ESA CONTRACT REPORT

Contract Report to the European Space Agency

Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction

WP-1000 report: Preliminary analysis and planning

M. Janisková and M. Fielding

ESA ESTEC contract 4000116891/16/NL/LvH

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF ESA Project Report Series

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

Contract Report to the European Space Agency

Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction

WP-1000 report: Preliminary analysis and planning

Authors: M. Janisková and M. Fielding

ESA ESTEC contract 4000116891/16/NL/LvH

August 2016

<table>
<tr><td colspan="4">ESA STUDY CONTRACT REPORT</td></tr>
<tr><td>ESA CONTRACT No. 4000116891/16/NL/LvH</td><td colspan="2">SUBJECT
Operational Assimilation of Space-borne Radar and Lidar Cloud Profile Observations for Numerical Weather Prediction</td><td>CONTRACTOR
European Centre for Medium-Range Weather Forecasts, UK</td></tr>
<tr><td>ESA CR( ) No:
................</td><td>STAR CODE
..................</td><td>No. of volumes: 1</td><td>CONTRACTOR'S REFERENCE: none</td></tr>
<tr><td colspan="4">ABSTRACT This report provides an overview of the Numerical Weather Prediction (NWP) system components that are crucial for assimilating radar and lidar observations of clouds in pre-operational set-up at ECMWF. The existing data assimilation system is outlined together with a more detailed plan for its extension towards the target configuration to be accomplished at the end of the project. The project plan accounts for the initial technical and scientific stages to be followed by testing that will allow conclusions to be drawn from the technical and scientific performance of the system.</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td colspan="4">NAMES OF AUTHORS: Marta Janisková and Mark Fielding

Responsible for the study: Marta Janisková</td></tr>
<tr><td colspan="2">NAME OF ESA STUDY MANAGER: Tobias Wehr
DEPARTMENT: Science and Applications (EOP-S)
DIRECTORATE: Earth Observation Programmes</td><td colspan="2">ESA BUDGET HEADING
................................</td></tr>
<tr><td colspan="4">APPROVED: ...</td></tr>
<tr><td colspan="4">AUTHORIZED: ...</td></tr>
</table>

Contents

1 Introduction

Numerical weather prediction (NWP) models have improved considerably over the past few years in the forecast of clouds thanks to progress in parametrizations. However, there is still need to explore new possibilities for model improvement through assimilation of data related to clouds from active and passive sensors. Observations providing vertical information on clouds from space-borne active instruments on board of CloudSat (Stephens et al., 2002) and CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations Winker et al., 2009) are already available and new ones, such as EarthCARE (Earth, Clouds, Aerosols and Radiation Explorer) should appear in the near future. This opens new possibilities for improvements of the atmospheric initial state and the model performance itself to be explored through assimilation of these data related to clouds.

The EarthCARE mission has the basic objective of improving the understanding of cloud-aerosol-radiation interactions by simultaneously measuring the vertical structure and the horizontal distribution of cloud and aerosol fields together with the outgoing radiation over all climate zones. The vertically resolved characterization of clouds will be provided by the combination of lidar (Atmospheric Lidar, ATLID) and a cloud profiling radar (CPR) as described in the ESA report (ESA, 2004).

A number of studies, including the ESA funded project Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles (QuARL Janisková et al., 2010), have shown that such observations are useful not only to evaluate the performance of current NWP models in representing clouds, precipitation and aerosols, but they have also a potential to be assimilated into these models to improve their initial atmospheric state.

In order to prepare for the exploitation of radar and lidar observations in data assimilation in the frame of the EarthCARE mission, the aim of the subsequent study funded by ESA (STSE Study EarthCARE Assimilation, Janisková et al., 2014) was the development of an off-line system to monitor/assimilate space-borne radar and lidar observations in clouds within the NWP model of the European Centre for Medium-Range Weather Forecasts (ECMWF). Although all developments and testing had to be done using CloudSat and CALIPSO data, since these are currently the only available real data from space-borne radar and lidar with global coverage, the work performed during the project paved the way towards the use of similar data from the future EarthCARE mission. The studies using a technique combining one-dimensional variational (1D-Var) assimilation with four-dimensional variational (4D-Var) data assimilation provided indications on the potential that assimilation of cloud information from active sensors could offer. The 1D+4D-Var approach used required to define errors for the pseudo-observations retrieved from 1D-Var by computing them from the 1D-Var analysis covariance matrix. This is quite expensive for profiling observations and not affordable for operational applications. For any future operational implementation, the use of a direct 4D-Var assimilation of cloud related observations would be much preferable. An important step before any observations can be assimilated into 4D-Var is a quality monitoring for these observations using a global NWP model.

The current project will focus on developments towards direct assimilation and monitoring systems to exploit cloud radar and lidar data for their assimilation in NWP models. This will require adjustment of assimilation tools developed during the STSE Study EarthCARE Assimilation (Janisková et al., 2014), such as observation operator, observation error definitions (namely, representativeness and forward operator errors), quality control, data screening and bias correction. The direct (in-line) data assimilation and monitoring systems developed during this project would allow extended research studies beneficial for future applications of EarthCARE ATLID and CPR data once available on the global scale.

An overview of the current knowledge and experience with assimilation of cloud radar and lidar observations is provided in Section 2. System components required for assimilation of these observations are summarized in Section 3. A more detail description of the assimilation system to be used for the project is also provided there.

Section 4 concentrates on detailed work plan for the project. A brief summary is in Section 5.

2 Overview of assimilation of cloud radar and lidar observations

A large part of forecasting deficiencies is connected with the imperfect assimilation of available observational data in the NWP process. Accurate initial conditions required by NWP models rely on the quality of observations and the quality of assimilating schemes. The potential increase of forecast skill through a more accurate treatment of clouds and their vertical structure can therefore be enormous.

2.1 General experience with cloud assimilation in NWP

In global data assimilation systems, despite the major influence of clouds and precipitation on the atmospheric water and energy balance, there is still no explicit analysis of clouds; cloud contributions to satellite radiances are mostly removed from assimilation systems. In mesoscale models, cloud analyses based on nudging techniques have been introduced (e.g. Macpherson et al., 1996; Lipton and Modica, 1999; Bayler et al., 2000). The use of cloudy observations has been explored in several experimental assimilation studies. For example, Vukićević et al. (2004) attempted to make use of visible and infrared cloudy satellite radiances in 4D-Var in a mesoscale model. On the global scale, the capability of 4D-Var assimilation systems to assimilate cloud affected satellite infrared radiances using observations from the narrow-band AIRS (Advanced Infrared Sounder) has been explored by Chevallier et al. (2004). Some experiments have been performed by Benedetti and Janisková (2008) assimilating optical depths retrieved from MODIS (Moderate Resolution Imaging Spectroradiometer) in the ECMWF 4D-Var system. The use of cloud retrievals from radar data has been investigated in 1D-Var assimilation studies (Janisková et al., 2002; Benedetti et al., 2003a,b; Benedetti and Janisková, 2004). Lopez et al. (2006) performed experimental assimilation of the ARM (Atmospheric Radiation Measurement programme) cloud radar observations combined with the ground-based precipitation measurements and GPS total column water vapour retrievals using a 2D-Var technique (the second dimension being time). An experimental procedure to assimilate cloud fraction from CloudSat cloud profiling radar was used by Storto and Tveter (2009) in the limited area 3D-Var system through the use of humidity pseudo-observations derived from a one-dimensional Bayesian analysis. A technique combining 1D-Var with 4D-Var data assimilation was experimentally applied in the first assimilation attempts with space-borne cloud radar observations (Janisková et al., 2012). More recently, this experimentation has been extended to the combination of cloud radar and lidar observations (Janisková, 2015).

