CMUG Deliverable

Number:	**D3.6**
Due date:	**June 2014**
Submission date:	**30 June 2014**
Version:	**0.1**

Climate Modelling User Group

Deliverable 3.6

Roadmap for data integration in Reanalysis

Centres providing input: ECMWF

Version nr.	Date	Status
0.1	30 Jun 2014	Initial template and draft RD

CMUG Deliverable
Number: D3.6
Due date: June 2014
Submission date: 30 June 2014
Version: 0.1

Deliverable 3.6

Roadmap for data integration in Reanalysis

Table of Contents

1. PURPOSE AND SCOPE OF THE DOCUMENT 3

2. REVIEW OF PROGRESS 3

2.1 PRIORITIES 3

2.2 GAP ANALYSIS 4

2.3 STATUS OF THE OZONE ASSIMILATION IN REANALYSIS 4

2.4 DEVELOPMENT OF THE OPERATIONAL CMF 8

2.5 OTHER 9

3. CONCLUSIONS 9

4. ROADMAP FOR FUTURE ACTIVITIES 10

5. REFERENCES 10

CMUG Deliverable
Number: D3.6
Due date: June 2014
Submission date: 30 June 2014
Version: 0.1

Roadmap for data integration in Reanalysis

1. Purpose and scope of the document

This short document aims at summarizing the activities performed during the three month contract extension of CMUG Phase 1 (1 April - 30 June 2014), and at providing a roadmap on how to fulfil the vision behind the ECMWF CMUG Phase 2 core activities and achieve their envisaged goals, as defined below:

Vision:
Fostering and promoting the usage and exploitation of the CCI ECVs within the Reanalysis community and the upcoming Copernicus Climate Change and Atmosphere Services.

Goals:

1. *To assess the quality of the information provided by the Ozone CCI team (in terms of observation values, uncertainties, and biases), and demonstrate its potential impact in Reanalysis.*
2. *To assess the quality of the information provided by the Ozone, Aerosol, and GHG CCI teams (in terms of observation values, uncertainties, and biases), and demonstrate their potential assimilation impact and mutual consistency.*
3. *To continue developing the Climate Monitoring Facility (CMF) to facilitate multi-year observation - model comparisons of level 3 statistical averages.*
4. *To continue managing, updating and extending the content of the CMF database.*

2. Review of progress

Most of the work performed during the three-month period extension consisted in technical preparation, preliminary scientific analysis and coordination that will ensure to realize the Phase 2 goals as given above.

2.1 Priorities

The activities were prioritized so that full advantage could be taken of upcoming initiatives and projects where the usage of the CCI atmospheric retrievals could and should be promoted.

It is noted that as part of the ERA-Clim project the next European atmospheric reanalysis that will cover the satellite era and replace the ERA-Interim datasets is expected to start between Q4 2014 and Q1 2015 at the latest.

It is also noted that plans for a long reanalysis of atmospheric composition still need to be defined. Such a reanalysis is not foreseen for the remaining of the MACC project. Although a new atmospheric composition reanalysis could be included in the Copernicus Atmosphere Services (CAS) activities (the work plans for CAS are still under discussion, and thus it should not be ruled out at this stage), its start is not regarded as imminent.

Based on these considerations, the necessary preparation for the exploitation of the Ozone CCI products in Reanalysis was given the highest priority. An account of the current status of the work is given in section 2.3 below.

Additionally, work has started on the operational CMF development and CMF database towards the achievement of goals 3 and 4. Details are provided in section 2.4 below.

2.2 Gap analysis

At the time the Second Supplement Contract was defined and approved, the ECMWF participation in the so-called "CCI for MIP" (CCI4MIP) activities was under discussion and proposed for the CMUG Phase 2. These activities intended to assess the usage of the CCI SST and the CCI SI datasets as forcing at the ocean surface in an Atmospheric Model Inter-comparison Project (AMIP) experiment using the ERA-Clim model. To that end, the Second Supplement Contract valid for the period April-June 2014 foresaw interaction with the reanalysis section at ECMWF and some preliminary work on the definition and set-up of the AMIP experiment.

