Towards coupled data assimilation systems

F. Bouttier - ECMWF

The current evolution of meteorological models and observing systems mean implies no single data assimilation system is going to be able to blend all the available data. Although 3D-Var/4D-Var are likely to remain at the core of atmospheric data assimilation, several ancillary assimilation systems need to be developed and properly coupled together. This paper reviews the issues, discusses some of the possible solutions and lists the associated problems.

1 Introduction: requirements for NWP

Meteorological models are being improved to represent an increasing variety of physical processes that go beyond the traditional synoptic-scale troposphere. This goes hand in hand with the growing availability of high-quality data, mainly from satellites, to observe the corresponding physical variables: clouds and precipitation including microphysics, state of land and sea surfaces including coverage of snow, ice, vegetation and parameters of the biosphere; chemical content of the stratosphere; dynamics of the deep ocean. The trend is to blend meteorological model design with expertise from other scientific communities to try and create so-called Earth Simulator Systems, which would represent all the important physical processes that interact on the global scale, for applications in forecasting on a wide variety of scales, up to seasonal prediction and climate studies.

In NWP (numerical weather prediction) this evolution has originally been fuelled by the recognition of the importance of various boundary forcings on the meteorological synoptic atmosphere (the troposphere at scales larger than 100km). The evolution of the atmospheric fluid is influenced by cloud processes, surface fluxes of heat, moisture and momentum, and radiation. The priority in research and development has naturally been given to the initialization and forecasting of those processes that have the greatest impact on weather forecasts. This leads to the following sorted list of basic requirements for the initialization of synoptic-scale NWP models: (roughly in decreasing order of importance)

1. very accurate initial tropospheric temperature/wind on synoptic scales,
2. acceptable diabatic forcing, notably tropical convection (weaknesses in the initialization of convection may cause substantial large-scale errors which affect the extratropical systems),
3. good surface forcing fields: orography, roughness, prediction of heat and moisture fluxes,
4. radiative forcing e.g. initialization of the cloud cover,
5. initial humidity (largely forced by the model dynamics),
6. chemistry (some knowledge of trace gases and aerosols is needed to compute the radiative forcing),
7. initialization of clouds and precipitation is regarded as a secondary problem in most NWP centres.

These requirements for forecast models are also valid for data assimilation which is normally formulated as a succession of short-range model forecasts. Slowly evolving imbalances (such as drifts in the soil moisture) may actually be more harmful in data assimilation that in pure forecast mode, because assimilations run for longer periods that can reach several years.

Data assimilation brings additional requirements in order to allow a correct use of observed data (i.e. a meaningful comparison between observed parameters and model fields from the previous short-range forecast): it needs accurate

1. mid-tropospheric temperature and wind (these have fast growing errors. Good background fields are essential for quality control and in 4D-Var for providing the correct flow-dependency to the analysis),
2. low-level and stratospheric temperature and wind (necessary for scatterometer wind dealiasing, the computation of radiative transfer observation operators, and a correct handling of tidal signals),
3. humidity (its assimilation is difficult, and there is little accurate data),
4. ozone (for radiative transfer observation operators and as a forcing to stratospheric chemical processes),
5. land surface variables: soil humidity (which has a long memory), radiative transfer variables (near-surface temperature, albedo, snow, ice...), are all essential for a correct handling of the diurnal cycle in data assimilation,
6. sea state (to compute surface fluxes),
7. clouds, precipitation, aerosols (poorly assimilated although they are essential for radiative transfer observation operators and for controlling the spin-up of the forecast model).

These requirements are particularly important for the use of satellite data. An example of concretisation of the above "requirements for today" is the ECMWF data assimilation system, which is mainly concerned with the skill of intermittent medium-range weather forecasts in the midlatitudes (figure 1):

- Its core is a low-resolution, numerically expensive 4D-Var analysis of air surface pressure, temperature, wind, humidity, in the troposphere and the stratosphere.
- The forecasts rely on a high-resolution global assimilating model, which was made as realistic as possible in order to improve the comparison with observations.
- An experimental stratospheric ozone analysis is coupled with temperature and wind through the model equations, 4D-Var and the radiative transfer observation operator.
- There is no cloud analysis per se, the cloud fields are forced in the model by the other atmospheric model fields during the assimilation.
- The main observations are the conventional in-situ measurements of pressure, temperature and wind, because of their high quality, and a selection of satellite data, because of their excellent coverage. In order to be useful, data must be reliable, frequent and timely.
- The ocean wave model has its own full-resolution wave analysis, which is useful to the atmospheric model's roughness, and to the wave forecasts themselves.

