ECMWF Feature article

from Newsletter Number 141 – Autumn 2014

METEOROLOGY

Twenty-five years of IFS/ARPEGE

sheilades/iStock/Thinkstock

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/ftu6mfvy

This article appeared in the Meteorology section of ECMWF Newsletter No. 141 – Autumn 2014, pp. 22–30.

Twenty-five years of IFS/ARPEGE

Jean Pailleux, Jean-François Geleyn, Mats Hamrud, Philippe Courtier, Jean-Noël Thépaut, Florence Rabier, Erik Andersson, David Burridge, Adrian Simmons, Deborah Salmond, Ryad El Khatib, Claude Fischer

The coding of the first version of the IFS/ARPEGE model was initiated by Philippe Courtier and Mats Hamrud at ECMWF in 1987 as a project involving ECMWF and Météo-France – IFS: Integrated Forecasting System and ARPEGE: Action de Recherche Petites Echelles Grandes Echelles. Many scientific projects, sub-projects, and operational and research options have been built around this initial code since then, both on data assimilation and forecasting aspects.

Here, we first describe the rationale for the project, which was originally based mainly on the limitations of the optimum interpolation (OI) assimilation systems that were widely used at the time. We then describe the operational implementation of IFS/ARPEGE before going into more detail on the developments concerning variational assimilation and the spectral model. This is followed by a discussion of various scientific projects associated with IFS/ARPEGE. Finally we describe the recent evolution of IFS/ARPEGE and future developments, and look back at what has been achieved over more than 25 years of cooperation.

The scientific rationale

The main scientific trigger for the project (which was new and somewhat revolutionary in 1988) was variational data assimilation. By that time several studies had already been performed on variational assimilation, most of them theoretical. However, variational techniques were seen as a promising way to cope with several important limitations in multivariate OI assimilation systems such as those then operational at ECMWF and elsewhere.

- Inconsistency between an OI initial state and the dynamics of the forecast model, which created spurious shocks called 'spin-up effects' during the first time steps of the model integration. This led to significant loss of observational information in the form of noise.
- Difficulties in using observations reporting meteorological quantities which are linked in complex ways to temperature, humidity, pressure or winds, especially satellite observations. Consequently, with OI there was only a small impact from satellite data (radiance measurements from TOVS), in spite of rapidly increasing numbers of good quality satellite instruments.
- Challenge of producing a good analysis at the large horizontal scales, as OI by necessity was becoming more and more localized with higher observation density from new satellite instruments.

A variational assimilation comprises solving a minimization problem either at a given time in three dimensional space (3D-Var) or over a time window preceding the model forecast (4D-Var). A 4D-Var algorithm requires the repetitive use of the forecast model over its time window, and of the adjoint of the forecast model (for computing the derivatives needed to perform the minimization efficiently). Such a 4D-Var system has to be coded around a forecast model, its tangent linear version and the latter's adjoint version – see Box A.

In 1987, at ECMWF, the need for a tangent linear and adjoint model was an important factor in the decision to code a completely new model rather than adapt the previous model code. The adjoint and tangent linear models were also seen as necessary ingredients because they are useful not only for running 4D-Var, but also for computing the forecast sensitivity to the initial state, and for computing Singular Vectors (SV) for an ensemble prediction system (these selectively sample initial perturbations with the fastest growth rate).

The development of ECMWF's forecasting system was therefore strongly based on the limitations of OI that was widely used up to that time. The introduction of variational data assimilation was built on theoretical ideas from universities in France and at Météo-France. This brought teams at ECMWF and Météo-France together into what became a collaborative effort that has been very productive now for 25 years. In fact a common global NWP system emerged called IFS by ECMWF and ARPEGE by Météo-France. To this date, new model cycles are linked having Reading or Toulouse variants.

AFFILIATIONS
Jean Pailleux (retired), **Jean-François Geleyn, Ryad El Khatib, Claude Fischer:** Météo-France, Toulouse, France
Mats Hamrud, Jean-Noël Thépaut, Florence Rabier, Erik Andersson, Deborah Salmond: ECMWF, Reading, UK
David Burridge, Adrian Simmons: retired from ECMWF
Philippe Courtier: Commissariat Général au Développement Durable, Paris, France

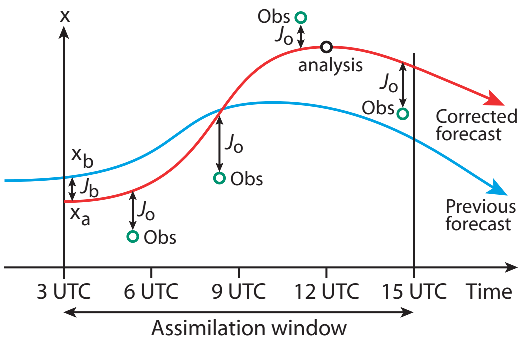
doi:10.21957/ftu6mfvy

Four-dimensional data assimilation

A

The dynamics and physics of the forecast model are an integral part of 4D-Var. Consequently 4D-Var propagates information horizontally and vertically and uses observations in a meteorologically more consistent way. The aim is to seek the initial conditions such that the forecast best fits the observations within the assimilation interval.

