the biases in current and future satellites (http://www.ecmwf.int/news/events/meetings/workshops/2005/NWP_SAF/presentations/presentations.html). In this paragraph we discuss the implementation of the method suggested by Harris and Kelly (2001) (off-line Bcor) and the variational one (Dee, 2004 and Auligné et al., 2007).

Additional programs for off-line estimation of the bias correction (Bcor) coefficients are included in the export pack of AAAH system. To take into consideration the problems described above, the export programs and routines should be modified. Contrary to OULAN, these routines/programs depend on ODB, so they must be recompiled for each cycle. Technical problems that we had to face during the local implementation of this method regarded the lack of the NAG package and some division by zero. The NAG package was needed for eigenvectors and eigenvalues computations, while the second problem occurs in case of missing radiance data in some assimilation time-windows, which is likely to happen often in LAM assimilation.

The setup process for the variational bias correction (VarBC) is much simpler since the algorithms are already implemented in the system and needs only namelist and script modifications. Auligné and Dee (2007) prepared a very short but comprehensive user’s guide about how to set up and deal with the variational bias correction with the IFS assimilation system. VarBC can be applied in combination with the off-line Bcor files; however, we implemented a pure VarBC. This means that we start with zeroed bias coefficients and cold start. Bias coefficient tables have been setup by performing analysis during a 2 months period. This period appeared to be enough to obtain “convergence” of bias values, not only for ATOVS (AMSU-A, AMSU-B/MHS) but also for IASI channels. Bias correction coefficients for SEVIRI infrared radiances are currently assumed to be latitudinally (scanline) constant, in both the bias correction methods.

**Bias correction for radiosonde data**

The bias correction for radiosonde data in AAAH is the same as the one used at ECMWF. We use statistical table files from ECMWF (Patrick Moll, personal communication). All what is needed is to activate some switches in the namelist for screening.

**Observation screening**

The screening software runs as standalone step just before the variational assimilation. Getting
further to bias correction related routines, main tasks in the screening are i) detection of cloudy radiances from satellite instruments; ii) initialization of observation errors for remote-sensed observations; iii) redundancy check and time-dependent selection for all the observations and in particular for moving platforms (aircrafts and ships observations); iv) use of full forward model to compute observation minus guess differences; v) background quality check of observations (i.e. rejection of observations whose distance from first guess is too large); vi) dependent decision routines, namely spatial thinning of observations. In particular, Table 2 reports the average spatial thinning for all the observations as adopted in the HARMONIE reference system.

Table 2: Use of observations in HARMONIE upper-air analysis: assimilated parameters and horizontal thinning

<table>
<thead>
<tr>
<th>Domain</th>
<th>Parameters (channels)</th>
<th>Horizontal thinning</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNOP LAND</td>
<td>Z</td>
<td>-</td>
</tr>
<tr>
<td>SYNOP SHIP</td>
<td>Z</td>
<td>-</td>
</tr>
<tr>
<td>AIREP/AMDAR</td>
<td>U, V, T</td>
<td>25 Km</td>
</tr>
<tr>
<td>AMV</td>
<td>U, V</td>
<td>25 Km</td>
</tr>
<tr>
<td>DRIBU/BUOY</td>
<td>Z</td>
<td>-</td>
</tr>
<tr>
<td>EUROPROFILER</td>
<td>U, V</td>
<td>-</td>
</tr>
<tr>
<td>RADIOSONDES</td>
<td>Z, U, V, T, Q</td>
<td>-</td>
</tr>
<tr>
<td>AMSU-A</td>
<td>5 to 13</td>
<td>80 Km</td>
</tr>
<tr>
<td>AMSU-B/MHS</td>
<td>3 to 5</td>
<td>80 Km</td>
</tr>
<tr>
<td>MSG/SEVIRI</td>
<td>5, 6, 7, 9, 10, 11</td>
<td>60 Km</td>
</tr>
</tbody>
</table>

Minimization
The core of the 3D-Var algorithm is the minimization performed by the M1QN3 algorithm (Gilbert et al., 1989). Variational quality control can be configured to run at minimization time, but it’s not implemented in the reference HARMONIE scripts. For more details about the variational algorithm in ALADIN the reader should refer to Fischer (2007).