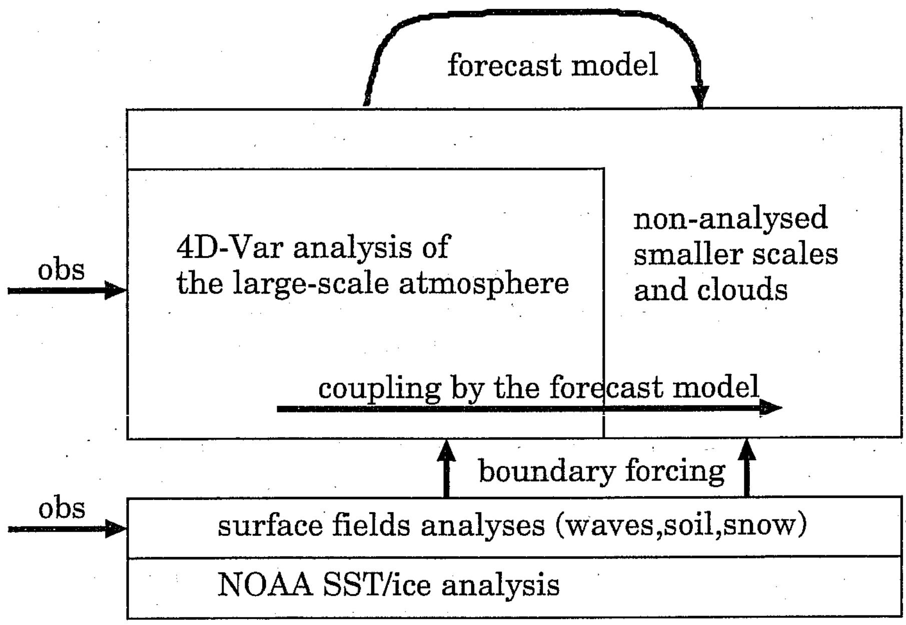

Figure 1: Schematic organisation of the ECMWF data assimilation system.

- Surface analyses of snow, ice, soil temperature and humidity are done at full resolution, using simple algorithms and data selection.
- The SST analysis is provided by NOAA.

Although the 4D-Var system has received most of the publicity, the other surface analysis modules above are important, too: they do not need to work at a very high precision or with sophisticated data, but they are carefully designed and monitored because they have the potential to harm the whole system if they go wrong.

2 Emerging needs and opportunities

Models are able to represent more and more physical processes. New satellites are providing enhanced data that yield information about new physical parameters. Using these data will require model improvements, too. Finally, new degrees of freedom in the models need to be kept under control by the extension of data assimilation to additional variables. It leads to a complex three-way interaction:

- NWP models are extended to represent new processes in assimilations and forecasts, allowing the generation of new products, and the public will expect their quality to keep rising.
- New observing systems and data assimilation techniques improve NWP forecasts and the assimilation of new fields.
- In order to use remotely-sensed data correctly, it is necessary to improve the representation of new processes in the model and data assimilation.

New issues become fashionable because of evolving user expectations, the evolution of satellite remote-sensing technology, or simply their recognition as a promising way to improve weather forecasts:

Clouds and precipitation are beginning to be well observed by satellites (e.g. TRMM, MSG, see the other contributions in this volume) and there is a demand for more reliable forecasts as well as global analyses for climate monitoring;

Ozone, aerosols and other chemicals are increasingly well observed, particularly in the stratosphere. They are connected with urgent environmental problems such as ozone depletion, pollution and UV forecasting, climate change, and their tracking could improve the estimation of atmospheric winds.

The actual weather on land surfaces can be better simulated thanks to new data on surface characteristics, which opens the door to coupling with biosphere models, and to applications in agriculture or seasonal forecasting.

Sea surface fluxes are better known through observations of the sea state and the boundary layers of the sea and the atmosphere. They are essential for deep ocean modelling and seasonal forecasting.

Fine scale modelling becomes possible thanks to new high-resolution, frequent and precise data, with applications for short-range NWP such as thunderstorm and flood nowcasting.

Extreme event forecasting needs to be improved. This probably calls for the targeting of forecasting resources and observations to the appropriate areas.

Errors in NWP forecasts can be better understood and controlled through bias correction of models and observations, provisions of estimates of the forecast and analysis quality, and observation targeting techniques to optimize the observing systems.