2.2 Outcomes from the previous ESA studies on assimilation of cloud radar and lidar observations

2.2.1 QuARL: ESA-ESTEC Contract 21613/08/NL/CB

The aim of the QuARL project (Janisková et al., 2010) was to explore ways of exploiting the new wealth of information about the vertical distribution of clouds and aerosols obtained from the new space-borne radar and lidar instruments. The work during the project included the adaptation and developments of observation operators for cloud radar and aerosol lidar observations. An important part of the project was also the development of a method for estimating the representativity error, which can be a dominant source due to the small footprint of the active satellite measurements. Verification studies have been performed for both clouds and aerosols using the CloudSat and CALIPSO measurements. These new data types have been also used for validations of cloud detection and cloud analysis from the passive instruments. The last part of the project focused on assimilation experimentation with the cloud radar observations using the 1D+4D-Var technique and on a demonstration of monitoring for CloudSat observations.

The performed studies have shown a number of areas in NWP system that can benefit from the space-borne radar and lidar observations. The comparisons between data from the ECMWF model and observations from the new instruments helped to identify some apparent weaknesses of the forecast model. This has given guidance for model developments and, as a consequence, some of the discrepancies between model and observations could already be reduced. Thus project studies indicated that there is a significant potential that the global datasets with high resolution vertical profiles of aerosol and cloud-related information (such as that from CloudSat/CALIPSO and EarthCARE) will provide an invaluable source of data to inspire and validate model parametrization schemes.

Generally, the work on assimilating the new observations has shown great potential. While the improvements found in the 1D+4D-Var assimilations performed for this project were relatively small, strong benefits of the new data was demonstrated in the context of the 1D-Var retrievals. The successful use of the new data by the 1D-Var system prepared the basis for future progress and showed that these data can be correctly interpreted in the form of model relevant parameters. The study has indicated that to achieve the impact of such information on the whole data assimilation system (as tested in the 1D+4D-Var studies) will still require improvements in several areas, such as building a suitable bias correction for these data, better screening of observations and improved observation error definition (including representativity error).

Results from monitoring studies suggested that potential problems with cloud radar observations can be identified when first guess departures are brought outside their typical range of variation in the time trends. This could provide a basis for a future alert system.

The work carried out for this project has not only demonstrated the usefulness of the new data types (through validation and data assimilation studies), but it has also helped in laying the technical and conceptual foundations for their future exploitation.

2.2.2 STSE Study EarthCARE Assimilation: ESA-ESTEC Contract 4000102816/11/NL/CT

The main objective of STSE Study - EarthCARE Assimilation (Janisková et al., 2014) has been the development of off-line data assimilation and monitoring system to exploit the potential of space-borne radar and lidar cloud observations for NWP models. The project focused mainly on improving the radar observation operator, the development of a lidar forward operator as well as tools for handling observations such as data selection (i.e quality control and data screening) and bias correction. Observation errors such as instrument, forward modelling and representativity (due to the narrow field of view of the space-borne lidar and radar instruments) have been estimated as well. Great efforts have also been put on the development of the off-line systems for monitoring and assimilation of both cloud radar and lidar data.

A basic framework for monitoring time series of cloud radar and lidar observations has been established based on time series of cloud observations from the CloudSat radar, the CALIPSO lidar and the corresponding reflectivity, respective backscatter first guess (FG) departures simulated from the ECMWF model. Experiments have been performed to assess the ability of the monitoring system to detect potential problems in the quality of observations.

In order to investigate the potential that assimilation of cloud information from active sensors could have for NWP models, feasibility studies have then been performed where pseudo-observations of temperature and specific humidity retrieved from 1D-Var were assimilated in the ECMWF 4D-Var system to study the impact of the new observations on analyses (i.e. initial conditions for NWP models) and subsequent forecast. Outcomes from assimilation experiments have shown that 1D-Var analyses get closer to assimilated and also independent observations. The performed 1D+4D-Var assimilation experiments have indicated a positive impact of the new observations on the subsequent forecast.

In the project, the feasibility of assimilating space-borne cloud radar and lidar observations has been demonstrated. The achieved results triggered the desirability to use these new type of cloud observations for assimilation. Though results were encouraging, to gain the full benefit from these observations in the operational context will require a substantial amount of work. The studies indicated that the 1D+4D-Var approach is only affordable for non-operational applications due to error definition for pseudo-observations retrieved from 1D-Var. These errors are computed from the 1D-Var analysis covariance matrix and such computation is quite expensive for profiling observations. Therefore the direct 4D-Var assimilation of cloud related observations should be considered for any future operational implementation.

This project has provided the necessary developments and experimentation prior to the future pre-operational assimilation/monitoring of EarthCARE observation.

3 System components for assimilation of cloud radar and lidar observations

3.1 Observations and model to be used for the project

All developments and testing will be done using CloudSat and CALIPSO data, since these are currently the only available real data from space-borne radar and lidar with global coverage. However, the work performed during the project will aim to prepare the use of similar data from the future EarthCARE mission. Currently we have archived one year of CloudSat and CALIPSO data for the period from August 2007 to July 2008. These datasets contain different products (see Tab. 3.1 summarizing the data products). These datasets should be sufficient for performing the proposed study.

The comprehensive Earth-system model developed at ECMWF in co-operation with Météo-France forms the basis for all our data assimilation and forecasting activities. All the main applications required are available through one computer software system called the Integrated Forecasting System (IFS). Since November 1997 the operational data assimilation system at ECMWF uses 4D-Var method (Rabier et al., 2000; Mahfouf and Rabier, 2000). This system is based on the incremental formulation proposed by Courtier et al. (1994). The model data and 4D-runs for this project will use the model version with a horizontal resolution at least of T799 spectral truncation (approximately corresponding to grid resolution of 25 km) and 137 levels in the vertical up to 80 km altitude.

CloudSat data in archive	**CALIPSO data in archive**
2B-CLDCLASS.R04	CAL_LID_L2_01kmCLay-Prov-V2-01
2B-CWC-RO.R04	CAL_LID_L2_05kmALay-Prov-V2-01
2B-FLXHR.R04	CAL_LID_L2_05kmCLay-Prov-V2-01
2B-GEOPROF.R04	CAL_LID_L2_05kmCPro-Beta-V2-01
2B-GEOPROF-LIDAR.P1.R04	CAL_LID_L2_40kmAProCal-Beta-V2-01
2B-TAU.R04	CAL_LID_L2_VFM-Prov-V2-01
MODIS-AUX.R04	

Table 3.1: List of CloudSat and CALIPSO data available for the project.

3.2 Data assimilation system at ECMWF

3.2.1 Description of the system

The operational data assimilation system at ECMWF is 4D-Var and it is based on incremental formulation. 4D-Var seeks an optimal balance between observations and the dynamics of the atmosphere by finding a model trajectory $\mathbf{x}(t)$ which is as close as possible, in a least-square sense, to the observations available during a given time period $[t_0, t_n]$. The model trajectory $\mathbf{x}(t)$ is completely defined by the initial state $\mathbf{x}_0$ at time t_0.

The misfit to given observations $\mathbf{y}^o$ and to an *a-priori* model state $\mathbf{x}^b$ called background, which is usually provided by a short-range forecast, is measured by an objective (cost) function. In addition to the background cost function $\mathscr{J}^b$ (measuring the distance between the initial state of the model $\mathbf{x}_0$ and the background $\mathbf{x}_0^b$) and the observation cost function $\mathscr{J}^o$ (measuring the distance between the model trajectory and corresponding observations), a constraint cost function $\mathscr{J}^c$ is used in 4D-Var to control fast gravity waves using the digital filter approach developed by Gauthier and Thépaut (2001).