Due to budgetary consideration, the ECMWF contribution to the CCI4MIP work package had to be withdrawn from the CMUG Phase 2 proposal. As a consequence, attention was paid in advancing the preparation in other areas, particularly in the integration of the ozone products in the reanalysis system.

2.3 Status of the ozone assimilation in Reanalysis

The aim of this work package is to assess the quality of the level 2 ozone retrievals produced by the O3 CCI team and demonstrate their usefulness in an atmospheric reanalysis context. The O3 CCI team has three lines of production: one stream for Total Column Ozone (labelled as TCO_3) and two streams for ozone profiles retrieved from nadir UV backscatter sensors (this product is referred to as NPO_3), and from limb and occultation sensors (referred to as LPO_3), respectively. For each line of production, measurements from a number of different sensors have been considered so that a temporal coverage that dates back to 1995 can be achieved for TCO_3 and NPO_3.

It is noted that, although some of the instruments can be regarded as one precursor of the others, differences between their corresponding retrievals may exist (e.g. as a consequence of anomalies that could have affected one particular instrument only), thus the assimilation impact of each dataset needs to be individually assessed.

For each dataset, a number of tasks has to be performed, many of which of technical nature to prepare the actual assimilation tests. A flowchart is provided in figure 1.

There are four main sub-packages to consider. The first three of them will be applied to each dataset individually. They consist of:

- Creation of a data converter and corresponding data preparation to allow ingestion in the IFS.
- Analysis of the data bias and uncertainties.

- Experimentation and analysis of the impact on the system (quality of the ozone analyses, impact on other used data, and impact on the forecast skill).

Once the results from the three above sub-packages are available, a further analysis can be carried out. The aim is twofold. On one hand, the impact of assimilating the CCI ozone datasets will compare against that of using the same datasets produced with competing algorithms, a sort of Round Robin exercise in assimilation mode (e.g. CCI SCIAMACHY TCO3 vs. KNMI SCIAMACHY TCO3). It is noted that in many cases, ECMWF has already some experience in using the datasets produced with the competing algorithms, and thus this analysis is required to more efficiently promote the usage of the CCI data. On the other hand, the impact of data vertical resolution and viewing geometry can also be assessed and provided as a further feedback to ESA, in particular, and Space Agencies and retrieval expert teams in general.