The forecast system
The forecast was the first tested at the beginning of the implementation. Cycle 31 was initially implemented, using downscaled IFS analyses and forecasts as initial and lateral boundary conditions, respectively. Digital filter is also used as an initialisation technique. We observed consistent and good-quality forecasts for the upper-air fields, but we found some large biases near the surface. The blending technique between 6 hours LAM forecasts and ECMWF analysis has been tested to decrease the bias near the surface, without improving significantly forecasts of surface parameters. It is important to mention that we concentrated more on the analysis system than on the forecast one.

Analysis system monitoring
It is important to have a system that is able to give fast and accurate overview of the functionality of the assimilation system. This is true not only during the implementation phase but also to follow up the operational assimilation cycles. To check the use of observations and the performance of the built analysis system, we implemented the web-based monitoring system developed in Hungary. We did succeed in adding the monitoring of METOP AMSU-A and MHS data to the system. Monitoring of the IASI data uses components of the system. Our future goal is to put the IASI data monitoring on the web site of the system.
3 Analysis and forecast systems evaluation techniques

**Verification**

The HARMONIE system includes two post-processing tools for verification purposes: the “HARMONIE verification monitor” (Andrae, 2007), that calculates verification scores against independent observations (synoptic reports and radiosonde observations) and an additional verification tool to compute bias and root mean square errors of forecasted fields against analysis (coming from a user-defined experiment or the ECMWF global model). Interpolation and extraction of parameters at observation location for the “HARMONIE verification monitor” is performed on-line as post-processing step at the end of each assimilation and forecast cycle, while verification against analysis runs offline.

**Observation contribution evaluation – Degrees of freedom for signal (DFS)**

A tool to calculate Degrees of Freedom for Signal (DFS) for different groups of observations was created. DFS are defined as the derivative of the analysis increments in observation space with respect to the observations, and are computed in practice through randomization techniques (Chapnik et al., 2006). The tool allows to compute the weight of all the observations grouped by observed parameter and observation subtype, providing valuable information about actual (Absolute DFS) and relative (Relative DFS, namely Absolute DFS divided by number of observations for each subgroup) importance of observations in the assimilation system.

**Sensitivity of the forecasts to different observations**

An experimental tool to evaluate the impact of observations on forecasts at 12, 24, 36 and 48 forecast hours is also included in the HARMONIE system, and is under validation at the moment. Perturbation of observations is exploited to study the sensitivity of forecasts to analysis with different observations types by evaluating the relative changes in root mean square error of the forecast. The bigger the error, the more sensitive the forecast to the given perturbed observation.

4 Selected results, concluding remarks and ongoing work

A new version of the observation pre-processor OULAN was created, which works with BUFR files. Since the BUFR format is widely used as meteorological data exchange format, the new version of the OULAN program may be of interest for other institutes/services.

A script system for running a number of configurations and steps of analysis and forecast system has been built, with the aim of being flexible enough for operational and research applications, not only at our Meteorological Institute.

Modifications, needed to solve problems related to the use of high density observations (aircraft and radiances) were introduced into the cycle 31 (CY31T1) of the AAAH export pack by Gergely Bölöni. One should refer to LAMSUB_FULL key to enable this option.

Our investigations showed that for the ALADIN11 domain it is important to take into account the seasonal dependence of background error statistics (B matrix) in order to obtain a better analysis. Figure 4 shows the positive impact of the winter B matrix. One can see that the difference is persistent from the analysis along the forecast ranges. The opposite situation is found in summer cases, where the use of “summer B” improves forecasts scores at all the time ranges.