Using the incremental approach, 4D-Var can be approximated to the first order by finding the analysis increment $\delta\mathbf{x}_0$ at initial time t_0 which minimizes the following cost function $\mathscr{J}$:

$$\mathscr{J}(\delta\mathbf{x}_0) = \underbrace{\frac{1}{2}(\delta\mathbf{x}_0)^T\mathbf{B}^{-1}(\delta\mathbf{x}_0)}_{\mathscr{J}^b} + \underbrace{\frac{1}{2}\sum_{i=0}^{n}\left(H_i'\delta\mathbf{x}_i - \mathbf{d}_i\right)^T\mathbf{R}_i^{-1}\left(H_i'\delta\mathbf{x}_i - \mathbf{d}_i\right)}_{\mathscr{J}^o} + \mathscr{J}^c \tag{3.1}$$

where at any time t_i,

- $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_i^b$ is the analysis increment and represents the departure of the model state ($\mathbf{x}$) with respect to the background ($\mathbf{x}^b$) which consists of temperature, humidity, vorticity, divergence and surface pressure in the current 4D-Var system;
- H_i' is the linearized observation operator providing the model equivalent to the observations and it also includes the spatial interpolations to observation locations as well as the propagation of the initial state to each observation time using the forecast model;
- $\mathbf{d}_i = \mathbf{y}_i^o - H_i(\mathbf{x}^b{}_i)$ is the so-called innovation vector providing the departure of the model background equivalent from the observation ($\mathbf{y}_i^o$);
- $\mathbf{R}_i$ is the observation error covariance matrix (including measurement and representativeness errors);
- $\mathbf{B}$ is the background error covariance matrix of the state $\mathbf{x}^b$ and is based on a wavelet formulation (Fisher, 2004) to introduce regime-dependent error statistics.

In the incremental approach the highest possible resolution is used for the computation of the model trajectory, and for calculating the departures between observations and model, whereas a lower-resolution model (its adjoint and tangent linear) are used for the iterative and relatively costly computation of analysis increments. The lower-resolution iterations (the inner-loops) can optionally be nested within a set of outer-loop iterations at full resolution. Apart from the resolution, the cost of the inner-loops will depend also upon the complexity of the inner-loop model, e.g. the use of simpler or more complete representations of the physical processes (Janisková and Lopez, 2013). The inner-loop resolution is increased with each iteration of the outer-loop using 'multi-resolution' extension to the incremental method (Veerse and Thépaut, 1998). Fig. 3.1 provides a schematic description of the incremental 4D-Var solution algorithm (IFS-Documentation, 2015).

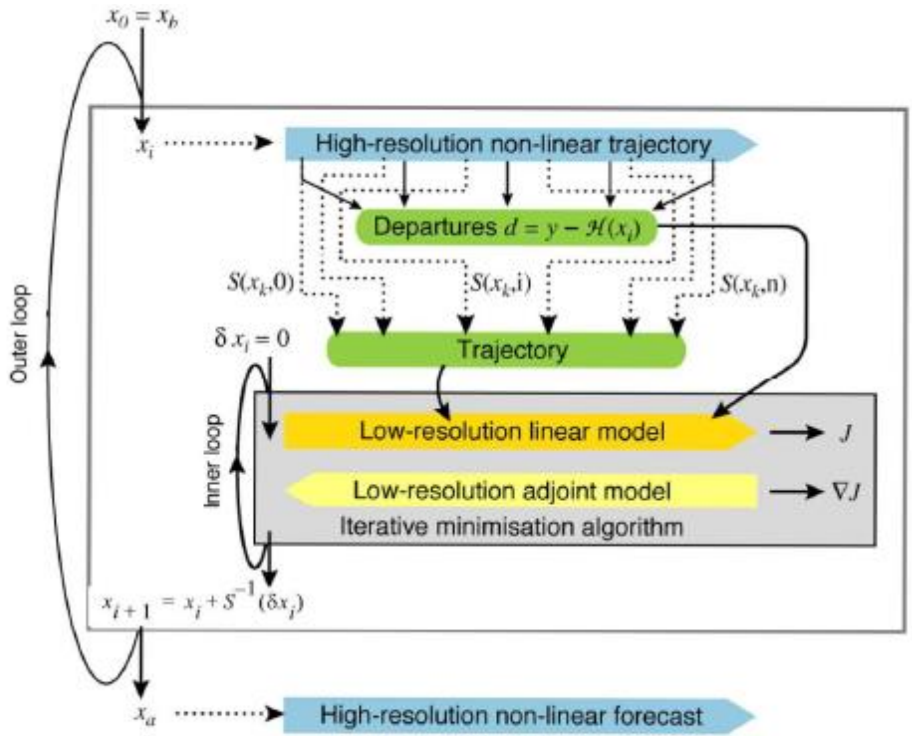

Figure 3.1: Schematic of the 4D-Var solution algorithm at ECMWF (IFS-Documentation, 2015). Outer loops are performed at high resolution using the full non-linear model. Inner iterations are performed at lower resolution (increased with each iteration) using the tangent-linear and adjoint versions of the forecast model, linearised around a 12-hour succession of model states ('the trajectory') obtained through interpolation from high resolution (S denotes the truncation operator, J the cost function and x the atmospheric state vector).

3.2.2 Particular steps of 4D-Var solution

The current length of the 4D-Var assimilation window at ECMWF is 12 hours, running from 09 UTC to 21 UTC to produce the 12 UTC analysis and forecast products, and from 21 UTC to 09 UTC for the 00 UTC production. Observations for assimilation system are organized in time-slots (currently half-hourly, i.e. each i in Eq. 3.1 corresponds to a half-hour time slot).

The incremental 4D-Var system is rather complex and it consists of several steps which are briefly summarized here:

i. Comparison of the observations with the background is done at high resolution to compute the innovation vector providing the departure of the model background equivalent to the observation. This procedure involves using appropriate observation operators which need to be properly included in the 4D-Var system.

The background departures must then be stored in the ODB (the observation database) for later use in the minimization. This job step also performs screening (i.e. blacklisting, thinning and quality control against the background) of observations. The screening determines which observations will be passed to the main minimization. Very large volumes of data are present during the screening run only, for the purpose of data monitoring.

The model trajectory is subsequently interpolated to the resolution of the next job step and written out.

ii. The first minimization at low resolution is used to produce preliminary low-resolution analysis increments, using simplified tangent-linear physics, and tangent-linear observation operators.

iii. First update of the high-resolution trajectory to take non-linear effects partly into account applies the analysis increments obtained in the first minimization and performs another forecast integration with comparison to observations. Departures from this new atmospheric state are stored in the ODB.

The analysis problem is then re-linearized around the updated model state which provides a new linearization state for the next minimization. The updated model trajectory is then interpolated to the resolution of the next job step and written out.

iv. The second main minimization is performed at increased resolution with a more complete representation of tangent-linear physics. The tangent-linear and adjoint versions of cloud radar and lidar observation operators will also need to be used in this minimization.

Steps (*iii*) and (*iv*) are repeated again in the current operational data assimilation system.

v. The production of the high-resolution analysis is finally carried out by adding the low-resolution increment to the background (at initial time), and integrating to the analysis times. The comparison of the analysis with all observations (including those not used by the analysis, for diagnostic purposes) is performed.

3.2.3 Data assimilation cycling

In June 2004, the early-delivery suite (Haseler, 2004) was introduced at ECMWF. This suite reorganises the data assimilation system in order to make the ECMWF operational products available earlier. In particular, it makes the 0000 UTC products available before the end of the working day, even in the easternmost ECMWF Member States where local summer time is three hours ahead of Universal Time.

