ECMWF
Copernicus
Procurement

COPERNICUS PROJECT

CAMS_61

Development of regional air quality modelling and data assimilation aspects

ITT Ref: CAMS_61
Submitted by: TNO
Date: 29 July 2019
Version: 1
TNO quotation no. 928399

Table of Contents

1 Executive Summary

Overall Objective

The proposed project will deliver development plans, guidelines, working examples and tools for the continuous upgrade of the CAMS regional service. This will encompass (i) a in depth assessment of the CAMS regional forecasts and a prioritized list of proposed model developments, (ii) best practices for coupling forecasts to analyses, and (iii) model-agnostic tools for the data assimilation of Sentinel-4 and 5p observations. The project will guide the shaping of the development plans for the CAMS regional systems. Through this, the project will lead to higher quality forecasts and analyses for the users of the CAMS regional service. The consortium consists of teams currently involved in the CAMS regional service, bringing together the required expertise and computational capabilities to perform the in-depth analyses, and tests under the strict configuration constraints. In combination with the regular interaction with stakeholders, the complementarity of the consortium teams ensures the deliverables will be applicable to all CAMS regional air quality systems.

Specific Objectives

The proposed consortium will investigate and test approaches for improving the CAMS regional ensemble forecasts and analyses. The project will focus on three specific objectives, each tackled by one of the work packages:

1) Identify and prioritise model developments, using the operational configurations from the CAMS regional air quality service, that are needed to improve the quality of the CAMS Regional Ensemble and the individual Regional Systems.
2) Evaluate the benefit and suggest the best practices of initiating regional air quality forecasts on the current analyses of (mainly) surface observations.
3) Prepare for the assimilation of observations from geostationary satellites (Sentinel-4).

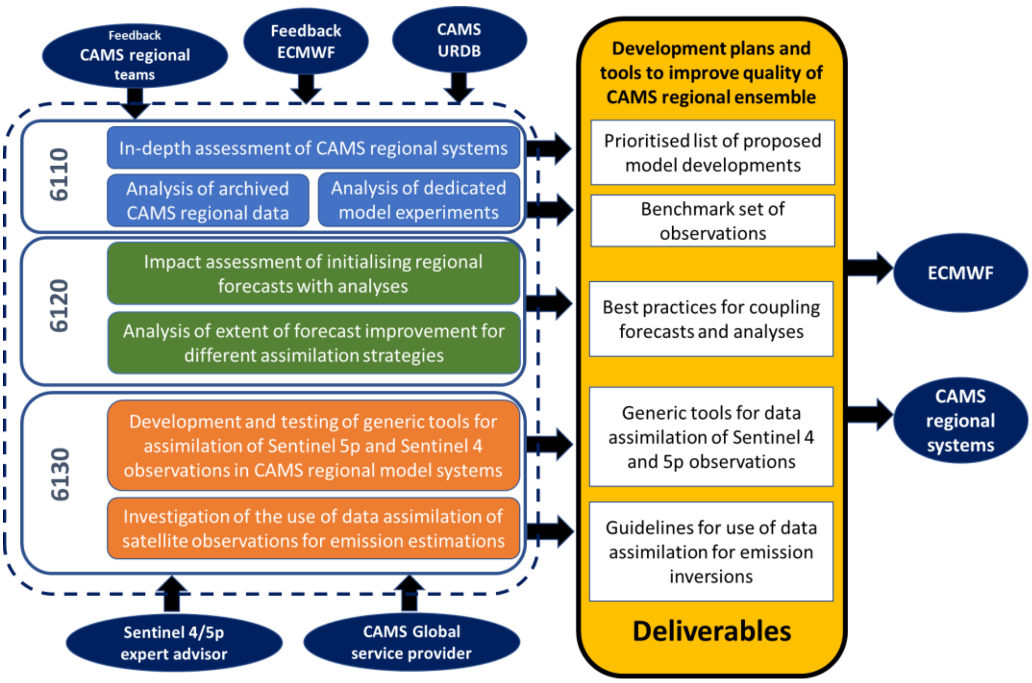


Figure 1. Overview and structure of proposed CAMS_61 project, highlighting the deliverables and interactions with stakeholders. URDB = User Requirements Data Base.

The **first objective** will be achieved by an in-depth evaluation of the CAMS Regional Systems forecasts. An overview of situations where individual models or the Ensemble deviate significantly from observations will be provided, including possible reasons for the discrepancies. Several CAMS regional systems represented in the consortium will perform dedicated model runs and sensitivity studies needed for identification of the reasons behind the model discrepancies and possible ways of resolving them. The in-depth evaluation will lead to a prioritized list of individual model developments as well as developments that are common to all models. Priorities will be set for the actions with the largest impact on model performance, specifically for the episodes with high concentrations of pollutants that are most relevant to users of the CAMS regional air quality services. Three meetings with ECMWF and all CAMS regional model teams are planned to discuss evaluation results, gather feedback, and increase the likelihood of implementation of suggested model developments.