- 4D-Var is based on minimization of a cost function which measures the distance between the model from the observations and from the background state. The cost function and its gradient are needed for efficient minimization.
- The tangent linear model provides a computationally efficient (although approximate) way to calculate the model trajectory, and from it the cost function.
- The adjoint model is a very efficient tool to compute the gradient of the cost function.



The figure shows a simplified view of the 4D-Var assimilation technique for a single parameter x. Over a given time window (12 hours here), the observations are compared at their appropriate time with a short-range forecast from the previous analysis. The model state at the initial time of the window is then adjusted to achieve a statistically good compromise between the fit J_o to the previous forecast x_b, and the fit J_o to the observations.

Among the other important scientific ingredients which justified the launch of a new forecast model at Météo-France were the needs for:

- A semi-Lagrangian version of a spectral code.
- A global ARPEGE model with a variable mesh (stretching the global coordinate and tilting the pole of dilatation towards the area of maximum interest).

An important new feature of the IFS/ARPEGE was its ‘integrated’ property (the ‘I’ of ‘IFS’). The idea was to have a code that is sufficiently general, flexible and modular to take an integrated approach to all the computations necessary for a global data assimilation and forecasting system: observation processing, assimilation, forecast model and post-processing feeding the data bases for forecasters (and other users). Previously, analysis, forecast model and postprocessing codes were developed independently, which often led to inconsistencies in the forecasting suite.

From the kick-off meeting of October 1988 until now, the IFS/ARPEGE system has been developed along the above-mentioned scientific lines. Also some additional projects that were not foreseen initially were initiated later on either for scientific reasons (e.g. development in France of a limited-area model called ALADIN) or for software reasons (e.g. change of computer type).

Operational implementation of IFS/ARPEGE

The first operational use of the IFS/ARPEGE code occurred in autumn 1992, both at Météo-France (replacement of the previous global spectral forecast model EMERAUDE by the ARPEGE code at the same T79 resolution, but without stretching-tilting) and at ECMWF (computation of singular vectors for the first operational implementation of the ensemble prediction system). 1993–95 was the period when major changes were made to the model component of the code.

- At Météo-France in October 1993, a stretched-tilted version of ARPEGE became operational in Eulerian mode.
- At ECMWF in March 1994, the IFS model code replaced the original operational spectral model.
- At Météo-France in October 1995, ARPEGE became semi-Lagrangian with a significant resolution increase.

From 1996 to 2000, the global analysis components were completely changed, both at ECMWF and Météo-France, from an OI to a 3D-Var system first, then from 3D-Var to 4D-Var. The operational implementation of 4D-Var occurred in November 1997 at ECMWF and in June 2000 at Météo-France. Although somewhat delayed with respect to the initial plans, they are the only two operational 4D-Var systems implemented before the end of the 20th century.

So far the 21st century has been a period of less intense development of the IFS/ARPEGE systems. However, numerous new observation types have become available, such as infrared hyperspectral sounders and GPS meteorological data. Research on these data types and then their operational implementation have shown the efficiency of variational assimilation schemes for drawing the benefits from these new observations.

The 4D-Var version has been developed to work on time windows of different length (typically of 6, 12 or 24 hours). Also the 3D-Var system has been adapted for the limited-area ALADIN model (used operationally in France and several other countries of the ALADIN consortium). IFS/ARPEGE developments also led to operational suites with increasing room given to ensemble systems and the characterisation of the uncertainty in the analysis and forecast. This is a general tendency in NWP, not limited to ECMWF and Météo-France.

Evolution of the horizontal resolution in the operational forecast models IFS and ARPEGE from 1992 to 2013

B

For about 20 years, both ECMWF and Météo-France have run daily a global high-resolution forecast model, called IFS at ECMWF and ARPEGE at Météo-France. This table shows the progress made on the horizontal resolution during this period for ARPEGE and IFS. Other versions of IFS/ARPEGE have been run operationally, such as ensemble systems (ENS at ECMWF and PEARP at Météo-France). The horizontal resolution of the ensemble forecast has been generally twice as coarse as the high-resolution forecast.

The resolution is noted Txxx/Lyyy/cz, where xxx is the spectral resolution, yyy the number of levels in the vertical, z the stretching factor (always 1, no stretching, omitted for the IFS). When the model formulation started to use a linear grid, instead of a quadratic one (in 1998), the truncation is noted T_L instead of T.