An overview of the data flow through the data assimilation system with the operational early-delivery configuration (IFS-Documentation, 2015) is shown in Fig. 3.2. The 12-hour 4D-Var analyses are run with a delayed cut-off time, in order to use the maximum possible number of observations:

- 0000 UTC analysis using observations from the time window 2101 - 0900 UTC;
- 1200 UTC analysis using observations in the window 0901 - 2100 UTC.

The extraction tasks for observations:

- in the periods 2101 - 0300 UTC and 0301 - 0900 UTC are run at 1345 and 1400 UTC respectively;
- in the periods 0901 - 1500 UTC and 1500 - 2100 UTC are run at 0145 and 0200 UTC.

The 0000 UTC 12-hour 4D-Var analysis generates two sets of analysed fields, at 0000 and 0600 UTC. A separate surface analysis is run every 6 hours. The final analysis is a combination of the fields from 4D-Var and from the surface analysis.
The first guess:

- for the 0000 UTC 12-hour 4D-Var analysis is the three-hour forecast from the previous days 1800 UTC delayed cut-off analysis;
- for the 1200 UTC 12-hour 4D-Var delayed cut-off analysis is the three-hour forecast from the 0600 UTC analysis.

The 12-hour 4D-Var delayed cut-off analyses are actually those which propagate information forwards from day to day.

The early-delivery analyses do not pass information from cycle to cycle. Each analysis is reinitialized with the best available model fields from the delayed cut-off assimilation:

- The 0000 UTC early-delivery analysis is a 6-hour 4D-Var analysis that uses observations in the time window 2101 - 0300 UTC. The cutoff time is 0400 UTC, and any observations which arrive after this time are not used by the early-delivery analysis. However, if they arrive by 1400 UTC, they are not lost, since they are included in the delayed cut-off 12-hour 4D-Var 0000 UTC analysis. The first guess for the 0000 UTC early-delivery analysis is the three-hour forecast from the previous day's 1800 UTC delayed cut-off analysis.
- The early-delivery 1200 UTC analysis is a 6-hour 4D-Var analysis that uses observations in the time window 0901 - 1500 UTC, with a cut-off time of 1600 UTC. Its first guess is the three-hour forecast from the 0600 UTC delayed cut-off analysis.

To summarize observation usage, the 4D-Var analysis waits about 5 hours to make sure that almost all available observations have arrived in order to ensure the most comprehensive global data coverage, including southern hemisphere surface data and global satellite-sounding data.

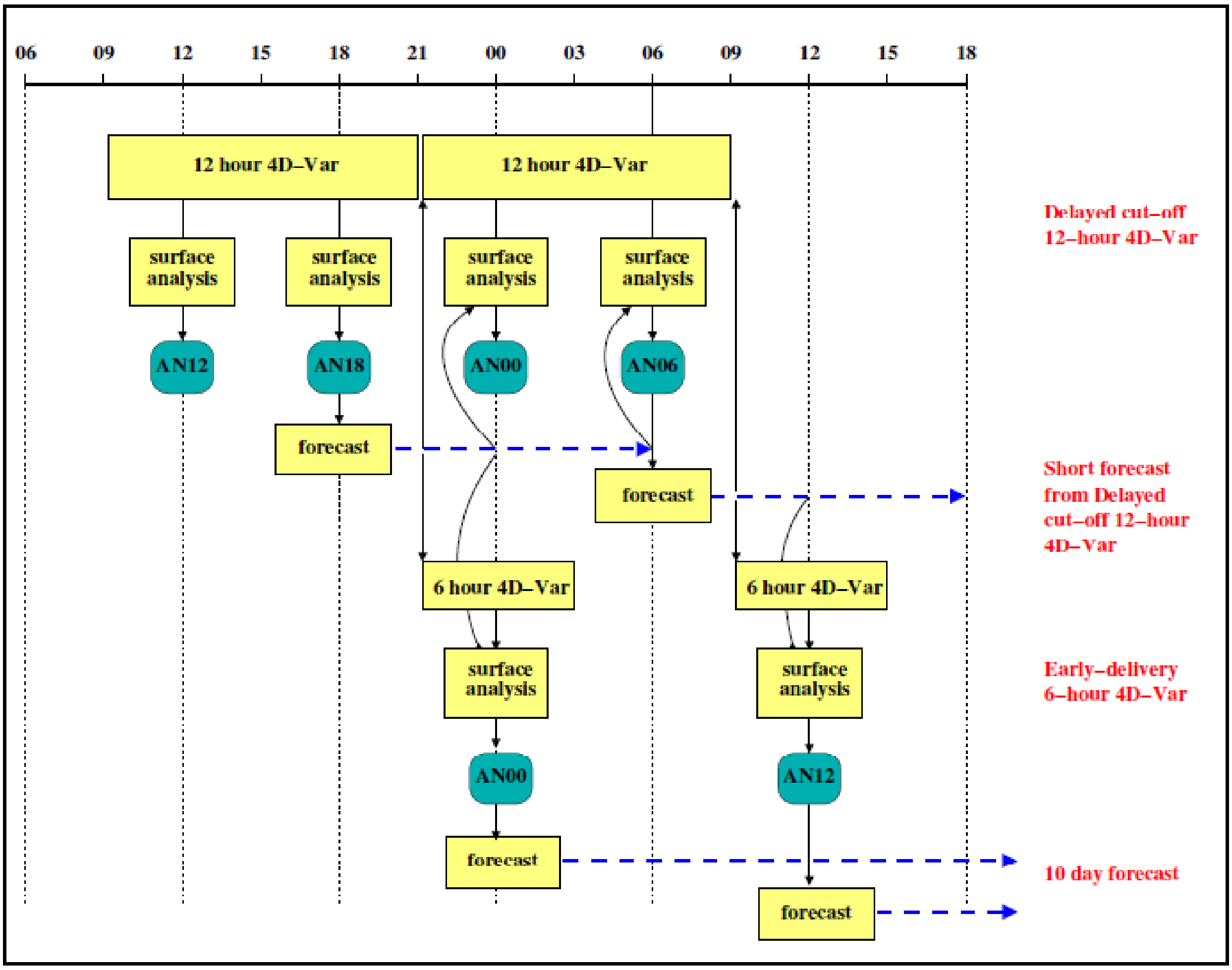

Figure 3.2: ECMWF Data assimilation cycling with the Early Delivery configuration.

4 Project plans for assimilation of space-borne cloud radar and lidar observations

The main objective of the project is to carry on developments towards an assimilation system that will be able to monitor the data quality of the EarthCARE radar and lidar profiles and assimilate them into an operational global weather prediction model. As mentioned above, data products of the CloudSat and CALIPSO missions will be used for test and demonstration cases. However, the outcome of the study and the developed tools should be applicable for EarthCARE radar and lidar products.

Exploiting EarthCARE-type data through assimilation and monitoring will not only be beneficial for the NWP community by performing the necessary developments and scientific testing, but would furthermore demonstrate the operational potential for the EarthCARE mission. The direct (in-line) data assimilation and monitoring systems developed during this project would allow extended research studies beneficial for future applications of EarthCARE lidar and cloud profiling radar data once available on the global scale. The study will be limited to cloud observations only.

4.1 Structure of the project

A flowchart of the overall system (Fig. 4.1) shows the individual components required to combine the EarthCARE observations (top row) and the model (bottom row) to enable routine EarthCARE radar and lidar observations of clouds to be assimilated in the ECMWF NWP model and to create a real-time monitoring system to detect instrument errors or degradation in the observations. The work will require to accomplish sets of the technical (pink) and the scientific tasks (blue). The work packages (with prefix WP) to which the tasks correspond, are also indicated on the flowchart.