More frequent NWP product updates are necessary for short-range forecasting, which will require frequent data and improved, quicker data acquisition and forecasting procedures.

Massive data monitoring will be required from the NWP centers to protect themselves against data quality problems, and as a feedback to data providers. This will be increasingly complex and expensive.

The above emerging requirements have several consequences. First of all, there are implications for the organisation of the work in NWP centres. The investment into each observing system and data assimilation module, is in proportion to the expected return on investment. Scientifically exciting instruments may not provide advances in weather forecasting performance. Due to the limited amount of manpower available for research and development in each NWP institute, there is a danger that issues that are regarded as unimportant, or too difficult, will not be seriously tackled by anyone, leading to persisting weaknesses. A good way of setting priorities for work is to run impact experiments using realistic data assimilation and forecasting systems; their scientific value is at risk if they are carried out by people who were hired to demonstrate the importance of some aspect of the observing system or data assimilation procedure. In an ideal world there should be no duplication of effort between NWP centres; they should coordinate their work, and exchange software as much as possible, so as to ensure that the broadest range of problems is explored. Some important scientific problems are outside the competence of NWP centres (ice sheets, properties of the vegetation...); their study should be left to specialist institutions. Studies need to be performed using a realistic experimental setup: low-resolution models, simplistic analysis systems or simulated observations usually lead to useless results. To allow good-quality scientific collaboration, more open NWP experimentation facilities should be developed, such as (perhaps) ECMWF's `prepIFS` system.

The trend towards more complex models and heterogeneous meteorological data networks raises some technical concerns, too. Data providers may not realize that an overcomplicated encoding of observations, or an encoding that keeps changing, does create major problems in NWP centres. This is often enough to make the difference between a useful and a useless observing system. Data that is not transmitted in real time is often lost for NWP applications. Data with irregular quality or reporting frequency cause extra monitoring work at NWP centres, work that could have been put to improving the use of the data.

One can foresee that the planned increases in satellite data volumes will cause specific difficulties. Data processing has a cost in storage and computing facilities (it vectorizes poorly on the supercomputers used for NWP). Ideally, the density in time and space of the transmitted data should match the resolution of the assimilation system and the scale of the physical phenomena of interest, notably in terms of error growth and propagation properties:

- SST and ice evolve slowly on the scales relevant to NWP;
- Soil wetness evolves quickly and on small scales, but it does not propagate much in space;
- Mid-tropospheric temperature and wind evolve quickly, on rather large horizontal scales but on small vertical scales, with very quick chaotic amplification of errors (locally, in a matter of hours);
- clouds and precipitation evolve very quickly and on small scales, in a chaotic manner;
- chemicals are believed to evolve rather slowly, on small scales (filamentation, plumes and local sources). How quickly do their forecast errors grow ?

The technical facilities must match the volume and complexity of the data that is being used. There is no point in insisting on using of new data types before appropriate diagnostic tools are available — if it takes too long to diagnose problems, people will not try and understand them. It is equally wrong to waste qualified manpower to very fine technical optimization of the CPU and memory used by the data processing and assimilation system.

3 What and how to assimilate

A data assimilation procedure attempts to accumulate observed information into a numerical model. It only works if the data frequency is comparable to error growth rate, and if the amplitude of the errors in the observations (including biases and errors in the observation operator) is comparable with model errors. It means that some fields and observations simply cannot be assimilated:

- Very slow processes: e.g. vegetation type or instrument biases require a specific, independent estimation procedure.

- Very fast processes: e.g. vertical air velocity, precipitation, cloud water and ice (so far) tend to be forced by the NWP model more quickly than they could be assimilated. However, if they are observed, it makes sense to try and correct the processes responsible for their evolution (McPherson 1999).
- Processes which are too poorly modelled: actual weather, lightning, visible satellite images cannot (yet) be directly forced into NWP models.
- Observations which are too poor: divergent wind, potential vorticity, momentum fluxes at the surface, temperature gradients, or soil moisture are all essential features of modern models but little or no direct measurements are available or even planned.

Some of these limitations can be circumvented by a suitable selection of favourable data, weather situation, and by retrieval methods.