September 1992	First operational version of ARPEGE forecast model which replaces the global model EMERAUDE. Same resolution, same geometry, no stretching (c=1). Semi-implicit Eulerian formulation.	T79/L15/c1
October 1993	First operational version of stretched ARPEGE forecast model (c=3.5) which replaces both the previous ARPEGE and the limited-area model PERIDOT.	T95/L21/c3.5
December 1993	The horizontal resolution of ARPEGE is increased from T95 to T119, and the vertical resolution from 21 to 24 levels.	T119/L24/c3.5
March 1994	First operational version of IFS forecast model which replaces the previous global ECMWF model. Same resolution, same geometry. semi-implicit, semi-Lagrangian formulation.	T213/L31
October 1995	ARPEGE becomes semi-Lagrangian.	T149/L27/c3.5
April 1998	IFS spectral resolution is increased by 50%. Introduction of a two-time-level semi-Lagrangian scheme; use of a linear grid (instead of quadratic).	T_L319/L31
September 1998	ARPEGE resolution upgraded.	T199/L31/c3.5
November 2000	IFS resolution upgraded.	T_L511/L60
January 2002	ARPEGE spectral resolution is increased by 50%. Introduction of a two-time-level semi-Lagrangian scheme; use of a linear grid (instead of quadratic).	T_L298/L41/c3.5
June 2003	The stretching of the ARPEGE grid is decreased from 3.5 to 2.4, while the average spectral resolution is increased to T_L358.	T_L358/L41/c2.4
February 2006	IFS resolution upgraded.	T_L799/L91
February 2008	ARPEGE resolution upgraded.	T_L538/L60/c2.4
January 2010	IFS resolution upgraded.	T_L1279/L91
April 2010	ARPEGE resolution upgraded.	T_L798L70c2.4

doi:10.21957/ftu6mfvy

Building an SI-SL spectral model with several numerical options

Since 1983, both ECMWF and Météo-France have used spectral models (with a triangular truncation) for operational NWP. Unlike the 1983 models, the design of the IFS/ARPEGE model included a semi-implicit semi-Lagrangian (SI-SL) version as the main numerical tool for high-resolution operational forecasting along with a tangent linear version and its adjoint.

The preliminary tests of 4D-Var assimilation used a barotropic model. Shortly afterwards, the baroclinic version of a primitive-equation Eulerian semi-implicit (SI) spectral model became available, together with its tangent linear and adjoint, at least for its dynamical part (almost no physical process included). This version was used for several years as the main research tool for 4D-Var and for ensemble prediction. It also paved the way for operational implementation of the ARPEGE and IFS forecasting models.

Rather than developing a single physical parametrization package for IFS/ARPEGE, dynamical–physical interfaces were set up to accommodate:

- An ARPEGE physical package, oriented towards short-range forecasting (the main job of ARPEGE).
- An IFS physical package, largely based on the physics which was operational at ECMWF at this time (in the predecessor of the IFS).

Both packages had many options, with several switches giving the flexibility to run different physical processes treated in different ways, and different techniques to link the physical and dynamical processes. The price to pay for this additional flexibility was increased complexity in the code.

At ECMWF, the operational stage for a semi-implicit, semi-Lagrangian (SI-SL) spectral model was reached in September 1991, not with the IFS code, but with its predecessor. However, March 1994 is considered as the birth of the IFS as an operational system. This mainly involved a change of code with no important changes to the science or resolution, which remained as T213L31.

At Météo-France, it appeared that combining the technical difficulties of a semi-Lagrangian scheme with the ones having a stretched geometry was leading to slightly more complexity than initially envisaged. Consequently, ARPEGE went into operations in four steps from September 1992 to October 1995. During this period ARPEGE went from a semi-implicit Eulerian formulation at T79L15/c1 to semi-Lagrangian at T149L27/c3.5 (see Box B for more details).

Further development was carried out on the IFS/ARPEGE numerical scheme to increase the spectral resolution by 50%. This was achieved with a two-time level semi-Lagrangian scheme operated on a linear gaussian grid (instead of a gaussian quadratic grid). The linear grid went into operations in April 1998 at ECMWF and in January 2002 at Météo-France. Since then the spectral resolutions have been noted 'T_L' rather than 'T' (see Box B). More recently, following the evolution already implemented in several limited-area models, a non-hydrostatic option has been developed in the IFS/ARPEGE model (see the section on ALADIN, below).

During the 25-year period, progress was made with various numerical aspects required by data assimilation. These are described in Box C.

We now consider the progress made with various scientific projects associated with IFS/ARPEGE. These concern an integrated post-processing package, dynamical core of ALADIN, ensemble prediction system, 3D-Var in ALADIN, and variational assimilation and use of new observations.

Progress made with various numerical aspects required by data assimilation C

- Development of normal mode computation and normal mode initialisation algorithm: such an initialisation process was necessary for most of the OI assimilation schemes.
- Development of Digital Filter Initialisation (DFI): this was originally developed in the ALADIN model version (becoming operational in May 1994), then adapted to the global IFS/ARPEGE model. In Météo-France it replaced the normal mode initialisation in the operational OI code in July 1996. The DFI technique has been used in 4D-Var as a weak constraint to filter out some short waves from the evolution of the model on the 4D-Var time window.
- Development of the tangent linear/adjoint of the semi-Lagrangian model: combining the semi-Lagrangian technical constraints with the tangent linear/adjoint constraints led to a complex development which could be achieved only in the mid-1990s.
- Development of simplified physical packages together with their tangent linear/adjoint counterparts: this was necessary in the incremental 4D-Var, where tangent linear and adjoint computations are performed many times with a simplified model (lower resolution, simplified physics) to run the minimization algorithm.