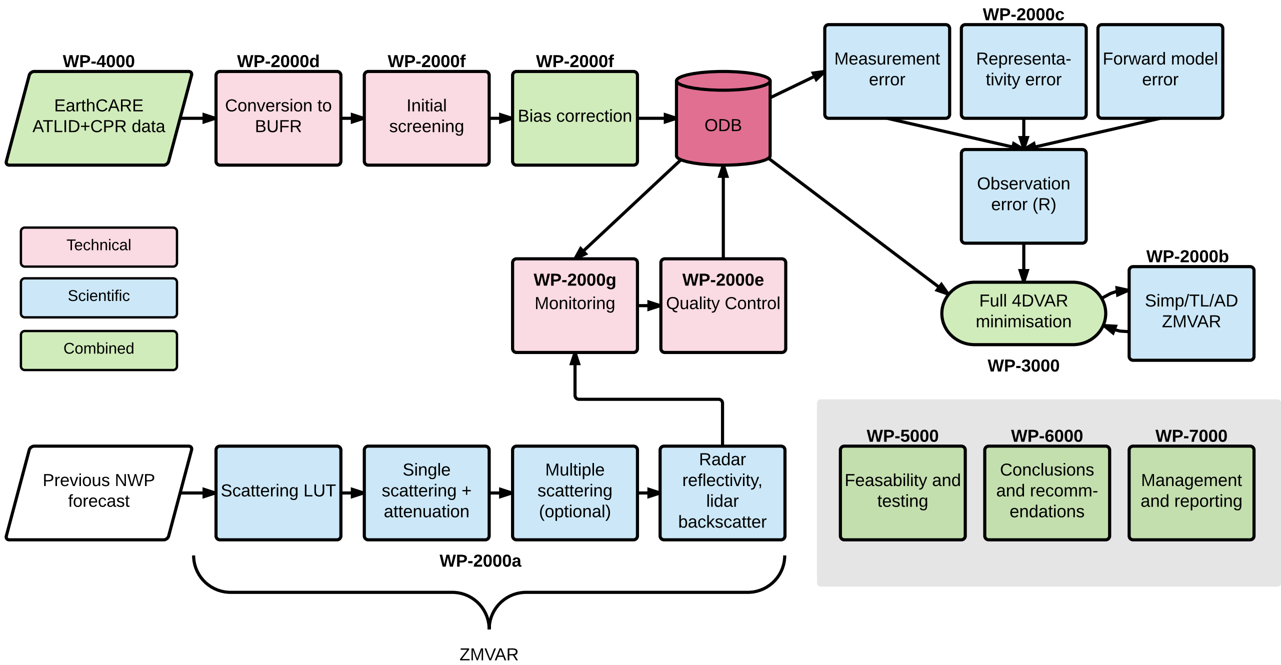

Figure 4.1: Structure of the project displaying the tasks to be done for direct (in-line) assimilation of EarthCARE cloud radar and lidar observations.

The majority of the technical aspect in WP-2000 is devoted to conversion of the EarthCARE observations into a format that is generally used by the ECMWF assimilation system. All observational data assimilated in the operational ECMWF system is processed in the same way, despite different data access mechanisms and incoming formats. Observational data is converted from the original to Binary Universal Form for Representation of meteorological data (BUFR) as prescribed by the World Meteorological Organization (WMO) standards. The output is then ingested in the Observational Data Base (ODB) that allows flexible data selection and easy association with model fields. In the post-processing step, analysis feedback information including data quality indicators is added to the database to provide a full trace of the effect of all observations in data assimilation. This capability is a necessary condition for data monitoring and impact assessment. The data handling according to this procedure will be extended to include radar and lidar observations for CloudSat/CALIPSO as well as for simulated EarthCARE observations.

The science aspect of WP-2000 covers:

- the definition of observational quality control and bias correction through a statistical analysis of CloudSat/CALIPSO data and model FG departures;
- the improvement, adaption and implementation of the existing observation operators (Z (reflectivity) model for variational assimilation, ZmVar) to match the technical specifications of the EarthCARE observations;
- the refinement of the corresponding observation errors, which consists of instrument, representativity and observation operator errors.

When the above technical and scientific developments have been completed, a significant amount of time and effort will be required to test various aspects of the updated assimilation system (WP-3000), including that:

- the new observations are correctly recognized and treated by both the model system and the ODB;
- the adjoint of the 4D-Var system is correct with respect to the new observation operator;
- the 4D-Var system converges with the new observations;
- the system is scientifically correct, i.e. checking that the system minimizes towards not only the new observations, but also towards all other observations used in the 4D-Var system.

Attempts will also be made to make improvements to the system indicated by any deficiencies discovered during the testing. Once the technical testing has been accomplished, a feasibility demonstration of the 4D-Var assimilation system will be carried out using CloudSat and CALIPSO observations (WP-5000). In particular, a few successive orbits will be used to demonstrate the system on a variety of meteorological scenes, focusing on scenes that have a higher probability of success. A validation will be performed using available independent observations.

While the feasibility and testing of the system is underway, preparations for the data handling of EarthCARE observations will run in parallel (WP-4000). This will involve identifying the exact EarthCARE data products that will be used, taking into account the data latency and reliability. A full consideration and analysis of the differences of EarthCARE data compared to CloudSat and CALIPSO will be made. A simulated EarthCARE observation dataset will be created using retrieved cloud properties along the CloudSat/CALIPSO track, which will be used to test the system for capability to assimilate EarthCARE observations.

At the end of the project, conclusions and recommendations (WP-6000 report) will be provided. They will focus on with assessing the suitability of using space-borne radar and lidar data for monitoring and assimilation in operational NWP models. The remaining work required to complete preparation for operational assimilation of the EarthCARE radar and lidar observations will be also discussed.

4.2 Main components of the project

4.2.1 Observation quality monitoring and pre-processing

During the STSE Study - EarthCARE Assimilation (Janisková et al., 2014), the 1D+4D-Var system for cloud radar and lidar assimilation using CloudSat and CALIPSO data has been prepared at ECMWF. Though the aim of this project is to develop the system for direct 4D-Var assimilation of these type of observations, the different tasks of WP-2000 will be built on already existing tools developed in the previous project. Various components requiring either adjustment, further development or to be newly built will be considered by this work package. These tools will also need to be adapted and included in the 4D-Var system.

Figure 4.2 shows the developments for WP-2000 separated into three distinct workflows:

