# Introduction

The present workshop was organised under the joint ECMWF/GEWEX[1] umbrella, thus expressing the importance of a global analysis of atmospheric water both for weather forecasting and climate studies. The humidity analysis is a topic that is currently receiving a great deal of attention at ECMWF. In several cases it has been demonstrated that the medium-range forecast-performance depended on the accuracy of the humidity analysis. Especially the model's development of intense weather systems can be linked with the distribution of humidity in the initial conditions. In the tropics, ECMWF and many other NWP centres report difficulties maintaining a correct intensity of precipitation in analyses and short-range forecasts – a problem referred to as spin-up/spin-down. It is expected that development and refinement of the humidity analysis will lead to improved short-range prediction of precipitation in general – and particularly so in the tropics.

There are good prospects for very significant improvements in the availability of high-quality humidity observations for NWP. The newest radiosonde technology provides accurate measurements, with fast response times. The prospects of having aircraft humidity observations with high accuracy and good coverage are also very promising. A rich source of humidity information is expected with the arrival of data from several new and planned satellite missions on the Earth's hydrological cycle. The aim of several NWP centres is to use the satellite humidity information not only in clear-sky conditions, but also in cloudy and precipitating conditions. To do this effectively, requires substantial development of the data assimilation systems, especially of the humidity analysis.

Thanks to GEWEX sponsorship it was possible to include within the present workshop a comprehensive discussion of the issues surrounding cloud and rain assimilation. These include development of linearized parameterisations of moist physical processes, development of fast radiative-transfer models incorporating the effects of clouds and rain, and their tangent linear and adjoint counterparts.

The workshop followed the usual format of invited lectures followed by discussions in working groups and it concluded with a plenary session. Groups were set up to consider the issues of "Humidity analysis formulation", "Consistency between model and observations", and "Assimilation of clouds and precipitation". The discussions and recommendations of the three working groups are summarised in the following three reports. In order to speed up the production of the workshop proceedings, the lecturers were given the option to either provide extended abstracts or full papers for this volume. The slides of the presentations have been posted on the www.ecmwf.int web site.

ECMWF and WCRP thank all the participants for contributing to a successful and stimulating workshop.

---

[1] GEWEX, the Global Energy and Water Experiment is a component of WCRP, the World Climate Research Programme.

# 1. Working group 1: Analysis formulation issues

*Present: Dick Dee (Chair), Tony McNally (Secretary), Ian Roulstone, Loïk Berre, Christina Köpken, Elìas Hólm, Mike Fisher, Adrian Simmons, Andrew Lorenc and Martin Miller*

## 1.1 Strengths of the current formulation

The current ECMWF humidity analysis system shows some encouraging characteristics. The assimilation of humidity-sensitive data has generally had a positive impact upon forecast quality. Other healthy diagnostics are: A high degree of consistency in the fit of the analysis to different data types; Realistic features in simulated imagery in comparisons with Meteosat data; and good results from comparisons with data from experimental field campaigns (e.g. MOZAIC and POLDER)

## 1.2 Choice of humidity variable

The group considered the appropriate choice of humidity variable an important pre-requisite for any future development of the humidity analysis. Apart from the more obvious desirable characteristics of being reasonably homogeneous, Gaussian and accounting for known flow dependence, it was thought that the variable should also have a well-understood and robust link to observable quantities for validation. The chosen control variable should also be amenable to the imposition of constraints and diagnostic relationships. A number of candidate variables (pseudo relative-humidity, total water, balanced and residual specific-humidity) are currently being investigated at other NWP centres, and ECMWF is encouraged to follow these developments.

## 1.3 Jb-formulation and background-error statistics

The group considered the specification of background-error correlations to be a higher priority for the development of the humidity analysis than background-error variances. In particular, improved specification of vertical correlations seems the most urgent issue. The vertical correlations should reflect important physical features such as the tropopause and the boundary-layer height, as this would be of benefit to the assimilation of low vertical-resolution data, such as satellite radiances from current instruments. While an appropriate choice of humidity control-variable will facilitate this work, some additional techniques, such as vertical co-ordinate transforms and explicit state-dependent parameterisations, should also be considered.