To exploit the impact of ensemble-derived B, we have to make some refinement in the analysis system. Nevertheless, ensemble based B gave better analyses and forecasts for humidity (results are not shown).
The introduction of CANARI in the assimilations tasks improves the surface and even the upper-air analysis, but the impact can be shown for very short-range forecasts only. We should improve/tune this scheme in the near future to have better impact in the forecasts.

Figure 5 shows the impact of the implemented methods for radiance bias correction. Use of the off-line Bcor resulted in less bias in radiance departures (observation-minus-first-guess), but the analyses increments (observation-minus-analysis) is biased especially in the stratosphere. Applying VarBC, we observe biased radiance departures, but less bias in the analysis increments. We observed a large positive impact of VarBC compared to the off-line technique. Nevertheless,
negative impact of VarBC on temperature in lower troposphere was found. This problem seems due to low-peaking AMSU-A channels (the problem can be seen through monitoring for the channels 5 and 6) which are not handled properly. Auligné (personal communication) proposed a possible solution for this problem.

The preliminary results obtained from computation of the DFS (Figure 6) on analyses showed that the aircraft wind, AMSU-A and AMSU-B/MHS have the biggest impact in absolute term, while in relative values the influence of rare observations (“single observations”) is remarkable - e.g. radiosonde humidity, geopotential from surface measurements and even more from buoys and ships. Concerning our investigation on the sensitivity of forecasts to the observations, the technique mentioned above showed that AMSU-A have the biggest impact on the forecasts for almost all the parameters (Figure 7), followed by radiosonde, whose impact was very strong in the high troposphere. Aircraft observations seem to be very important for short-range forecasts, especially for temperature fields, while AMSU-B data influence mostly the low and high level humidity fields.

In accordance with our tasks at Met.no, our work focuses mostly on the assimilation of satellite data. We will continue to work on the assimilation of humidity and temperature retrievals derived from cloud observations (CloudSat and nowcasting products), GPS-related observations (ZTD and RO), and new instruments aboard Metop (ASCAT and IASI data) and also tuning the use of AMV from different satellite platforms.

To learn more about the ALADIN/AROME/HARMONIE system it is possible to consult the following pages:
- RCLACE web page, containing mostly report on work done in LACE countries, but there are also comprehensive documentations: http://www.rclace.eu/
- ARPEGE/ALADIN/AROME web page, containing lot of useful documentations: http://www.cnrm.meteo.fr/gmapdoc/

The following page is a restricted one for authorised users only. Here one can find documentation for the HARMONIE system: https://hirlam.org/trac/wiki/HarmonieSystemDocumentation
5 Acknowledgements

The authors would like to thank the Météo-France/GMAP staff for their help, in particular Bernard Chapnik for his suggestions in the computations of DFS, Loïk Berre for his precious comments on the background error statistics, Vincent Guidard and Elisabeth Gérard for the implementation of the VarBC and Claude Fischer, Karim Yassad and Pierre Bénard for the fruitful discussion on the use of stereographic projection in ALADIN/HARMONIE system. Help from Alena Trojáková (CHMI) and Gergely Bölöni (Hungarian Met Service) in the assimilation of Seviri data and CANARI implementation is also acknowledged. The authors want also to thank Lars Isaksen of ECMWF for having kindly provided the ensemble analysis used for computing background error covariances. Andrew Collard and Tony McNally of ECMWF are highly acknowledged for their help in the implementation and assimilation of the IASI data. The present study was supported by the Norwegian Research Council through its International Polar Year programme and the project IPY-THORPEX (grant number 175992/S30): www.ipy-thorpex.no and by the EUMETSAT Fellowship program that is funding the "Assimilation of binary cloud cover" (ABC) project. The authors are grateful for encouragement from HIRLAM colleagues, especially from Xiahoua Yang, Ulf Andrae and Magnus Lindskog, which inspires them to continue this work.

References


Auligné, T. and D.P. Dee, 2007. VarBC user’s guide. Available at ECMWF.