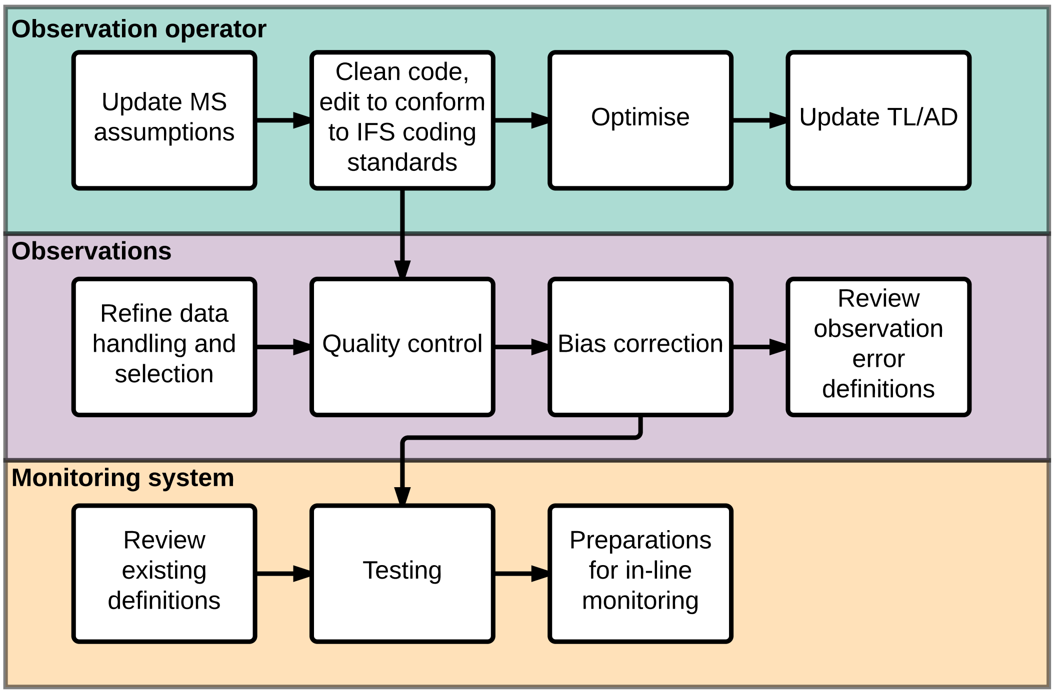

Figure 4.2: Workflow for tasks related to observation quality monitoring and pre-processing.

(a) Observation operator

The role of observation operator is to provide model equivalent to observations. The main structure of the observation operators for radar reflectivity and lidar backscatter due to clouds will likely remain unchanged from the developments in previous projects (Janisková et al., 2010, 2014). However, unlike previous off-line experiments, the computational cost of additional components to the observation operators will need to be balanced by the benefit they bring. Therefore, a review of the cost of accounting for multiple scattering for lidar (PVC method; Hogan, 2008) and radar (TDTS method; Hogan and Battaglia, 2008) will be made so that a decision can be made on their inclusion in the full 4D-Var system. A breakdown of the cost of each element in the observation operators will also be made and the most costly routines will be optimised. The code will be rewritten to conform to the strict coding standards of the IFS. The corresponding tangent linear and adjoint models will have to be updated and tested with any changes made to the forward observation operators. Finally, the observation operators with their tangent-linear and adjoint versions will need to be tested in the full 4D-Var system.

(b) Observation handling

Data handling is an important and necessary step for further testing of the assimilation/monitoring system. For observations to be included in the 4D-Var system, they must first be converted to Binary Universal Form (BUFR). Existing code will then need to be modified to process the new observations and transfer them between the ODB and the model in both directions. A review of the screening criteria for blacklisting of observations that are likely to have a negative impact on the assimilation will be made as a part of observation handling.

The data quality control system is needed to remove not only observations flagged as of poor quality (e.g. weak signal, surface contamination) or leading to excessive first guess (FG) departures (i.e. differences between observations and the model first guess) but also to avoid cases of large model uncertainties or excessive non-linear effects in the observation operators, which can lead to very large differences between observed and corresponding simulated parameters. Thus quality control not only aims at discarding erroneous data but also to avoid the data assimilation system performing sub-optimally. The strategy for quality control investigated in the previous project (Janisková et al., 2014) will be re-assessed. It also needs to be carefully implemented and tested for global applications.

Once the observations have been screened, they must also be corrected for any biases since any variational data assimilation system relies on the assumption that both observations and model background are unbiased quantities. The bias correction will be based on a year-long statistical analysis of the first guess departures with CloudSat and CALIPSO observations during 2007-2008. This will require running the latest IFS model version for a series of 12 hour forecasts on the historical analyses. Having a year-long statistical analysis will allow the bias correction to use indicators such as seasonality and hydrometeor content, as well as temperature and altitude as in Janisková et al. (2014). Preparations and necessary adjustments for including the scheme for bias correction in 4D-Var system will also be done.

Another vital component of observation processing is the definition of observation errors. Along with the background error, these errors define the weight that is placed on each observation. The most difficult to define is the representativity error (Stiller, 2010), which accounts for the active sensors much narrower field-of-view relative to the model grid box. These errors will need testing and adjusting for use in the 4D-Var system.

(c) Monitoring system

Monitoring of observations in an NWP analysis system is fundamental for understanding the consistency between model and observations. By tracking the FG departures, the routine monitoring system ensures the correctness and consistency of both the model and the observations. This procedure also provides a unique tool to routinely check for instrument deficiencies using the model as a reference point.

The system prepared by the STSE Study - EarthCARE assimilation (Janisková et al., 2014) will be revised and improved by using the year-long statistics generated for the bias correction in the observation handling section. This will allow smaller drifts or errors in instrument calibration to be detected. The system will be re-tested and preparations for in-line monitoring will be made.

4.2.2 Assimilation system developments and testing

Adjustments to the operational data assimilation system at ECMWF need to be done to account for the new observation types of cloud radar reflectivity and lidar backscatter. All modifications will subsequently need to be tested, corrected and further improved based on the results of validation tests. Figure 4.3 provides an overview of assimilation system development and testing.

a) Developments for 4D-Var assimilation system

The incremental 4D-Var system is rather complex and it consists of several steps which are briefly summarized in Section 3.2.2. For the new cloud radar and lidar observations, the following modifications and developments will be required (numbering corresponds to the steps in Section 3.2.2):

i. Computation of the innovation vector providing the departure of the model background equivalent to the observation at high resolution involves using appropriate observation operators which need to be properly included in the 4D-Var system. These background departures for the new observations to be later used in the minimization must be stored in the ODB. Screening (i.e. blacklisting, thinning and quality control against the background) of these observations to determine which observations will be passed to the main minimization will also need to be implemented. Data present during the screening run will also be usable for the purpose of data monitoring.

ii.,iv. The tangent-linear and adjoint versions of cloud radar and lidar observation operators are required to be incorporated to the system for any minimization at low resolution which is used to produce preliminary low-resolution analysis increments.

iii. The procedure that performs another forecast integration for comparison to observations and stores departures from the new atmospheric state (obtained after update by increments from minimization) will have to involve the new observations after any update of the high-resolution trajectory.

v. The final step, which is production of the high-resolution analysis and the comparison of the analysis with all observations, will also be required for the new observations.

Another development required for 4D-Var system will involve:

- passing appropriate model information needed as input to observation operators (such as cloud fraction, liquid or ice water) to the observation space, where the comparison between observations and the model background equivalents can be performed;
- the bias correction for new observations to be included in the observational term of the 4D-Var cost function to account for possible systematic errors in these observations and/or observation operators;
- observation errors for the space-borne cloud radar and lidar observations to be incorporated in the 4D-Var system.

b) Testing of the updated assimilation system

All the above development will first require technical testing to assure that the new observations are properly passed through the system, i.e. they are correctly recognized and treated by both the model system and the ODB. The correctness of adjoint of the 4D-Var system with the new observation operators will have to be tested before attempting to start with any minimization trials. Afterwards convergence of the 4D-Var system with the new observations and operators included will have to be tested.

After successful technical validation, a basic scientific assessment of the system will need to be done. This will involve testing of data assimilation cycling, i.e. checking that the system minimizes towards not only the new observations, but also towards all other observations used in the 4D-Var system. This evaluation may help to reveal some deficiencies in the applied blacklisting, quality control, bias correction or error definitions for the new observations.

Attempts will be made to make improvements to the system indicated by any deficiencies discovered during testing of the updated data assimilation system. Overall extent of these improvements may be limited by the length of the project contract and complexity of the problem to be solved.

4.2.3 Feasibility demonstration of the system

This activity will demonstrate the feasibility of assimilating EarthCARE observations and will provide an initial assessment of the benefits for NWP in preparation for the potential future operational use. This will be achieved through impact experiments executed using CloudSat and CALIPSO observations for a set of meteorological scenarios. The impact of the assimilation of CloudSat and CALIPSO data on the forecast skill shall be experimentally quantified.