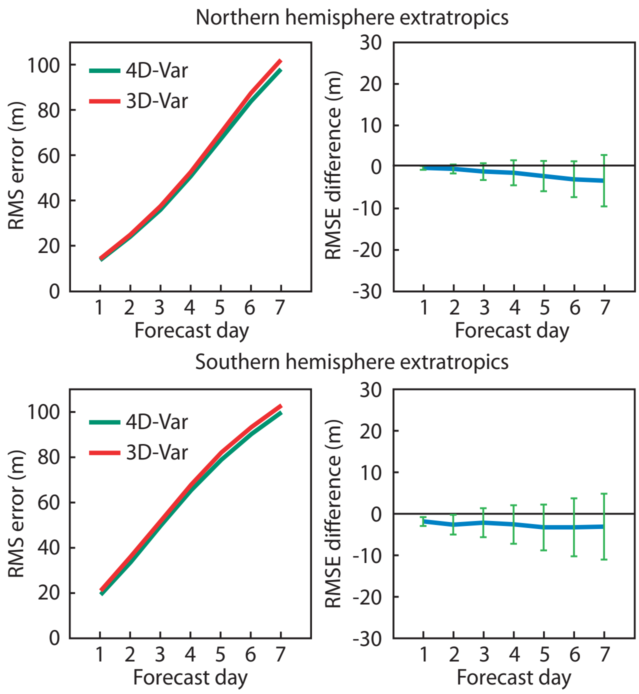


Figure 1 Comparison of root-mean-square (RMS) error for a set of pre-operational forecasts started from 3D-Var and 4D-Var analyses (left) and corresponding differences in RMS error between the two sets of forecasts with 95% confidence intervals (right) for the northern hemisphere extratropics (top) and southern hemisphere extratropics (bottom) (taken from Bouttier & Rabier, 1998, *ECMWF Newsletter No. 78).*

From optimal interpolation (OI) to variational assimilation

There was a general recognition that initial data (the analysis) is of critical importance for success in forecasting the medium range. However, as already explained, from about 1986 there had been growing frustration with operational OI assimilation schemes.

Initially (in 1986), the variational approach was seen as an effective means to provide a global assimilation scheme without the need for local observation selection. Also, it was noted that 4D-Var had the potential to introduce consistent use of the dynamics in assimilation. A year later it was recognised that variational methods would provide a solid foundation for the assimilation of satellite data. For example, experimentation with so-called 'physical retrievals' at the UK Met Office or 1D-Var at ECMWF had demonstrated the benefit of using model fields as first guess for the retrieval of temperature and humidity profiles from the observed radiances – this would be possible also in 3D-Var.

Early on during the development of the variational scheme, ECMWF was able to demonstrate that 4D-Var could generate well-balanced fields on the large scales, create flow-dependent, vertically sloping corrections from single-level observations, and induce wind-field information from a time-sequence of humidity-sensitive satellite radiance observations. These important results showed that the theoretical advantages of 4D-Var would be realized in practice.

All the main building blocks of 4D-Var were already in place when 3D-Var was implemented at ECMWF in 1996. Once sufficient computer power became available to run 4D-Var, intense experimentation demonstrated that all the benefits of 3D-Var were present using 4D-Var with the additional benefits of improved dynamical consistency. Figure 1 shows the impact on the 500 hPa geopotential height of going from 3D-Var to 4D-Var with everything else being kept identical. Though the impact might look small, it was an essential step towards further improvements of the assimilation system.

The development of the concept of 'incremental' 4D-Var was important for enabling an efficient first operational implementation of 4D-Var at ECMWF in 1997. At Météo-France, a few years later, the concept was expanded to keep the cost increase within a factor of 3 for its operational implementation. This led to the so-called 'multi-incremental' 4D-Var. In addition, it was necessary to introduce some filtering as part of the minimization process as well as a small degree of explicit filtering for the final result of the whole 4D-Var procedure. Last but by no means least, major effort went into developing a sophisticated linear physics package to improve the realism of the tangent-linear and adjoint models used in the 4D-Var minimization.

doi:10.21957/ftu6mfvy

Dynamical core of ALADIN

An exciting possibility was offered by the development of the stretched-tilted version of ARPEGE, thanks to the novelties (with respect to IFS) of a variable map factor and of a grid-point by grid-point compass. This led to the development of a limited-area model named ALADIN. It has about 90% of its dynamical software in common with IFS/ARPEGE. Furthermore, ALADIN can share with ARPEGE the characteristics of the physics-dynamics interface as well as links with variational tools and the post-processing package known as FULL-POS (see Box D).