## 1.4 Diagnostic Jb-verification

More diagnostic information should be extracted from the wealth of information available from ECMWF's recent ensemble of assimilations, to diagnose the accuracy and appropriateness of a chosen Jb-formulation. In particular, higher-order statistics of the PDF of background errors should be investigated. The group also felt that more use could be made of the existing innovation (departure) data, multi-centre analyses and observation withdrawal experiments to diagnose the performance of the humidity analysis – both in terms of background error and observation errors specification. The diagnostic adjoint sensitivity techniques, to initial conditions and to observations, were considered useful tools by the group, and these should be explored.

## 1.5 Observation error statistics

More effort should be directed towards improving the specification of observation error statistics for humidity-related observations. A better characterisation of observation errors will improve quality control and analysis accuracy through the formulation of the Jo term, and it will benefit observation-based diagnostic Jb-verification. Accurate knowledge of observation error statistics is also a prerequisite for the development of adaptive estimation schemes that can be used to cycle the background-error variances during data assimilation.

## 1.6 Systematic errors

A number of serious problems with the humidity analysis are associated with systematic errors, in either the humidity observations, the observation operators or the background humidity fields. It was strongly felt that the handling of the biases should be unified and possibly incorporated within the analysis estimation itself. The main difficulties in this are likely to be the parameterization of the bias and a lack of independent information (from observations) to uniquely estimate values of the parameters. The group encouraged an audit of existing information and recommended that WMO and the satellite agencies be involved in funding and participating in special campaigns to investigate systematic errors and biases.

**Recommendation:** That WMO and the satellite agencies be involved in funding and participating in special campaigns to investigate systematic errors and biases in humidity.

## 1.7 Balance and physical constraints

The imposition of obvious physical constraints (e.g. that relative humidity remains between 0 and 100 %) should be performed as far upstream in the processing chain as possible. The statistical correlation of humidity with other variables should, as far as possible, be built into the choice of humidity variable, but other more explicit ways may also be considered. Functional balance (i.e. some humidity analogue to geostrophy relating humidity to vertical velocity) should be investigated, but unfortunately the group could not say what exactly this functional balance should be. The group did agree, however, that these issues could be pursued in collaboration with other centres. If it were possible to observe all aspects of the hydrological cycle then the need for prior constraints in the assimilation system would be reduced.

## 1.8 Stratospheric humidity

The group indicated that good observations related to stratospheric humidity are soon becoming available (e.g. ENVISAT and GPS). These data may be used to assess issues of cross-tropopause humidity transport, which is a major uncertainty in current models and assimilation systems.

# 2. Working group 2: Consistency between model and observations

*Present: Sue Ballard (Chair), Tomas Jung (Secretary), Tadashi Tsuyuki, Ole Vignes, Ralph Petersen, Mariella Tomassini, John Nash, Per Kållberg, François Lalaurette, Antonio Garcia-Mendez, Erik Andersson, Anton Beljaars, Stephen English, Reima Eresmaa and Sakari Uppala.*

## 2.1 Long-term trends in precipitation

A long-term trend of increasing precipitation has been identified in ERA-40 reanalyses. This is also seen in ECMWF's operational system, in the Met Office operational system, and to a lesser extent at NCEP. All three centres now seem to have global precipitation rates higher than indicated by the most reliable climatologies. This is an issue that requires investigation:

- **Is there a real trend in the atmosphere?** The rainfall climatologies should be checked as well as time-series of accurate observations at different locations, such as the tropics, for validation. To be done in collaboration between ERA-40, GPCP and climate groups.

- **Is there a trend in the satellite data?** We recommend that the satellite community should calibrate and cross-calibrate their data to ensure that artificial trends are removed. The data needs to be checked to see if there is a volcanic-ash or forest-fire effect that needs to be accounted for in the observation operators and the forecast model.