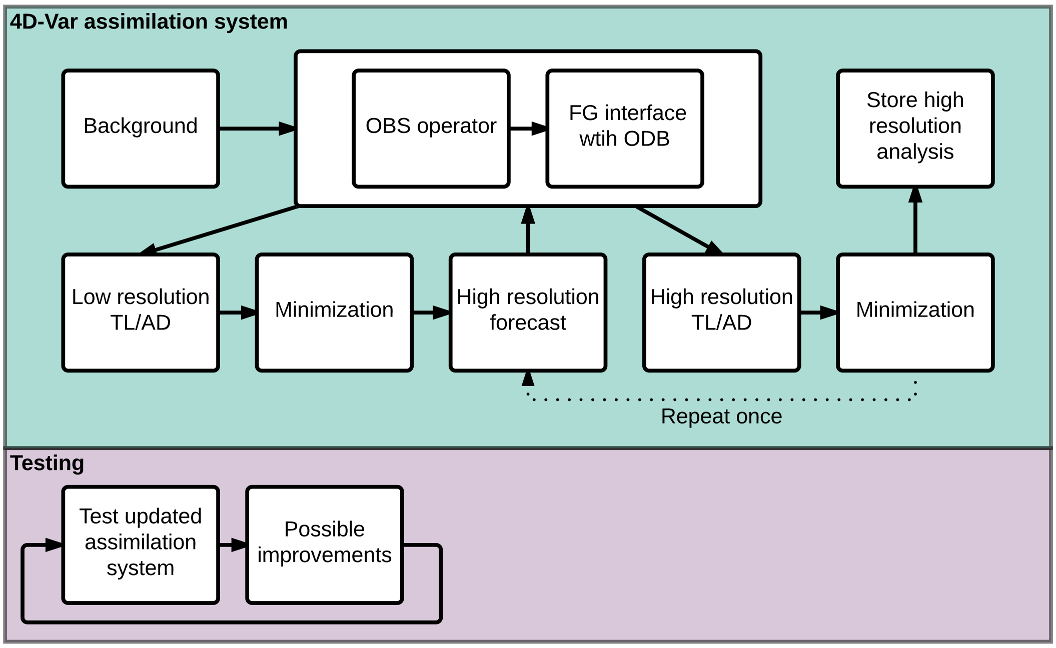

Figure 4.3: Assimilation system development and testing.

The experimental set-up will be based on experience gained during the STSE study (Janisková et al., 2014). However in contrast to that study, the observations from radar and/or lidar will be assimilated directly in 4D-Var in a similar way to the operational assimilation of all-sky radiance observations (e.g. Geer et al., 2014) at ECMWF. Geer et al. (2014) have demonstrated the feasibility of this transition from a two-step to a one step approach for radiance observations and this work package will demonstrate whether this transition is also feasible for EarthCARE radar and lidar observations.

The 4D-Var assimilation will be performed with the aim of adjusting the model temperature, specific humidity and wind profiles, as well as surface pressure in a consistent way and without degrading the fit to other observations. The proposed assimilation approach will be tested for specific situations over a limited period of time to study the impact of the new observations on 4D-Var analyses and subsequent forecasts. Due to lack of spatial coverage of cloud radar and lidar observations from space two successive orbits are not likely to cover the same cloud system so that the cloud observational constraint in 4D-Var will not be strong. Therefore the best approach to test feasibility is to use a few successive orbits and focus on the given meteorological scenes in these orbits. This is the natural approach for any future pre-operational application since, in general, assimilation systems are usually using observations without targeting specific situations. Therefore a few selected successive orbits (if possible) covering several cloud scenes will be used.

In planned experimentation, CloudSat radar reflectivities and/or CALIPSO cloud lidar backscatter coefficients (i.e. level-1 products) will be assimilated directly in 4D-Var. A control 4D-Var experiment based on the full observing system will be run for the same selected periods as the experiments, as a reference baseline. Assimilation tests shall be done for CloudSat and CALIPSO data individually, as well as simultaneously. Assimilation runs over several months may be considered but initial experimentation will concentrate on detailed studies over a few successive orbits.

Analysis increments of cloud, water vapour, temperature and wind vectors will be assessed to quantify the impact of the assimilated observations on the control variables of the 4D-Var system. This will allow a better understanding of the contribution of the integration of the dynamical model in the 4D-Var system to the temperature and moisture analysis.

A validation will be performed using available independent observations and other validation tools that are generally used for analysis and forecast performance evaluation of the ECMWF model. The evaluation of analyses is usually performed with all used observations assuming that better analyses will produce a generally and consistently better fit of the model fields when compared to observations. This comparison will be performed against all conventional and satellite observations for both analysis and first-guess, i.e. the short-range forecast that produces a first estimate of the actual state and that has been initialized from the previous analysis. This method is very stable and generally considered unambiguous, though it shows the change in the quality of the analysis, and may not necessarily imply forecast improvements. Forecast evaluation is more difficult because different reference standards (own or operational analysis) can produce rather different results.

5 Summary

The overview description of the NWP system components crucial for assimilating radar and lidar observations of clouds was provided for a pre-operational setup at ECMWF. The detail plan for extension of the existing system towards the targeted configuration to be accomplished at the end of project was outlined.

At the end of project, conclusions from the work carried out will be drawn with an aim to assess and demonstrate the suitability of using space-borne cloud radar and lidar data for assimilation and monitoring in operational NWP models. This will be based on the outcomes of performed 4D-Var assimilation test studies using CloudSat and CALIPSO cloud observations during the project.

Requirements for suggested additional technical and scientific developments needed for potential operational data assimilation and/or data quality monitoring of EarthCARE radar and lidar cloud observations will be derived taking into account our long experience from the operational assimilation and monitoring of a large variety of different satellite observations (Thépaut, 2003; Bauer et al., 2006a,b; Kelly and Thépaut, 2007; Bauer et al., 2010; Collard and McNally, 2009; McNally, 2009, and many others) and results of the performed study in this project.

Suggestions and recommendations for potential additional work to be done by the space agencies in order to refine their observation quality monitoring and also data dissemination systems to facilitate observation usage by NWP centres will also be provided.

List of Acronyms

1D-Var	One-Dimensional Variational Assimilation
4D-Var	Four-Dimensional Variational Assimilation
ARM	Atmospheric Radiation Measurements
ATLID	ATmospheric LIDar
BUFR	Binary Universal Form for the Representation of meteorological data
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CloudSat	NASA's cloud radar mission
CPR	Cloud Profiling Radar
EarthCARE	Earth, Clouds, Aerosols and Radiation Explorer
ECMWF	European Centre for Medium Range Weather Forecasts
ESA	European Space Agency
FG	First Guess
IFS	Integrated Forecasting System of ECMWF
NASA	National Aeronautics and Space Administration
NWP	Numerical Weather Prediction
ODB	Observation Data Base
PVC	Photon Variance-Covariance
STSE	Support-to-Science-Element
TDTS	Time-Dependent Two-Stream
Z	Radar reflectivity
ZmVar	Z (reflectivity) Model for Variational assimilation at ECMWF
WMO	World Meteorological Organization

References

Bauer, P., A. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation, *Q. J. R. Meteorol. Soc.*, **136**, 1868–1885, doi:10.1002/qj.659.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2006a: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. Part I: 1D-Var, *Q. J. R. Meteorol. Soc.*, **132**, 2277–2306.

Bauer, P., P. Lopez, A. Salmond, D.and Benedetti, S. Saarinen, and M. Bonazzola, 2006b: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. Part II: 4D-Var, *Q. J. R. Meteorol. Soc.*, **132**, 2307–2332.