Fundamentally, data assimilation is about interpolating between observations which are available at points in time and space. The main problem is to determine the most appropriate interpolation structure: intuitively, one should know about the typical scales of errors for each type of weather, the existence of homogeneous air masses, the presence of coastlines and mountains, and even cloud edges. The next problem is, how can we represent what we know about these structures numerically. This is the purpose of background error modelling, which in turn has led to developing a variety of analysis algorithm, each with its own pros and cons. One- or two-dimensional problems can often be treated empirically using simple techniques such as linear regressions, interpolation by polynomials, Cressman technique, splines or krigeing. The multivariate three- and four-dimensional problems that are found in the modelling of the atmosphere and the ocean normally require more complex statistical interpolation techniques, which rely on the proper modelling of forecast and observation error covariances. The most popular at the time of writing are:

Optimal Interpolation is quickest to set up,

3D-Var (three-dimensional variational analysis) avoids the data selection noise of OI and can handle weakly non-linear problems.

4D-Var (four-dimensional variational data assimilation) is expensive and thus only justified when the following aspects are important:

- an analysis which is consistent with the dynamics of the model in unstable regions such as developing storms,
- a good consistency in the use of observations distributed over time (which limits interpolation errors, and allows observation errors to be averaged out over time),
- a reduction of model imbalances in the forecasts started from the analysis,
- the multivariate coupling between tracer and wind analysis through the model's transport scheme.

Kalman Filtering is even more expensive and only seems warranted for problems with low error dimensionality. Theory predicts that KF is more optimal than 4D-Var, but most KF implementations so far have been too approximate.

All these algorithms attempt to solve the same mathematical problem, which would yield optimal results (i.e. analyses that are as realistic as possible) if they were not affected by the following common approximations and mistakes:

- A misspecified observation/background error ratio for ANY observation or model variable will lead to valuable information being neglected, to overuse of bad data, or even to failure of the numerical algorithm in the case of variational methods (Andersson et al 2000). The solution is to check this ratio regularly for all observations.
- A misspecified background error correlation. When correctly set up, it should enforce a smoothing length scale and some balance properties in the analysis. Mistakes can cause noisy (i.e. unphysical) analyses, spurious corrections of unobserved variables, physical imbalances in forecasts, and a bad observation/background error ratio for complex observation operators such as the ones involving column averages. A solution is to check the structure functions for all observations, by running test analyses with one observation at a time.
- Biases in the model or in the observations can cause the model to drift in the assimilation, or to some fighting between observing systems. This is not always obvious because biases may be situation-dependent and create subtle problems through multivariate coupling. The solution is to estimate and remove the biases from observations and model, taking care not to mistake one from the other (Dee and Da Silva 1998, Dee and Todling 2000).
- Observation-background error correlations are frequent in data that have been retrieved or that involve some background-dependent quality control. These correlated observations often look very good when compared with the model background field, lulling the end user into a misleading sense of security. Actually, such observations may prevent valuable information from entering the assimilation system. The solution is to estimate the correlations, and to reduce the observation weight accordingly (highly correlated observations should be given very small weights).
- Variables with inhomogeneous error statistics will produce a suboptimal or even unphysical analysis. For the statistical interpolation technique to work, the error magnitude of fields and observations should remain stable during the analysis e.g. the background error covariance should always imply physically admissible perturbations of the background fields. This is a problem for variables such as humidity or ozone which are bounded, and which have errors of widely different orders of magnitude depending on the area or layer considered. A solution is to remap the observed and/or model variables to better conditioned ones in the analysis. Unfortunately, bounded variables will remain a problem, as it is not compatible with practical analysis algorithms (an example is binary variables such as ice cover or occurence of rain). Empirical solutions must be sought.
- Linearization and model errors in 4D-Var and KF will lead to unpredictable distortions of the analysis increments. This is a fundamental weakness of these algorithms. A solution is to monitor the linearity for realistic model perturbations, and to avoid trusting 4D-Var or KF blindly in non-linear situations.
- Mutually correlated observation errors are widespread in remote-sensed data, but they rarely are a significant problem. An easy ad hoc solution is to reduce the weight of correlated data. Note that most observation error correlations are the manifestation of a bias problem, which is better handled by bias correction.

- Underestimating analysis errors when cycling the assimilation system will cause background errors to be underestimated, leading to an underuse of observations. This is a common problem when theoretical Hessian-based formulae are used to estimate the analysis error of 3D/4DVar: real analysis algorithms are less optimal than theory would suggest, and their errors are much bigger than what the Hessian indicates.
- Underestimating model error in cycling. Some parts of meteorological models often exhibit less variability than nature. In particular, many (if not most !) model errors originate in the boundary conditions.

The bottom line is that the performance of data assimilation algorithms does not necessarily grow with their sophistication. It is important to question their behaviour in physical terms, in particular whether the information travels in a sensible way from the observation to the model fields.