- **Is there a problem with feedback in the circulation?** The analyses should be checked to see if there is a trend in the vertical motion and both branches of the Hadley circulation (i.e. ascent and subsidence regions) should be checked separately.

## 2.2 Investigations of spin-up/spin-down of precipitation

There is a marked spin-down of tropical precipitation in ECMWF's current operational system and in some periods of its 40-year re-analysis (ERA-40), at the start of forecasts and with forecast range. Different centres see different effects of similar type. The spin-up/spin-down is sensitive to initialization, e.g. nudging of analysis increments (such as IAU) which reduces excessive convection at the start of forecast in the Met Office 3D-VAR system. Spin-down at ECMWF is reduced when satellite humidity from HIRS and SSM/I are removed, and it tends to be less in 4D-Var than in 3D-Var. The spin-up/spin-down effects are sensitive to assimilation and model changes. ECMWF have spin-up of large-scale precipitation and spin-down of convective precipitation. The Met Office sees the same if nudging initialization is not used. In the ERA-40 assimilation large-scale precipitation exhibits a non-negligible spin-up during the first 24 hours of the forecasts, particularly in the mid-latitude baroclinic storm-track areas. This is particular to 3D-Var and may be connected to the intermittent nature of the data assimilation. In the long-term, biases in the model and data need to be removed but in the short-term empirical tuning may be required to reduce the problem. Points to investigate include:

- Initialization of cloud water to be consistent with analysed humidity and the model's cloud parameterisations

- Ensuring better vertical distribution of humidity in the analysis and the appropriate stabilisation of profiles, by introducing temperature changes that are consistent with the humidity analysis - including techniques to modify boundary layer depth.

- Ensuring better vertical distribution of humidity in the analysis and the appropriate stabilisation of profiles, by introducing temperature changes that are consistent with the humidity analysis - including techniques to modify boundary layer depth.

- Including estimation of bias parameters within the variational scheme

- Including treatment of model error in 4D-VAR to remove the mismatch between data and the model. There is even a possibility to provide information on what aspects of the physics the mismatch is sensitive to if errors are defined in terms of tuneable coefficients in the parameterisations through introduction of extra control variables.

- Field studies (e.g. ARM and GATE) should be used to get better information on the vertical profiles of temperature and humidity in problem areas where there is a large mismatch between model and observations of total-column water vapour. Analysis, background and longer-range forecasts should be studied, as they are likely to have different characteristics, in this respect.

- Errors in the model forecasts and climatology need to be reduced to make better use of the observations. For example errors in the diurnal cycle of convection are bound to cause difficulties in assimilation of cloud and precipitation data. Infrequent updating in radiation can be a contributing factor.

## 2.3 Accuracy of radiosonde humidity data

To get the most out of the existing radiosonde network requires access to the original high-quality measurements. Much valuable information is currently lost in conversion to the restrictive format currently used for distribution of the data on the GTS. **It is recommended:** that BUFR-coded reports should become the standard reporting method, as soon as practically possible. The BUFR-coded reports are required to include information on sonde type. It should contain the relative humidity data as directly measured with super-saturation retained, at more levels in the vertical with the horizontal position and actual time recorded.

The night-time sonde reports generally have better quality than day-time soundings. The most accurate sonde types are Vaisala's RS90 and RS80H. Data from these Vaisala sondes can reliably be used up to 150 hPa. However, stratospheric humidity cannot be trusted. The Japanese radiosonde system should be of similar quality. The Suppican sondes are not reliable at low humidity and Russian, Chinese and Indian sondes can be used to indicate cloud layers, but otherwise are not reliable. Improved statistics are needed to monitor the quality of all sonde types.

**It is recommended:** that NWP centres review their usage of radiosonde humidity data, to reflect the known properties and error characteristics of the different sonde types used around the world.

**It is recommended:** that a workshop should be organised, to specifically discuss radiosonde characteristics and how to correct errors or allow for them in data assimilation.