The basic development of ALADIN took place from 1991 to 1994 by a consortium of ten National Meteorological Services (NMSs) coordinated by Météo-France. Six other NMSs have joined the consortium since then. A key step was the signing of a software agreement between ECMWF and Météo-France. Consequently, ALADIN could benefit from most of the progress made with IFS/ARPEGE after only a small delay. But there was also a symmetric advantage: several innovations could be tested at an early stage in the ALADIN framework rather than in global conditions (e.g. parametrization schemes targeted at higher resolutions than currently used). Figure 2 shows the geographical locations of the ALADIN models that are run operationally in 2014.

A non-hydrostatic variant, known as ALADIN-NH, was developed from 1993 to 2003; it had only minor deviations from the basic hydrostatic system used in ALADIN. This development was then brought back to the global framework and is currently being tested at ECMWF as one possible option for a new IFS configuration. ALADIN-NH has been operational at Météo-France since 2008 as the dynamical backbone of the AROME mesoscale forecasting system.

AROME and variants of the model are now being used operationally or tested by many partners within the ALADIN and HIRLAM communities.

FULL-POS: an integrated post-processing package

D

The post-processing package was integrated software inside the IFS, but limited to a small set of fields and only vertical interpolations on pressure levels. However, Météo-France raised the need to develop a more comprehensive post-processing package, especially to filter out the spurious waves and provide 'modern' dynamical fields such as potential vorticity or fields on isentropic levels. The forecasting model already provided the necessary tools so an extended post-processing package was developed using the same routines and numerical computations as the main model and data assimilation.

Today, FULL-POS is intensively used for its original purposes, but also for new ones. In addition, it has been regularly expanded to output new fields.

Recently, FULL-POS has been greatly modernised in the context of the OOPS project (see below the section on recent evolution) and to face the challenges of producing very high resolution forecasts on future computer architectures.

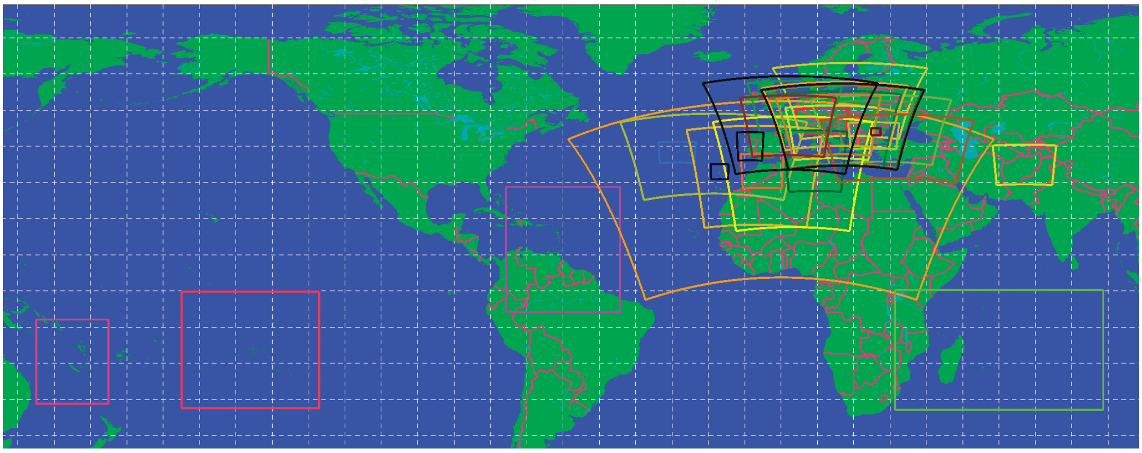


Figure 2 Map showing the domains covered by operational runs of the ALADIN model in weather services of the consortium as at spring 2014 (http://www.cnrm.meteo.fr/aladin/).

Ensemble prediction systems

Ensemble forecasting was already being studied at ECMWF in the 1980s before the start of IFS/ARPEGE. In 1992, the first ensemble forecast was run operationally at ECMWF. Different model runs (or members) were started from different initial states. These initial states were constructed by adding various perturbations to the standard analysis, with the perturbations intended to span the uncertainty of the initial state. The computation of these perturbations was the first operational use of the IFS/ARPEGE code at ECMWF.

Progressively, the ECMWF ensemble forecasting system improved: more members, uncertainty introduced into the forecast model, and better techniques for computing the perturbations and calibrating the ensemble scatter. In particular, the construction of initial perturbations has made increasing use of the perturbations inherent to the Ensemble of Data Assimilations (EDA). In 2014 at ECMWF, the ensemble forecast (now referred to as ENS) is run twice daily up to day 15 with 51 members (i.e. 51 IFS integrations at T_L639) and is coupled with an ocean circulation model.