4 Limitations of 4D-Var and KF

While the above remarks are valid for most analysis algorithms, there are specific problems regarding 4D-Var and the Kalman Filter. Since they are going to be widely use in the near future, it is important to examine them in the light of our hopes for the use of satellite data.

4D-Var and KF rely on assumptions about the linearity of the evolution of errors, proper modelling of model error, and, the approximation of the forecast model by a simpler one with less physics or lower resolution. The tangent linear and adjoint models (including the observation operators) must be linearized in the vicinity of a realistic trajectory. Small-scale phenomena tend to be less linear than large-scale ones, meaning that there is a tradeoff between model resolution and quality of the linearisation. Linearity is not only a property of the flow, it is a function of the amplitude of background errors and frequency of observations, so we may hope that in the future it will be possible to apply 4D-Var or KF more widely, as models and observations improve. Linearity is also a function of the field structure itself: although 4D-Var is able to couple tracer observations with the wind analysis, this will only work for smooth tracer fields, not for discontinuous and small-scale fields such as visible cloud observations. One may wonder whether 4D-Var should use a linearized model that is as realistic as possible, or a simplified one that provides a more robust estimation of the derivatives of the cost function.

Although 4D-Var is able to generate flow-dependent structure functions, it still relies on a proper modelling of background error covariances, unless all degrees of freedom of the model are observed in each assimilation cycle, and the model is perfect. Limits on the linearity and the presence of model error mean that 4D-Var cannot be applied on very long periods, and that its cycling will remain a crucial problem. The presence of model errors will limit the applicability of 4D-Var and KF until they are properly represented in these algorithms. Due to their inherent numerical cost, 4DVar and KF will never be applied at the same resolution as the forecast models.

Because of these limitations, 4D-Var and the Kalman filter cannot be used to assimilate observations related to the smallest scales and to the less linear features of the forecast model (figure 2). This problem will not disappear, because as numerical resources increase, the models are refined to include even smaller scales and less linear phenomena. A specific treatment of small scales is needed outside 4D-Var or the KF, which can only be achieved by coupling them with ad hoc, full-resolution assimilation modules (see e.g. McPherson et al 1996).