**It is recommended:** that routine monitoring of sonde humidity should be instigated, ideally under the auspices of the WMO (similar to what is already in place for temperature and wind measurements). This will provide feedback on the accuracy of humidity measurements to the data providers. Currently systematic monitoring of humidity data is not performed. Systematic biases need to be monitored and this is best done by the data assimilation centres rather than the observation providers. These need to be stratified by sonde type.

**It is recommended:** that (e.g. within the CIMO framework) an expert should spend time at ECMWF to work with the current observation data set to develop methods providing informative monitoring statistics and common formats for exchange of the monitoring results between NWP centres.

## 2.4 Absolute calibration standards

The humidity biases in the ECMWF forecast model (as verified against in-situ data from field campaigns) and also biases between the main observing systems are generally within 5% relative humidity, i.e. within the absolute calibration accuracy of radiosondes. To make further progress in parameterisation of moist processes and in the bias correction of humidity data, data with higher absolute calibration accuracy are required.

It has been suggested that ground-based GPS might provide a useful standard for the absolute calibration of instruments measuring total column water vapour. This is because the cost is relatively low, and a very good accuracy can be expected. In the tropics, for instance, errors in the GPS retrieval of the total-column water vapour could be as low as 1 kg/m2. If such accurate data were available in larger numbers then these would be immensely useful for calibration of SSM/I data, and for investigation of the spin-down of tropical precipitation in several NWP models.

From commercial aircraft, good quality in-route and ascent/descent measurements of relative humidity are expected, in particular with the Water Vapor Sensing System-2 instrument, under development. Such data are likely to become very valuable to NWP applications. The aircraft data are relatively cheap and the measurements are expected to be of high quality. **It is recommended:** that ECMWF stresses the likely importance of aircraft humidity data in support of the ongoing initiatives undertaken by NOAA and the WMO AMDAR-Panel.

Ground-based infrared and microwave radiometer (vertically oriented) have been found to be an extremely important source of humidity data. Among others this is due to the high sampling rate (many times per hour) which allows, for example, to address the representativeness problem, and diurnal variations.

**It is recommended:** that for the time being any bias tuning based on radiosonde humidity data should better be done using night-time data, only. This is because of better accuracy of relative humidity observations during the night.

**It is a recommendation to the WMO:** that the requirement for absolute error in relative humidity from radiosondes should be tightened to 2% (rather than the 5%, currently). Moreover, the working group recommends that radiosonde reports in the future provide measurements of total-column water vapour, in addition to the profile data.

## 2.5 Main inaccuracies and limitations in 4D-Var observation operators

Fast radiative transfer models have now reached a very high level of accuracy for both temperature and humidity. The largest errors in modelling observations from current and future sensors lie in errors in the underlying spectroscopy rather than errors in the fast fit to results of line-by-line models. The accuracy of the underlying spectroscopy in the infrared requires further improvement and is being investigated elsewhere.

The main sources of error of relevance for the 4D-Var assimilation of humidity information from radiance measurements are thought to be in the areas of cloud detection, in the modelling of the diurnal cycle of convection, and in data representativeness.

**1) Cloud detection.** Any residual cloud signal in assimilated IR data (e.g. due to failure of cloud detection) will appear as a cold bias in brightness temperatures, corresponding to a wet bias in humidity. In the microwave, influences of ice-particles and precipitation can contribute significant error.

**2) Modelling of the diurnal cycle**. Current forecast models do not represent accurately the diurnal variations in convection. This may introduce significant biases in humidity that vary during the day. Such biases are detectable with ground-based GPS, and from geostationary platforms. Improvements to the convection scheme, more frequent calculation of radiation, and higher resolution radiation calculation may be required, in the physics of the model.

As convection is inherently less predictable than the dynamics at mid-latitudes (for example), and more non-linear, the length of the assimilation window may need to be re-evaluated (shortened).