At Météo-France, the development of a global ensemble prediction system started later than at ECMWF, based on stretched-tilted versions of ARPEGE, and mainly oriented towards short-range forecasting (up to day 4) for Western Europe. The ARPEGE ensemble prediction (called PEARP: Prévision d'Ensemble ARPEGE) became operational at Météo-France in 2004. Nowadays it is run twice daily up to 108 hours, with 35 members (i.e. 35 stretched ARPEGE integrations at T_L538/c2.4).

Ensemble prediction systems based on a limited-area model (LAM) have been used at several NWP centres based on the ALADIN model. In particular, the large overlap between both spectral representations (global and limited-area) allows use of the so-called 'blending technique' (merging the interpolated large scales of the global model with the smaller LAM scales) as a vehicle for developing more consistent perturbations in LAM ensembles. At Météo-France a higher resolution ensemble system has been developed based on the AROME model.

Some key articles about the development of IFS/ARPEGE

- Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. 1986, *Tellus,* **38A**, 97–110.
- Application du contrôle optimal à la prévision numérique en météorologie. 1987, *Thèse de doctorat de l'Université Pierre et Marie Curie (Paris VI).*
- ECMWF's 4D-Var data assimilation – the genesis and ten years in operations. 2008, *ECMWF Newsletter No. 115,* 8–12.
- A global numerical weather prediction model with variable resolution - Application to the shallow-water equations. 1988, *Q. J. R. Meteorol. Soc.,***114,** 1321–1346.
- A strategy for operational implementation of 4D-Var, using an incremental approach. 1994, *Q. J. R. Meteorol. Soc.,* **120,** 1367–1388.
- Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. 1995, *Mon. Wea. Rev.,* **123,** 515–535.
- Improving the efficiency of a digital filtering scheme for diabatic initialization. 1997, *Mon. Wea. Rev.,* **125,** 1976–1982.
- Ten years of research and operational activities with the IFS. 1997, *ECMWF Newsletter No. 75,* 2–7.
- Simplified and regular physical parametrizations for incremental four-dimensional variational assilimation. 1999, *Mon. Wea. Rev.,* **127,** 26–44.
- The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. 2000, *Q. J. R. Meteorol. Soc.,* **126,** 1143–1170.
- Impact of the digital filter as a weak constraint in the pre-operational 4D-Var Assimilation System of Météo-France, 2001. *Mon. Wea. Rev.,* **129,** 2089–2102.
- An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system. 2005, *Q. J. R. Meteorol. Soc.,* **131,** 3477–3492.
- The use of an ensemble approach to study the background error covariances in a Global NWP Model. 2006, *Mon. Wea. Rev.,* **134,** 2466–2489.
- ECMWF's 4D-Var data assimilation – the genesis and ten years in operations. 2008, *ECMWF Newsletter No. 115,* 8–12.
- Dynamical kernel of the Aladin-NH spectral limited-area model: revised formulation and sensitivity experiments. 2010, *Q. J. R. Meteorol. Soc.,* **136,** 155–169.
- A new method for generating initial condition perturbations in regional ensemble prediction system: blending. 2014, *Mon. Wea. Rev.,* **142**, 2043–2059.

doi:10.21957/ftu6mfvy

3D-Var in ALADIN and AROME

Efforts to implement a LAM configuration of the variational assimilation code of the IFS/ARPEGE software started in the years 1996–1997. However, before this time, a LAM data assimilation already effectively existed in ALADIN as a part of the OI code. A major motivation for the development of these systems was to enable the assimilation of local (national, say) observations in a regional data assimilation system in addition to any observation type already made available to the LAM via the shared IFS/ARPEGE framework. Another motivation was to complement the global assimilation system with a local one, whose corrections to a given forecast state would encompass much finer scales than those of the global system.

As a concrete outcome, the first two 3D-Var LAM data assimilation suites became operational in 2005 in Hungary and at Météo-France. As of today, about half of the sixteen ALADIN partner NMSs run 3D-Var operationally.

At present, ALADIN 3D-Var is still operated at Météo-France for its overseas applications. An even more appealing application is the 3D-Var system becoming the backbone of the AROME data assimilation for the convection-permitting forecasting system. For example, it has been extended to assimilate ground-based radar observations and work is ongoing to implement very high frequency 3D-Var assimilation with a grid-mesh of about one kilometre. Furthermore, ensemble-based techniques for a convection-permitting LAM and software aspects remain important topics for collaboration in the coming years.

ALADIN and HIRLAM communities now join forces to further develop LAM applications in the same framework as the IFS/ARPEGE code, including the data assimilation component.

Development of variational assimilation and use of new observations

Collaboration between ECMWF and Météo-France greatly helped develop the Ensemble of Data Assimilations (EDA) for estimating the errors of the short-range forecast (background errors) used in the assimilation. This technique was first used at ECMWF for the production of climatological homogeneous covariances (and then heterogeneous correlations) of background errors.