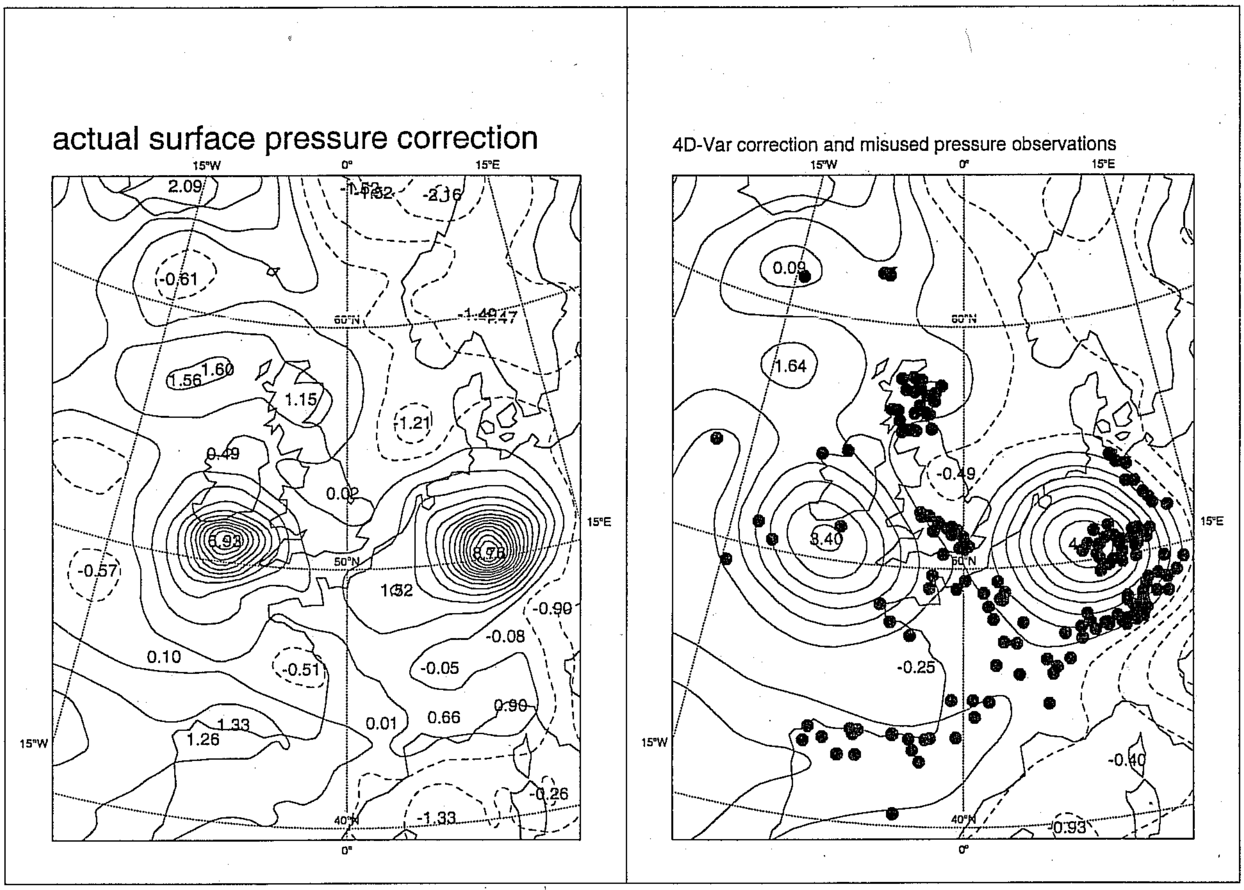

Figure 2: Example of inconsistencies caused by the linearisation and incremental formulation in the ECMWF 4D-Var (12-hour window, spectral resolution of T63 in 4D-Var and T319 in the forecast model, i.e. the version used operationally in October 2000). The case is the analysis of a particularly violent storm over Europe on 27 December 1999 at 12:00 UTC. The maps show the surface pressure correction (in hPa) produced by the analysis (left panel), and the correction that was internally produced by the 4D-Var algorithm (right panel). Due to inconsistencies in model resolution and lack of linearity, the attempted 4D-Var correction of the storm structure (South of Ireland and on Germany) was wrongly applied to the forecast model, leading to serious inconsistencies in the use of many surface pressure observations (dots on the right panel) in the most active areas.

5 Issues in coupled data assimilation systems

This section deals with the problem of making several data analysis systems work together. We will call *analysis module* a self-contained analysis system, with its own observation processing and its own analysis algorithm that converts background fields into analysed fields. These fields may be prognostic fields in a forecast model that couples them. A typical example is soil state (soil temperature and humidity), which is coupled to the model atmosphere through the diurnal radiative forcing and the parameterized surface fluxes of water, sensible and latent heat.

The simplest situation is when the modules do not really interact with each other. This can be implied by the physical properties of the problem: if one can break down the analysis problem into smaller, independent units, it will be easier to manage. For instance, to some extent soil state analysis can be regarded as an independent problem for each model surface grid point, because soil temperature and humidity do not travel much horizontally. In this framework, the time dimension of the assimilation needs to be accounted for, but the full 3-D structure of the atmosphere does not. Such properties simplify the specification of background error covariances. One then needs to rely on the model forecast to blend the information from the independent analysis modules — it is important that the model be robust enough to prevent internal drifts from developing.

The analysis problem may sometimes be broken down using assumptions of scale separation. In the incremental 4D-Var (Courtier et al 1994), only the larger scales are analysed, and it is left up to the model to force the smaller scales to a realistic state. Satellite bias estimation, which is in itself an analysis problem (Dee and Da Silva 1998), can be separated from the atmospheric analysis by assuming it to be slowly evolving and to have homogeneous large-scale properties implied by the form of the bias correction equations,

It may not be obvious whether it is advantageous to couple analysis modules together. The question is, are the errors to correct mutually correlated between several modules, and do we hope to gain anything by explicitly accounting for these correlations. The expected advantages of a coupled system are:

- that it can better account for the physical interactions that account in nature,
- to reduce problems of inconsistency in interpolating data and fields between different modules,
- to make better use of data by the representation of multivariate background error correlations and consistent observation quality control.

The expected problems are:

- more worrying possibilities of problems spreading from one module to the others,
- some likelihood of double-counting of observations when the same data is used in two modules,
- a more complex management of the resulting technical system,
- possible feedback loops between the different modules that amplify analysis weaknesses if one is not careful.

The first step is to select, for each identified module, the best analysis technique:

- In the atmosphere, cycled 3D- and 4D-Var are usually the preferred algorithms. They can be extended to include the estimation of slowly-varying parameters such as bias predictors, or model error parameters.