**3) Representativeness errors.**

- In radiosondes the representativeness error is thought to be small, about 5% over 40km, as the extra-tropical atmosphere to a large extent is layered. Extreme situations of large representativeness problems exist, but these are of little significance for synoptic-scale meteorology, and monitoring at regional scales.

- The current practice of using radiances only in clear areas is likely to be a significant source of analysis bias. Progress with cloudy radiances will alleviate this important problem.

- The representativeness of space-based GPS is uncertain at present. It depends on the accuracy of ray-tracing methods, and its ability to produce the appropriate horizontal and vertical averaging of the occultation information.

- The actual information-content in other surface-based observing systems require further investigation (ground-based GPS, radiometers, SYNOP).

- Representativeness error is large with respect to sub-gridscale islands – both for surface- and space-based observations.

**Recommendations:**

- Investigate the consequences of the current one-sided (clear-only) use of radiances, and its contribution to bias and spin-down.

- The representativeness of surface-based observing systems require further investigation (ground-based GPS, radiometers, SYNOP 2m dew-point).

- Seek improvements to the model's representation of the diurnal cycle of convection.

- Determine the appropriate length of the 4D-Var assimilation window given the non-linear interactions between humidity and the dynamic variables.

# 3. Working group 3: Assimilation of clouds and precipitation

*Present: Luc Fillion (Chair), Jean-Noël Thépaut (Secretary), Russ Treadon, Dongsoo Kim, Veronique Ducrocq, Godelieve Deblonde, Peter Bauer, Philippe Lopez, Frederic Chevallier, Marta Janiskova, Martin Sharpe, Adrian Tompkins and Emmanuel Moreau*

## 3.1 Cloud/Rainrate vs. radiance assimilation

Two methods for rain-rate assimilation have been studied at ECMWF (papers by Marecal and Mahfouf). In a first set of experiments, 1D-Var retrievals of TCWV were produced (from humidity profiles that produce model rain in better agreement with the observed rain-rate), followed by 4D-Var assimilation of the retrieved TCWV. In a second set of experiments the observed rain-rates were used directly in 4D-Var, using a linearized convection parametrisation as observation operator. The experiments identified several difficulties with the 4D-Var approach (in its current form). These were (1) different representation of specific humidity in inner and outer loops, (i.e. spherical harmonics versus grid-point representation), (2) validity of linearity with respect to convection over 12 hours, (3) difference between cloud schemes (prognostic vs. diagnostic) used in inner and outer loops in the 4D-Var at the time.

A third and a fourth method can be similarly defined, using rain-affected microwave radiances rather than rain-rates. The advantages of radiance assimilation are (1) the easier definition of observation errors, (2) the larger flexibility with respect to channel and sensor usage, and (3) the ability to generate rain where the model first-guess has no rain.

For non-precipitating clouds, the issue of cloud layering and the assumptions of cloud overlap has to be investigated, because infrared and microwave sensors provide little information on cloud structure (except cloud top height). Additional difficulties for non-precipitating clouds using the 1D+4D approach may arise because of lack of feedback with the dynamics.

## 3.2 Recommendations:

- 1D-Var+4D-Var vs. 4D-Var has to be tested in the context of radiance assimilation in presence of precipitation.
- It is recommended to unify radiative transfer for clear-sky, clouds and precipitation.
- Revisit the incremental strategy in the context of 4D-Var assimilation of precipitation.
- A priori, it is felt that similar approaches should be followed for cloudy radiance assimilation as for rainy radiances.

## 3.3 Merits of total water assimilation

The use of total water ($q_{tot}$) as control variable in 4D-Var, has been advocated at the Met Office. One advantage is that it saves on computational resources compared to adding extra cloud variables in the model. The use of total water introduces memory of cloud properties during the assimilation window, which cannot be ensured in a formulation based on specific humidity and a diagnostic cloud scheme. The modelling of background errors for total water is likely to be of similar difficulty as for specific humidity, as the specific-humidity component of $q_{tot}$ is likely to dominate the total variance, at least in the lower atmosphere.