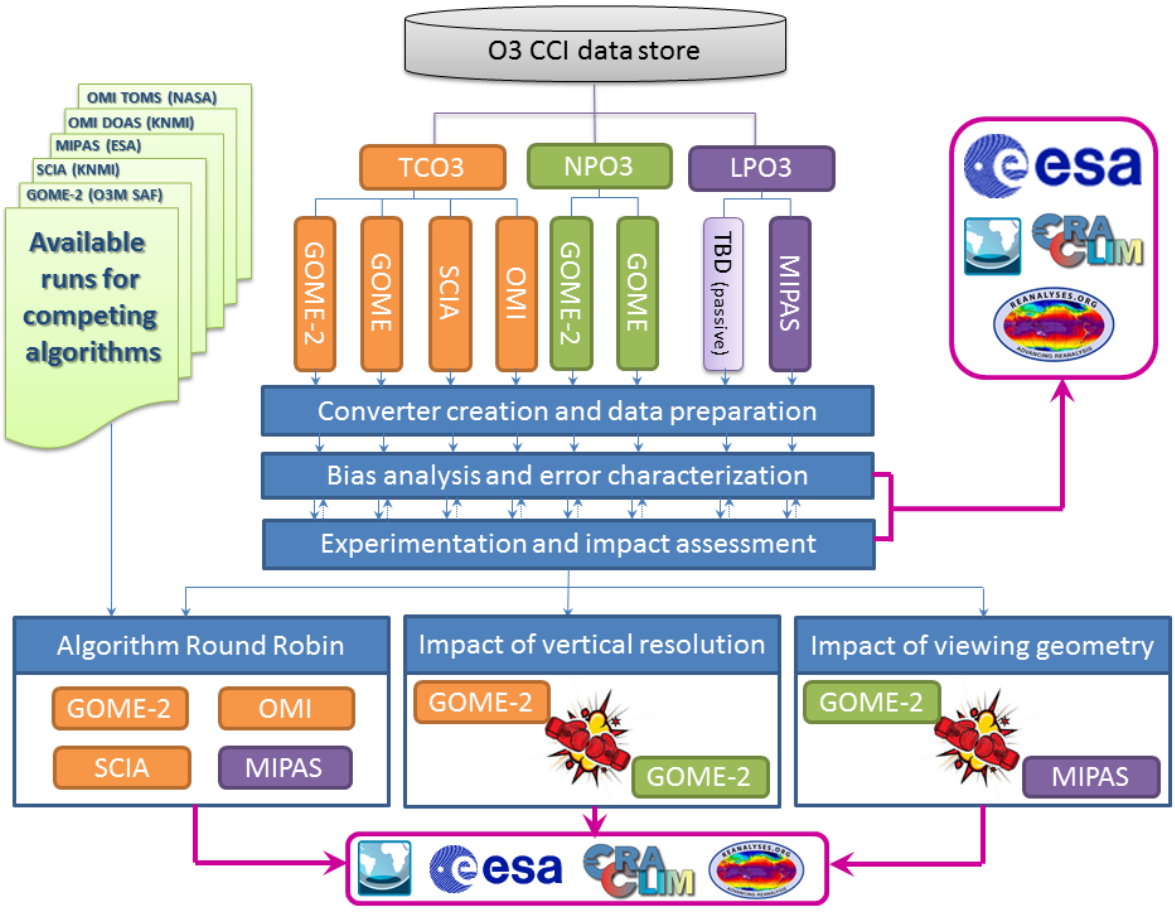

Fig 1: Flowchart of the main sub-packages to be performed. The magenta arrows symbolise feedback to a number of possible stakeholders. Here, the "reanalysis.org" is used to represent the wide reanalysis community.

Although only very preliminary results are available, as none of the experiments has finished yet, some first indications can already be seen. As an example, figure 2 shows the temporal mean zonal average difference between the ozone profiles retrieved from the Microwave Limb Sounder (MLS) and co-located ozone analyses obtained from three assimilation

experiments. While the MLS ozone profiles cannot be regarded as a validation truth, they are arguably one of the most accurate globally available ozone products. Thus, reducing the difference between MLS and co-located ozone analyses can be regarded as an indication of an improved quality of the ozone analyses themselves.

The top panel plot of figure 2 was obtained using the analyses from a control experiment (CTRL). This control run assimilated all observations available in the ECMWF data archive for the non-ozone related observations. In contrast, the baseline ozone observations used to constrain the ozone analyses were chosen to be the SCIAMACHY TCO3 retrieved at KNMI, low vertical resolution profiles from three SBUV/2 instruments flying on NOAA-16, -17 and -18, and ozone-sensitive channels in the IR from AIRS, IASI, and HIRS sensors.

The middle and bottom panels in figure 2 were obtained using ozone analyses that were constrained by the same observations used in CTRL plus the reprocessed GOME-2 TCO3 from the O3M SAF and CCI, respectively. These two sets of analyses are labelled as G2SAF and G2CCI, respectively.

All experiments are set-up to cover a period of three months, from August to October 2008. Due to data availability, figure 2 refers to August 2008.

Although preliminary, figure 2 suggests that the assimilation of the CCI GOME-2 TCO3 has a substantial impact on the ozone analyses while that of the O3M SAF dataset is negligible compared to the CTRL. This means that the assimilation system is able to extract more information from the CCI dataset than it can do from the O3M SAF product, and that this extra information is used to reduce the differences with the MLS ozone profiles at most levels in the lower and middle stratosphere.