- Slowly-evolving atmospheric parameters and ancillary variables such as observation bias estimators, or some model error parameters, can be included into 3D/4D-Var as extensions to the control variable of the variational analysis, if a suitable background error model is supplied.

- Memoryless ancillary variables such as skin temperature over land can also be included into 3D/4D-Var. However, a correct background term is still needed to preserve the good behaviour of the analysis.
- Slowly-evolving surface variables such as large-scale SST or ice can be estimated in a largely independent way using two-dimensional interpolation techniques.
- Prognostic surface variables such as the state of soil and snow can be estimated as a two-dimensional interpolation. Since the time dimension seems to be important, it has been proposed to estimate them using one-column variational techniques. How to enforce spatial consistency in this framework is not clear.
- non-linear, fine scale atmospheric fields such as clouds and precipitation can be reasonably well analysed using empirical techniques such as nudging or single-column variational analysis.
- chemicals such as ozone or nitrogen oxides raise special issues if complex and expensive chemistry evolution models are used. It has been proposed to use variational techniques along trajectories predicted by the atmospheric model.

The same observation may be useful for two modules; it is not always clear then how to prevent the same information from being used twice in the assimilation system (the double-counting problem). An example is the use of geostationary satellite images for wind retrievals, or for direct radiance assimilation. Another is the use of rain-related data for humidity retrieval, or for cloud nudging. The theoretical solution is to use realistic observation error correlation models, but this is difficult to implement. A more ad hoc solution could be to artificially separate the scales that can be handled by each analysis module, through a filtering of the observed data.

A related problem is, how to ensure that the same observation is not used in contradicting ways by two analysis modules. The only solution may be to enforce a data selection technique to ensure that it does not happen.

A third problem arises when the same model variable is analysed in two different modules. Discrepancies are liable to occur, and one must decide which analysis is the best. The answer may depend on the scale, the area, and it may require some independent cross-validation data. Some examples of this situation are: the analysis of low-level atmospheric temperature and humidity in the atmospheric 3D/4D-Var and in the soil state analysis, the distribution of clouds implied by satellite radiances or precipitation data, and the surface skin temperature that can be estimated from satellite radiance inversions or directly in SST and soil state analysis.

6 Some coupling strategies

Given the above discussion one can classify the main options for coupling in data assimilation as follows:

uncoupled analyses: each analysis module works independently on the same background state. It is left up to the model to blend the information from the resulting analyses. This approach is