Treatment of super-saturation for ice remains problematic once observations that are sensitive to ice are to be used. Physical parameterisations dedicated to assimilation might be easier to derive using an integrated quantity such as total water. The error statistics of $q_{tot}$ are thought to be more Gaussian close to saturation than those for q. An intermediate alternative might be to separate the ice phase from the total, and treat it separately.

**Recommendations:**

- Experiments are required to investigate the properties of the inversion from total-water to q and T (specific humidity and temperature).
- The Met Office approach should be looked at very closely.
- ECMWF's new approach for the humidity-analysis control-variable should be extended to total-water, and its properties investigated.

## 3.4 Required properties of cloud and convection scheme for TL/AD

Most schemes have upper bounds of about 0.1 K and 0.1 g/kg for T and q-perturbations up to which the linearity assumption holds. Bounds of 1K and 1g/kg would be more ideal as these higher values correspond to the amplitude of typical 4D-Var analysis increments. However, these higher bounds might not be achievable. Issues are the strong sensitivity of parameterised precipitation to humidity at lower levels, which in the full non-linear model is likely to be compensated by the mixing effects of vertical diffusion. It should be stressed that the single-column sensitivity is crucial for the performance of the retrieval in 1D-Var.

**Recommendations:**

- Physical parameterisation schemes should be developed in dialogue between physics and assimilation community.
- The issue of cloud/rain triggering needs to be further studied.

## 3.5 Benefit of cloud top pressure/temperature

Relating model cloud profiles with retrieved cloud-top information from satellites is difficult. Estimation of these quantities is interesting for validation purposes. Exact cloud top location cannot be retrieved easily from current infrared sounders and microwave imagers/sounders. Therefore a benefit from including cloud top information in a radiance assimilation framework is not evident.

## 3.6 Nudging techniques

Nudging represents an alternative if the variational approach does not bring all the expected benefits. Important efforts are undertaken at Met Office and Météo-France in the context of mesoscale assimilation with good results. Including time information seems crucial for the choice of methodology, i.e. nudging or 4D-Var rather than 3D-Var. JMA has demonstrated improvement of precipitation forecast skill with 4D-Var, hence, nudging does not appear a priority for ECMWF.

**Recommendation:**

- The development of nudging techniques at ECMWF should not be a priority at the current time.

## 3.7 Issues of "low predictability" of cloud and precipitation and its impact on assimilation

It is a generally accepted fact that cloud properties become slaved to the dynamical fields after 6-12 hours. This effectively reduces the choice of control variables to be moisture, temperature and wind rather than cloud liquid water (or rain rate itself). Investigations on this issue in a 1D-Var context have shown no advantage of including cloud variables and hydrometeor parameters as control variables. For the very short range (3-6 hours), the benefit could certainly be larger. Definition of background errors and correlation between hydrometeor types would also be quite difficult.

**Recommendations:**

- The primary control vector elements should be moisture (or total water, see section 3.3), temperature, and the dynamical variables

## 3.8 Non-Gaussian error distributions

It is expected that the Gaussian assumption for errors in rainy and cloudy radiances may turn out to be inadequate, as errors in the forward model can be multi-modal, or skewed. For the background errors it is expected that the modelling of its non-Gaussian properties near saturation is of prime importance.

**Recommendations:**

- Continue the validation of forward model operators in cloud and precip to better quantify the associated errors

## 3.9 Importance of $J_b$ at higher resolutions

As long as there are interactions between different variables, at scales that are not fully observed, Jb remains a fundamental component of the assimilation system. Flow dependent and sharp structure functions are required when going towards a higher-resolution system.

**Recommendations:**

- Continue the work on gradually incorporating non-homogeneous, flow-dependent structure in the background error covariance matrix (i.e. pursue the wavelet Jb formulation). Investigate whether moist singular vectors can provide indication on the structure of humidity background error correlations.

- Improve the variance description for saturated regimes