Figure 3 shows instead the fit in terms of RMSE of the three sets of analyses to a number of ozone sondes available for August 2008. The fit is given in terms of RMSE, thus the smaller the values the better the agreement between the ozone analyses and the ozone sondes. As in the case of the comparisons with MLS, a better fit to the ozone sondes is regarded as an indication of improved ozone analyses.

The comparisons with sondes generally confirm the results of figure 2, with the Ozone CCI product producing a large improvement in the stratospheric ozone analyses. In the troposphere, some degradation is seen below 500hPa at mid and high latitudes in the SH. It is important to consider that the figures refer to the month of August (i.e. winter in the SH), thus at least the high latitudes in the SH might not be under the best sunlight conditions for UV measurements. This might require some further analysis in terms of error characterization and bias that will need to be taken into account for assimilation.

Despite the small problems in the SH troposphere, these preliminary results are encouraging, and if they will be confirmed and at the same time no degradation will be found on the rest of the system, the GOME-2 TCO_3 retrieved by the Ozone CCI can be a good candidate to be assimilated in the forthcoming reanalysis production.

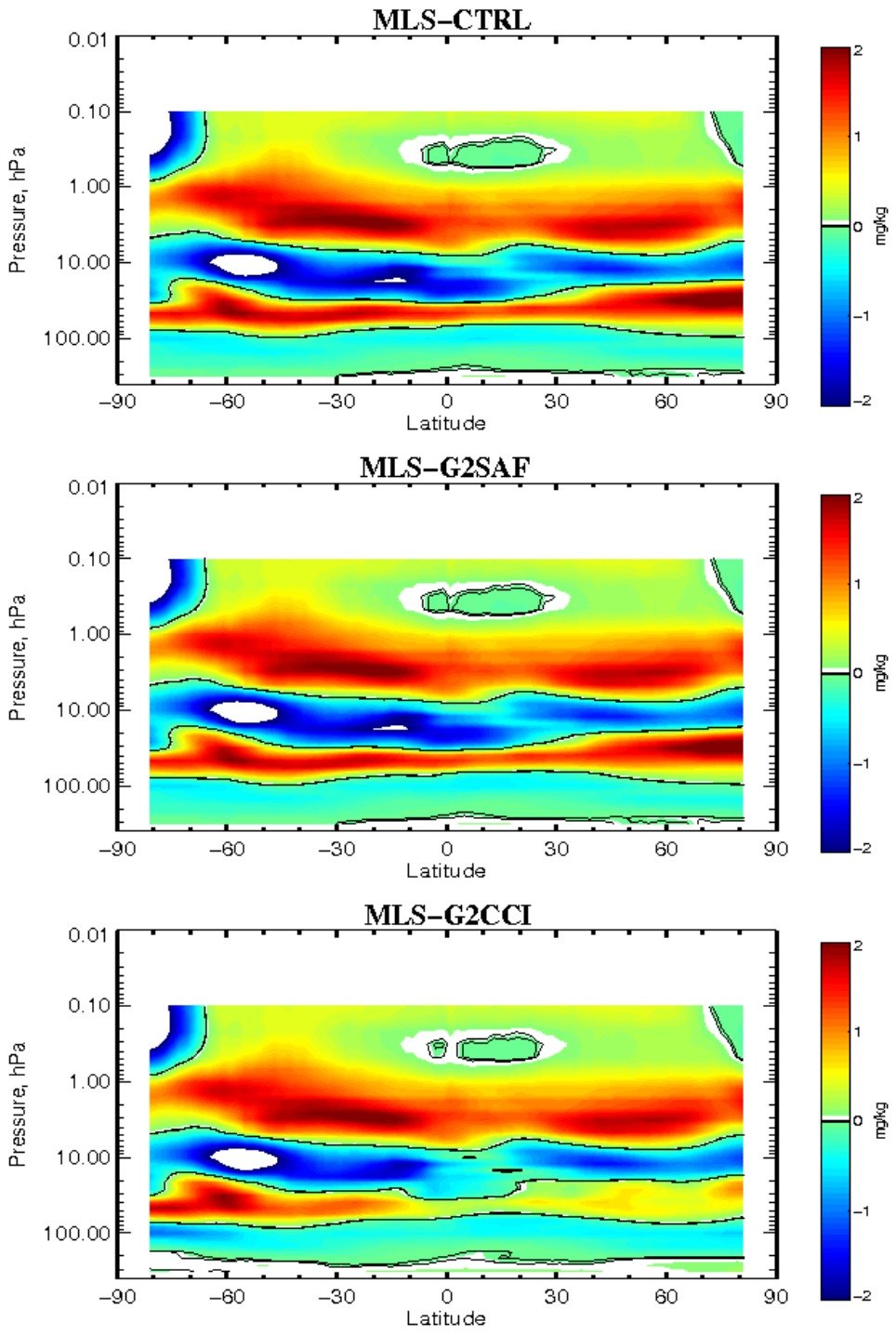

Fig. 1: Temporal mean zonal mean difference between the consolidated MLS ozone profiles and collocated ozone analyses from the control (top panel), and from the two perturbation experiments using the O3M SAF (middle panel) and the CCI (bottom panel) reprocessed GOME-2 TCO3. The plots were obtained for August 2008. Data are in mg/kg.

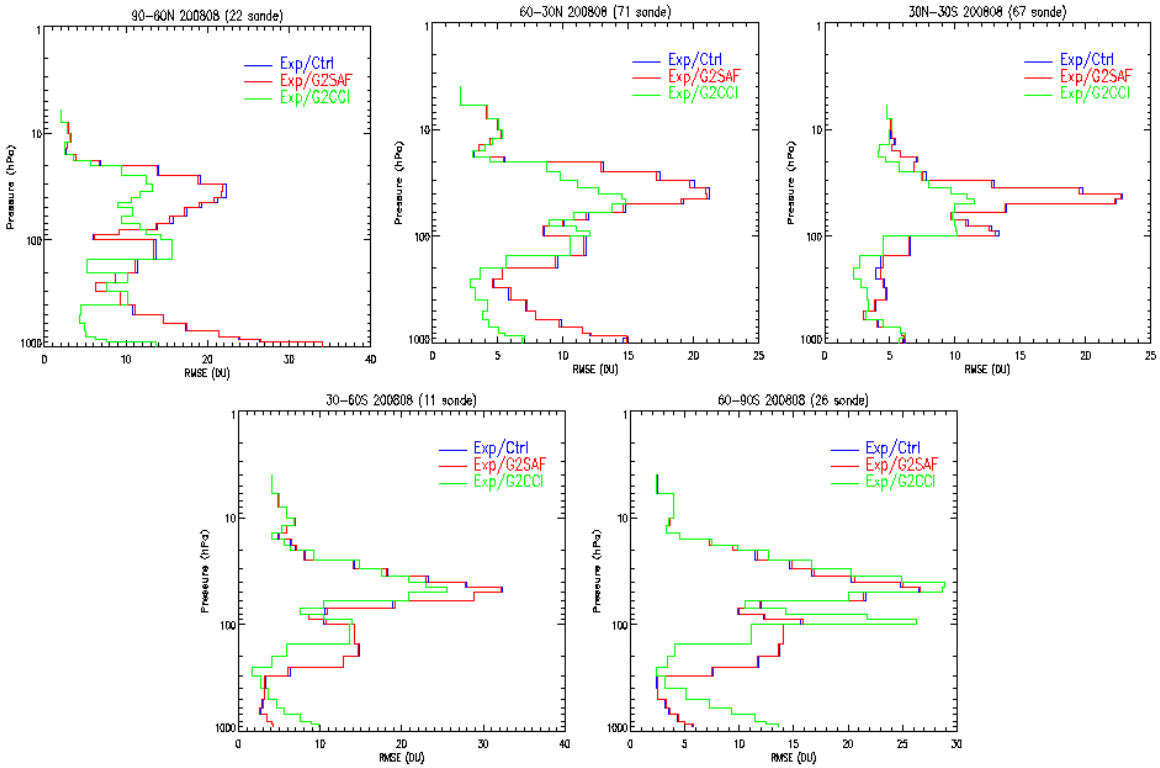

Fig. 2: RMS fit of the CTRL (blue lines), G2SAF (red lines), and G2CCI mean ozone analyses to ozone sondes averaged over five latitudinal bands for August 2008. The number of ascents included in the average can be found in the title of each panel. Data are in DU.