At Météo-France there was the implementation of climatological but heterogeneous error variances based on EDA, allowing data density and storm track features to be represented. Research then progressed until it was possible to use the EDA to produce 'errors of the day', quantifying appropriately the errors as a function of the actual weather conditions (e.g. typically, errors are larger in active weather conditions than in anticyclones). The implementation of elaborate filtering methods was a key development at Météo-France, allowing daily estimation of background error variances from a small ensemble (six members when first implemented in 2008). This advantage is unmatched today by Ensemble Kalman filter systems, which require much larger ensemble sizes. Close collaboration on the EDA ensued with ECMWF, where the errors of the day were implemented in 2011. Since 2013, both centres have been using the EDA to produce flow-dependent correlations for the background error specification.

In the 1990s there was impressive progress at ECMWF in the use of satellite observations when going from using retrieved profiles to the direct use of radiances. The use of 'raw' AMSU-A data from the NOAA/ATOVS satellite started in 1999. This paved the way for steady advances throughout the 2000s when the number and variety of satellite observations increased tremendously. In particular, the hyper-spectral sounders AIRS and IASI could be used with a large impact on forecast performance. Météo-France directly benefited from these developments, and started a major activity on the use of satellite observations, once 4D-Var was implemented in 2000. Several important research activities initiated at Météo-France benefited both systems (e.g. development of the assimilation of AMSU radiances over land and the assimilation of ground-based GPS data).

Figures 3 and 4 show the impressive increase in the amount of data assimilated at ECMWF since the implementation of 3D-Var in 1996. Similar results can be found for the ARPEGE assimilation on *http://www.meteo.fr/special/minisites/monitoring/menu.html*.

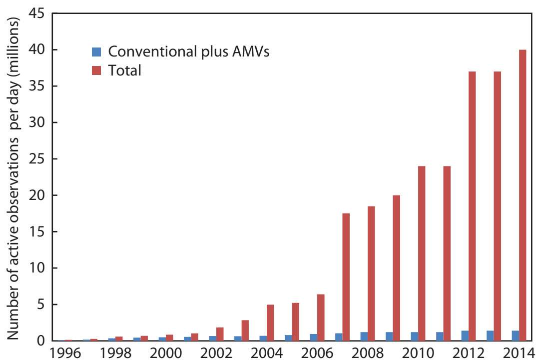


Figure 3 Evolution of the number of 'active' observations within ECMWF's 4D-Var system. Shown are the total number of observations and the number of conventional observations plus atmospheric motion vectors (AMSs). Observations are considered active if they are effectively used, indirectly used (for cloud detection, emissivity retrieval, etc.) or quality controlled and bias corrected ready to be activated (provided by Mohamed Dahoui, ECMWF).

doi:10.21957/ftu6mfvy

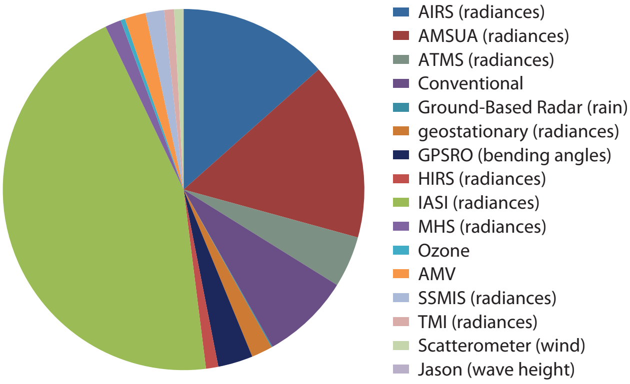


Figure 4 Contribution of observation data types to the total of 'used' data at ECMWF. The values are representative of April 2014 (provided by Mohamed Dahoui, ECMWF).

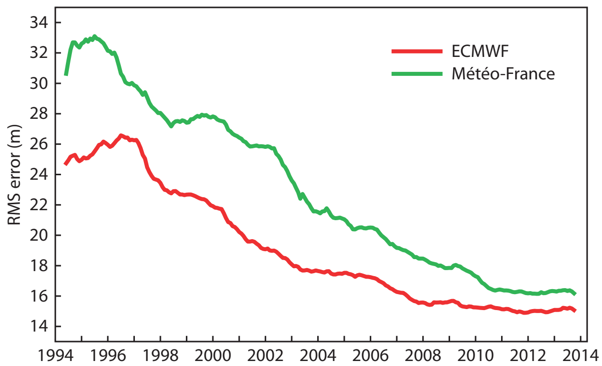


Figure 5 Root-mean-square errors for the geopotential height at 500 hPa at the 48-hour range over the northern hemisphere with respect to radiosonde observations for ECMWF (IFS) and Météo-France (ARPEGE) using a 12-month centred running mean (provided by Martin Janousek, ECMWF and Francis Pouponneau, Météo-France).

Recent evolution of the IFS/ARPEGE system and outlook

Over many years it has been necessary to face the challenge of adapting to and seizing opportunities offered by new computer architectures. A particular issue has been ensuring optimal code efficiency for different applications and platforms on each side of the Channel. Over the years one of the important aspects of IFS/ARPEGE has been the computer performance.