Bayler, G. M., R. M. Aune, and W. H. Raymond, 2000: NWP cloud initialization using GOES sounder data and improved modeling of nonprecipitating clouds, *Mon. Weather Rev.*, **128**, 3911–3920.

Benedetti, A. and M. Janisková, 2004: Advances in cloud assimilation at ECMWF using ARM radar data, *Extended abstract for ICCP, Bologna*.

Benedetti, A. and M. Janisková, 2008: Assimilation of MODIS cloud optical depths in the ECMWF model, *Mon. Weather Rev.*, **136**, 1727–1746.

Benedetti, A., G. Stephens, and V. T, 2003a: Variational assimilation of radar reflectivities in a cirrus model. Part I: Model description and adjoint sensitivity studies, *Q. J. R. Meteorol. Soc.*, **129**, 277–300.

Benedetti, A., G. Stephens, and V. T, 2003b: Variational assimilation of radar reflectivities in a cirrus model. Part II: Optimal initialization and model bias estimation, *Q. J. R. Meteorol. Soc.*, **129**, 301–319.

Chevallier, F., P. Lopez, A. Tompkins, M. Janisková, and E. Moreau, 2004: The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances, *Q. J. R. Meteorol. Soc.*, **130**, 917–932.

Collard, A. D. and A. P. McNally, 2009: The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF, *Q. J. R. Meteorol. Soc.*, **135**, 1044–1058.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, **120**, 1367–1387.

ESA, 2004: EarthCARE - Earth Clouds, Aerosol and Radiation Explorer, Reports for mission selection, the sixcandidate Earth explorer missions, ESA SP-1279, ESA Publications Division c/o ESTEC, Noordwijk, The Netherlands.

Fisher, M., 2004: Generalized frames on the sphere, with application to the background error covariance modelling, *Proc. Seminar on Recent Developments in Numerical Methods for Atmospheric and Ocean Modelling, Reading, UK, ECMWF*, pp. 87–102.

Gauthier, P. and J.-N. Thépaut, 2001: Impact of the digital filters as a weak constraint in the preoperational 4D-Var assimilation system of Météo-France, *Mon. Weather Rev.*, **129**, 2089–2102.

Geer, A. J., F. Baordo, N. Bormann, and E. S. J., 2014: All-sky assimilation of microwave humidity sounders, ECMWF Technical Memorandum 741, 57 pp.

Haseler, J., 2004: Early-delivery suite, ECMWF Technical Memorandum 454, 37 pp.

Hogan, R. J., 2008: Fast lidar and radar multiple-scattering models. Part I: Small-angle scattering using the photon variancecovariance method, *J. Atmos. Sci.*, **65 (12)**, 3621–3635.

Hogan, R. J. and A. Battaglia, 2008: Fast lidar and radar multiple-scattering models. Part II: Wide-angle scattering using the time-dependent two-stream approximation, *J. Atmos. Sci.*, **65 (12)**, 3635–3651.

IFS-Documentation, 2015: - CY41r1: Part II: Data assimilation, http://www.ecmwf.int/sites/default/files/elibrary/2015/9209-part-ii-data-assimilation.pdf, 103 pp.

Janisková, M., 2015: Assimilation of cloud information from space-borne radar and lidar: Experimental study using 1D+4D-Var technique, *Q. J. R. Meteorol. Soc.*, **141**, 27082725, doi:10.1002/qj.2558.

Janisková, M., S. Di Michele, and E. Martins, 2014: Support-to-Science-Elements (STSE) Study - EarthCARE Assimilation, ESA Contract Report on Project 4000102816/11/NL/CT, 225 pp.

Janisková, M. and P. Lopez, 2013: Linearized physics for data assimilation at ECMWF, in *S.K. Park and L. Xu (Eds), Data Assimilation for Atmospheric, Ocean and Hydrological Applications (Vol II), Springer-Verlag Berline Heidelberg*, pp. 251–286, doi:10.1007/978–3–642–35088–7–11.

Janisková, M., P. Lopez, and P. Bauer, 2012: Experimental 1D+4D-Var assimilation of CloudSat observations, *Q. J. R. Meteorol. Soc.*, **138**, 1196–1220, doi:10.1002/qj.988.

Janisková, M., J.-F. Mahfouf, and J.-J. Morcrette, 2002: Preliminary studies on the variational assimilation of cloud–radiation observations, *Q. J. R. Meteorol. Soc.*, **128**, 2713–2736.

Janisková, M., O. Stiller, S. Di Michele, R. Forbes, J.-J. Morcrette, M. Ahlgrimm, P. Bauer, and L. Jones, 2010: QuARL - Quantitative Assessment of the Operational Value of Space-Borne Radar and Lidar Measurements of Cloud and Aerosol Profiles, ESA Contract Report on Project 21613/08/NL/CB, 329 pp.

Kelly, G. and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the Global Observing System Experiments, ECMWF Newsletter **113**, 16–28.

Lipton, A. E. and G. D. Modica, 1999: Assimilation of visible-band satellite data for mesoscale forecasting in cloudy conditions, *Mon. Weather Rev.*, **127**, 265–278.

Lopez, P., A. Benedetti, P. Bauer, M. Janisková, and M. Köhler, 2006: Experimental 2D-Var assimilation of ARM cloud and precipitation observations, *Q. J. R. Meteorol. Soc.*, **132**, 1325–1347.

Macpherson, B., B. J. Wright, W. H. Hand, and A. J. Maycock, 1996: The impact of MOPS moisture data in the U.K. Meteorological Office mesoscale data assimilation scheme, *Mon. Weather Rev.*, **124**, 1746–1766.

Mahfouf, J.-F. and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Part II: Experimental results with improved physics, *Q. J. R. Meteorol. Soc.*, **126**, 1171–1190.

McNally, A. P., 2009: The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var, *Q. J. R. Meteorol. Soc.*, **135**, 1214–1229.

Rabier, H., F.and Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics, *Q. J. R. Meteorol. Soc.*, **126**, 1143–1170.

Stephens, G., D. Vane, R. Boain, G. Mace, K. Sassen, Z. Wang, A. Illingwort, E. O'Connor, W. Rossow, and S. Durden, 2002: The CloudSat mission and the A-train, *Bull. Am. Meteorol. Soc.*, **83(12)**, 1771–1790.

Stiller, O., 2010: A flow-dependent estimate for the sampling error, *J. Geophys. Res.*, **115 (D22)**, doi: 10.1029/2010JD013934.

Storto, A. and F. Tveter, 2009: Assimilating humidity pseudo-observations derived from the cloud profiling radar aboard CloudSat in ALADIN 3D-Var., *Meteorol. Appl.*, **16**, 461–479.

Thépaut, J.-N., 2003: Assimilation of remote sensing observations in numerical weather prediction, in *Procceeding of the NATO Advanced Study Institute on Data Assimilation for the Earth System, Acquafreda, Maratea, Italy, 10 May - 1 June 2002*, pp. 225–240.

Veerse, F. and J. N. Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation., *Q. J. R. Meteorol. Soc.*, **124**, 1889–1908.

Vukićević, T., T. Greenwald, M. Županski, D. Županski, T. Vonder Haar, and A. Jones, 2004: Mesoscale cloud state estimation from visible and infrared satellite radiances, *Mon. Weather Rev.*, **132**, 3066–3077.

Winker, D., M. Vaughan, A. Omar, Y. Hu, K. Powell, Z. Liu, W. Hunt, and S. Young, 2009: Overview of the CALIPSO mission and CALIOP data processing algorithms, *J. Atmos. and Ocean. Tech.*, **26(7)**, 2310–2323.