2.4 Development of the operational CMF

The project aims at developing the operational Climate Monitoring Facility, or CMF, i.e. an operationally maintained interface to visualize and facilitate model-observation confrontation for L3 products. The focus is on multi-year variability of statistical averages (monthly/regional means) that are pre-calculated and stored in a dedicated data store, the CMF database (CMFDb). The development will start from the CMF prototype described in CMUG (2013a, b).

2.4.1 Project organisation

In order to manage the development and implementation of the operational CMF, a team of experts has been set-up. The project sponsor is Florance Rabier, Director of the Forecast Department (FD) at ECMWF, and the project manager is John Hodkinson (team leader in the Production Section within FD). The project board comprises a Senior User (Dick Dee, head of Reanalysis Section) and a Senior Supplier (Sylvie Lamy-Thépaut, senior analyst in the Web Services unit). The role description follows the standards defined by the PRINCE2 project management (PRojects IN Controlled Environments, http://www.prince2.com/).

The project team includes a database expert, a web designer and a graphic expert, as well as researchers from the reanalysis section. The ECMWF expert working on the CHARMe project will also collaborate with the rest of the team to integrate a feedback and metadata facility in the operational CMF.

Project assurance will be achieved by using the PRINCE2 project management methodology and principles.

2.4.2 CMF Database

The CMF Database includes statistical averages an excess of 100 distinct variables defined over 32 different geographical regions, 12-18 layers (if applicable), several data streams (various reanalyses and several CCI datasets).

A problem was found in the CMFDb data produced for one of the algorithms used by the GHG CCI team to retrieve CO_2 as a consequence of differences in the data format/name variables used that was identified after some iterations with the GHG CCI team. The error was corrected and the CMFDb for CO_2 was regenerated. The GHG CCI team reassured that time during Phase 2 will be devoted to improve the product homogenisation by using a common specification at least for the most important parameters.

In parallel, a review of the content and of the data schema for the CMFDb has also started.

2.5 Other

During the three month period covered by the Second Supplement Contract, interaction with the CCI teams continued, particularly with the GHG CCI (as discussed in section 2.4.2), and the Ozone CCI.

In the latter case, it is worthwhile to mention that an agreement was reached on an early provision of the OMI TCO_3 retrievals for the period August-October 2008 so that initial assimilation tests can be performed and results collected before the start of the next reanalysis.

3. Conclusion

The three month period covering the Second Supplement Contract has been spent to carry out a number of technical activities and coordination required to realize the Phase 2 goals described in section 1. The work mostly concerned with 1) the preparation for assimilation tests that aim at utilizing as many Ozone CCI products as possible in the next reanalysis production - due to start between Q4 of 2014 and Q1 of 2015 -, and 2) collaboration with the Forecasts Department in establishing the infrastructure for the development of the operational CMF. Both aspects of the work are progressing as expected.

A gap analysis was performed between the accepted work package for the Second Supplement Contract and the actual performed activities. The former included preliminary work to define the set-up for an AMIP experiment using the CCI SST and SI products (CCI4MIP). However, as mentioned above, due to budgetary reason the ECMWF contribution in the CCI4MIP work package had to be withdrawn at a later stage. As a consequence, the time initially accounted for setting up the AMIP run was devoted to progress the integration of the CCI ozone products in reanalysis. This permitted to have to date some preliminary results based on the first month of the GOME-2 TCO3 assimilation. These preliminary results are very encouraging. Not only do they show that the GOME-2 TCO3 retrieved by the Ozone CCI team can substantially improve the agreement of the stratospheric ozone analyses to MLS

and ozone sonde data compared to a baseline experiment that did not use GOME-2 at all, it actually suggest that the amount of information that can be extracted from the CCI retrievals is much higher and can better constrain the ozone analyses than their O3M SAF equivalent, which instead has a neutral to marginal positive impact.