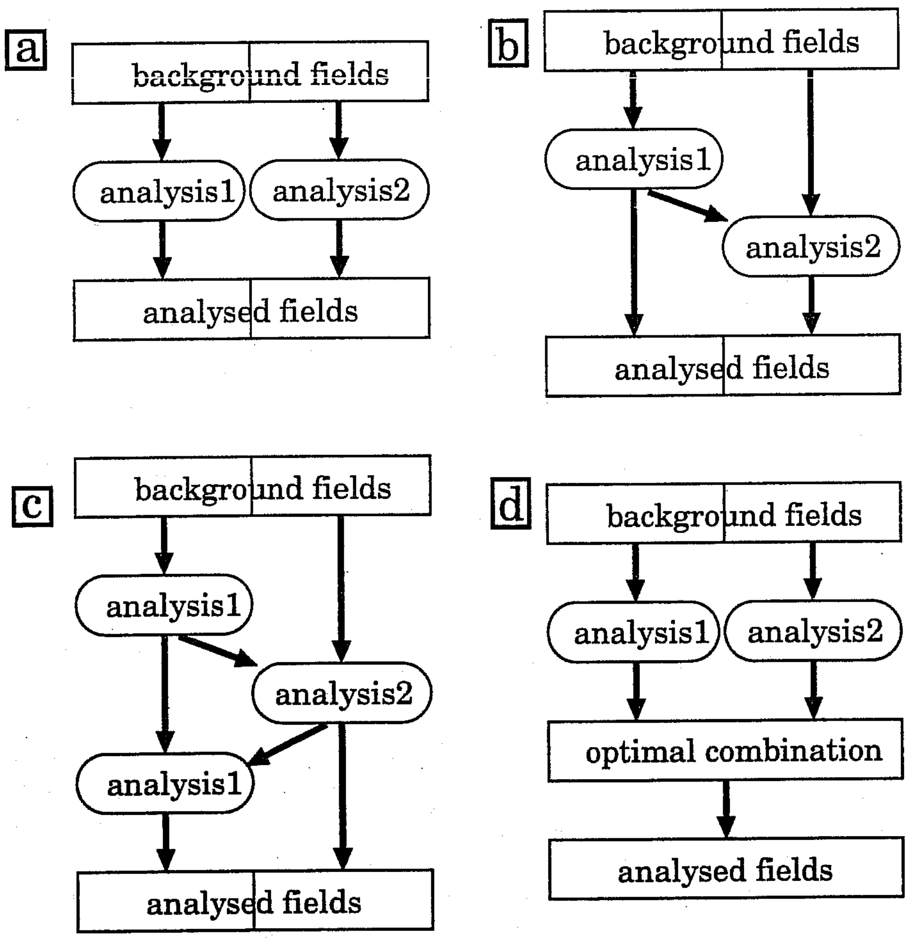

Figure 3: Four data assimilation coupling strategies : (a) uncoupled analyses, (b) one-way coupling, (c) iterated coupling, (d) enforced consistency.

simple but prone to inconsistencies between the different modules. It is not optimal in that the information used by each module is hidden to the other ones at analysis time.

one-way coupling: each analysis is designed independently from the other, but some modules may use as pseudo-background the analysis provided by other modules. This provides some limited consistency between modules, but the first analyses to run have no access to the information from the latest running modules. It may be difficult to specify the error statistics for pseudo-background fields (which are similar to retrievals). This technique is widely used at ECMWF.

iterated coupling: each analysis is done in turn, possibly several times until all modules yield consistent fields. It may be a good technique if one can guarantee that it converges towards a meaningful analysis. It is used in the incremental 4D-Var procedure at ECMWF, and convergence problems have been found. One could incorporate a variety of high-resolution analysis modules inside the incremental 4D-Var, because when the model is relinearized in the first incremental update, its high-resolution fields could be corrected as well. It is probably a good way to implement a cloud and precipitation analysis that is consistent with the atmospheric dynamics.

enforced consistency: in theory, one could perform some uncoupled analyses and then combine the inconsistent fields together in a statistically optimal way. This can be formulated as a special analysis problem, in which the input is a set of preliminary analyses, each with their own uncertainties. This could be a theoretically optimal technique, if one can manage to model the relative uncertainty of analyses, and in particular the cross-correlation of errors between analyses from different modules.

full coupling: of course, if one can afford it and there are no fundamental incompatibilities between the different analysis modules, the best approach is still to formulate the whole coupled system as one single analysis problem. As explained in previous sections, a problem there is the formulation of a suitable background error covariance model.

7 Conclusion

Remote-sensed data from future satellite systems have the potential to provide extra value to atmospheric data assimilation and forecasts through better precision and coverage. However, 3D-Var and 4D-Var analysis systems will remain blind to most of the interesting information because of the underlying linearization hypotheses, the limitations of the models and the low effective resolution of these algorithms. High-quality ancillary fields will be needed to select and to use the new data, notably clouds and precipitation, aerosol distribution, atmospheric chemistry, and surface properties. Tomorrow these fields will be regarded as regular NWP products by the users, and the NWP models will be more and more sensitive to their correct assimilation.

These two requirements (using the new data and providing new products) can only be fulfilled by combining heterogeneous analysis techniques to form coupled data assimilation systems. The 3D/4D-Var paradigm of blending all data directly into the model will be at the heart of such systems, but complementary techniques need to be developed. This will require innovation, the development of a suitable theoretical framework, of algorithms and (perhaps most importantly) of appropriate monitoring tools that guarantee the good behaviour of complex analysis suites.

Such coupled data assimilation systems will be required as the foundation of the planned Earth Simulator Systems of the future, but they will present many difficulties that remain to be investigated.

References

Andersson, E., M. Fisher, R. Munro and A. McNally, 2000: Diagnosis of background errors for observed quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Quart. J. Roy. Met. Soc.,* **126,** 1455-1472.

Courtier, P., J.-N. Thépaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-VAR, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120,** 1367-1388.

Dee, D. and R. Todling, 2000: Data assimilation in the presence of forecast bias: the GEOS moisture analysis. *Mon. Wea. Rev.,* **128,** 3268–3282.

Dee, D., and A. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Met. Soc.,* **124,** 269-295.

McPherson, B., 1999: Operational experience with assimilation of rainfall data in the UK Met Office mesoscale model. *UKMO NWP forecasting research tech report no.289.*

McPherson, B., B. Wright, W. Hand and A. Maycock, 1996: The impact of MOPS moisture data in the UK Meteorological Office mesoscale data assimilation scheme. *Mon. Wea. Rev.,* **136,** 1746–1770.