Since 1990, in terms of floating-point operations per second, the operational forecast has increased from about a Gflop to 5 Tflops running on one hundred times as many cores. Much effort has been spent on optimising and adapting the code for each new generation of computer architecture. Consequently, tight coordination for deciding on code evolution, frequency and level of code pruning, code reorganization, etc. is required. On the other hand, there is the benefit that one centre can learn from the other about strengths and weaknesses of software running on different computer systems (including the computers of the ALADIN community where the same code is run).

Radical changes to the IFS/ARPEGE system architecture have recently emerged from an initiative called OOPS (Object Oriented Prediction System). This project was initiated at ECMWF in 2010, in concert with Météo-France and the ALADIN and HIRLAM communities, based on two considerations: the IFS/ARPEGE code has increased in complexity and the current implementation of 4D-Var was considered to not be sufficiently scalable.

Object-oriented programming seemed to be a natural way to respond to the need for a more flexible, efficient and reliable code, and OOPS was initiated to address this. It is worth pointing out that although ECMWF and Météo-France are exploring different ways for evolving their global hybrid 4D-Var data assimilation systems to address the lack of scalability, OOPS will offer both organizations a common high-level flexible framework for testing alternative approaches. The OOPS development fits very well with the spirit of the long-standing collaboration between the two organizations.

More generally, improving scalability of the various components of the IFS/ARPEGE system, including a full rethinking of the dynamical core, grid mesh structures, etc., is an important challenge for NWP.

doi:10.21957/ftu6mfvy

A final look back

When the IFS/ARPEGE project was initiated in the late 1980s, operational NWP was still a young scientific activity, about 20 years old in some NWP centres like Météo-France. ECMWF was about 12 years old and had been operational for less than 10 years. Forecast and analysis models used to be developed as individual pieces of software which were then assembled in an operational suite together with pre-processing and post-processing software. NWP models used to live between 5 and 10 years, sometimes less; then they were almost completely recoded either for installation on a new computer or because the code was not flexible enough to accommodate new scientific ideas. When the IFS/ARPEGE models were implemented, the EMERAUDE and PERIDOT models (predecessors of ARPEGE) had been operational for 7 to 8 years at Météo-France, and at ECMWF the predecessor of IFS had been operational for almost 11 years when replaced.

IFS/ARPEGE was designed with a different philosophy. Although the project was triggered mainly by new scientific ideas, there was a strong will to make the IFS/ARPEGE system general, flexible and modular. When the first version of IFS/ARPEGE was designed it was hoped it would live longer than its predecessors, say about 15 years.

25 years later, IFS/ARPEGE is still going strong, and both systems have shown consistently improved scores within the period (Figure 5). The current IFS/ARPEGE system is quite different from what it was 25 years ago. One could claim that some important parts of the system have been recoded several times. One could even claim that the current system should be renamed as a consequence of all these code changes, with a new name replacing 'IFS/ARPEGE'. Computer evolution has obviously had a big impact on the system. Each change of computer type required some recoding, at least to keep the operational code efficient, which could be interpreted as a lack of generality. On the other hand, to survive for more than 25 years is a good sign that the IFS/ARPEGE code has been kept sufficiently general and robust, as anticipated in its original design. Like the Phoenix of Greek mythology, it was cyclically killed and reborn.

Thinking nowadays about the scientific and technical environment when the IFS/ARPEGE project was launched more than 25 years ago, it appears very risky for at least two reasons: there was no initial guarantee that a 4D-Var assimilation was operationally feasible and many decisions had to be taken simultaneously to create a new forecast model with a stretched-tilted geometry and a new semi-Lagrangian scheme.

In addition to these two scientific challenges, the cooperative project entailed some practical considerations, such as the necessity of rapid exchange of information between tens of scientists at ECMWF and Météo-France. This aspect was helped by the explosive development of Internet and email communication which occurred at the same time.

The risks and the uncertainty were discussed intensively for several years around the launch of the project. Those who took the decision to go ahead were vindicated: a 4D-Var assimilation has now been implemented in most of the NWP centres in the world and global operational forecasts have improved considerably. But it is interesting to consider the following questions. In 1990, was there an alternative for efficient use of the new sets of satellite data? Was there an alternative for having a cost-effective high-resolution model?

Recent experiments performed at very high resolution (T7999) with fast Legendre transforms have shown that there is still a future for spectral models such as IFS/ARPEGE, in spite of the difficulties of achieving scalability with a large number of processors. The main strength of IFS/ARPEGE has been to provide a very flexible and efficient dynamical core, together with several hooks to allow the plugging of many applications, especially the more advanced data assimilation tools. Some atmospheric composition components are now also included in the system. One key feature of the system is that most of its scientific and technical developments have been a 'bottom-up endeavour' with 'top-down control' based on experience.

doi:10.21957/ftu6mfvy