The results so far are very encouraging. If they will be confirmed and the further analysis will show no degradation in the rest of the system, then the GOME-2 TCO3 product retrieved by the CCI team will represent a good candidate to be exploited in the forthcoming ERA-SAT reanalysis.

Based on these preliminary results, an agreement was reached with the O3 CCI team on an early provision of the OMI TCO3 retrievals for the period August-October 2008 so that initial assimilation tests can be performed and compared with other available OMI TCO3 retrievals, so that a decision can be made on the best dataset to use in the next reanalysis ahead of its start.

4. Roadmap for future activities

Table 1 provides an indication of the work performed so far and a roadmap for what is required in the next months (see “Remaining activities”, column 5) for the integration of the Ozone CCI products in the reanalysis system. The prioritization of the various sub-packages and remaining tasks generally follows the same order that they are given in table 1. The only exception could be a review of the bias and error characterization that may also be carried out to optimize the assimilation in the forthcoming reanalysis if the assimilation results may indicate some problems.

In the more general context of the ECMWF contribution to the CMUG Phase 2 activities, the work package on the integration of the Ozone CCI products in reanalysis will still be given the highest level of priority, so that as many products as possible will be assimilated in the forthcoming reanalysis. Not only is this regarded as a major point to permit the exploitation of these datasets in the forthcoming European reanalysis, but also to promote their usage to the wider reanalysis community through a number of successful stories. The integrated exploitation of the O_3, Aerosol and GHG products will start afterwards.

In parallel, the development of the operational CMF will continue. This will be achieved through a coordinated work performed by the team of experts from both the Forecasts Department (that will ultimately support this facility in future), and the reanalysis section (that will provide expertise and guidance on the data usage and visualization).

5. References

CMUG, 2013a. Status of the ECMWF Climate Monitoring Database Facility. Technical Note/Deliverable 3.5, May 2013.

CMUG, 2013b. The role of Reanalysis in the production and quality assessment of CDRs. Technical Note/Deliverable 3.4, May 2013.

Sub-WP description	Data stream	Tasks	Current status	Remaining activities
Data preparation	TCO3	• Creation of NCDF2BUFR converter • Dataset conversion	• Done for all three streams. • GOME-2 (TCO3 and NPO3), SCIA TCO3, and MIPAS LPO3	Conversion of GOME TCO3 and OMI TCO3 (when available)
	NPO3			Conversion of GOME NPO3
	LPO3			Conversion of an additional limb dataset
Bias analysis and error characterization	TCO3	Comparisons with ozone (re)analyses	Started for GOME-2 and SCIAMACHY.	• OMI • GOME
	NPO3		To be done	• GOME • GOME-2
	LPO3		On-going MIPAS.	• Finish MIPAS • Additional limb
Active / passive assimilation experiments & output assessment	TCO3	• Analysis quality • Impact on used data • Impact on forecast score	Preliminary analysis of the1st month from SCIAMACHY and GOME-2 runs.	GOME
	NPO3		To be done	GOME and GOME-2
	LPO3		Preliminary analysis of the1st month from MIPAS run	One additional limb (passive)
Further analysis	• TCO3 vs. NPO3 • NPO3 vs LPO3	• Impact of data resolution • Impact of viewing geometry	To be done	To be done
	TCO3	Algorithm Round Robin	SCIA, GOME-2 in progress	OMI if CCI data available on time
	LPO3		MIPAS in progress	N/A

Table 1: Break-down of the main tasks to be performed, with a distinction between the on-going activities and those that will yet need to be start.