The **second objective** will be achieved by performing several experiments with the forecasts initialised by analysis based on surface observations. The setup of the experiments will be based on the experience of the consortium members and operational configurations of the models but also account for the recommendations of a literature review. Based on the experiments we will evaluate the benefit of initialising the forecasts with the analyses for different pollutants, forecast start time, air quality regimes, and lead times of the forecasts. One of the main aspects is the limited "memory" of the air quality modelling systems regarding their initial state. To address this issue, we will also investigate the benefit of assimilation of model parameters and/or emission fluxes and inheritance of new emission/parameter estimations into the forecasts.

The **third objective** will be achieved by the development of model agnostic tools for handling observations from geostationary satellites. The tools will comprise of a pre-processor for satellite data enabling the download, selection, and conversion of observations into a common format and reduced file sizes, and a generic observation operator that simulates the satellite observations (including pixel footprints, vertical column and vertical sensitivity of the instrument) from the modelled concentration fields. The technical functionality of the tools will be evaluated by all five CAMS_61 modelling systems, to ensure applicability with the variety of models in the CAMS Regional Ensemble. Three systems with different assimilation techniques will go further by adapting the assimilation configurations for analysis of Sentinel 5p and synthetic Sentinel 4 observations. The adapted configurations should ensure that satellite data with dominant volume does not overrule the valuable but smaller set of surface data, and will serve as guidelines for assimilations by other CAMS regional systems.

The tasks to reach the three objectives will be performed by a consortium of five complementary model teams (representing four operational and one candidate model from the CAMS regional air quality service) with large knowledge and experience in model development and evaluation, data assimilation, emissions, satellite observations, development of tools and ground observation classification and screening. The inclusion of CAMS regional systems in CAMS_61 means there is a strong incentive in the consortium to deliver products that will lead to the desired model improvements. It also enables dedicated model runs for extensive evaluation and ensures pre-existing availability of model systems in the configurations required by the call (reducing the start-up time in this relatively short project period to a minimum), as well as the applicability of developed tools to the majority of CAMS regional systems. To ensure the utilisation of most recent specifications for Sentinel 4 and 5p, an external advisor with the required knowledge will be engaged and invited to attend relevant meetings. One or two external advisors will be engaged to review the planned deliverables. The advisors will be selected after consultation with ECMWF.

2 Track Record

2.1 General

The presented proposal is led by the Netherlands Organisation for Applied Scientific Research (TNO). An institute with experience in the coordination of large international projects and partner in many CAMS projects (CAMS_50, CAMS_71, CAMS_73, CAMS_81, CAMS_95). TNO's subcontractors in the framework of the CAMS_61 tender are:

- The Norwegian Meteorological Institute (MET Norway)
- The Finnish Meteorological Institute (FMI)
- The National Institute for Industrial Environment and Risks (INERIS)
- The Barcelona Supercomputing Centre (BSC)

The consortium members were carefully selected based on their complementary expertise, facilities, and experience related to the scope of the ITT. The ITT asks for a mix of expertise on model evaluation, data assimilation of surface and satellite observations, and skills regarding project management, development of generic tools, and data management. The proposed consortium includes institutes with large experience in dynamic and diagnostic model evaluation activities, such as EURODELTA and AQMEII, data assimilation of satellite observations of various species with a variety of data assimilation techniques, and in the development of generic tools for handling data. In addition, the consortium has experience in coordinating and collaborating in international projects. The members of the consortium mostly hold permanent positions. A description of each institute with key facilities, infrastructure and equipment, and previous and current involvement in projects related to this tender is given in the next subsections.

The institutes are complementary to each other regarding the use of different data assimilation systems (which allows testing of developed tools within different types of systems) and specific expertise relevant for the proposed activities. Besides the common expertise on air quality modelling and model evaluation, specific complementary expertise for each institute is:

- TNO has large experience in leading large projects and expertise concerning active data assimilation with the Ensemble Kalman Filter (EnKf) technique. They also have the in-house expertise on bottom-up emissions as provided in CAMS_81 (emission services).
- MET Norway has large expertise and a long track record concerning the evaluation of model performances. They are exploiting 3D-var assimilation within their EMEP model. They are partner of CAMS_84 (validation services) strengthening the use of CAMS_84 findings and recommendations within CAMS_61.
- FMI has expertise in both variational (3D-Var and 4D var) and EnKf/EnKS data assimilation and the inversion of emissions from the use of satellite observations. They also have experience in the development of generic observation operators from the EUNADICS-AV project. FMI a.o. also contributes to the new CAMS-44 developing the GFAS fire system and to the forthcoming CAMS-63 that will be optimising the regional ensemble algorithms.
- BSC has specific expertise on the modelling and evaluation of dust, modelling for the Mediterranean region, and on ground observation classification and screening. They are partner of CAMS_84 strengthening the use of CAMS_84 findings and recommendations within CAMS_61.
- INERIS has a large experience in intercomparison exercises (e.g. EURODELTA) and are leading the scientific development work in CAMS_50 (regional services).

More details on specific expertise and individual track records are given in the following subsections and section 3.

In addition to the 4 subcontractors TNO will engage an external advisor on satellite observations (specifically Sentinel 4 and 5p) among others to ensure the utilisation of most recent specifications for the satellite instruments considered in this proposal.

The teams involved have access to computing facilities with proven operational capabilities and large storage capacity, enabling the storage of large observational datasets from satellite instruments. The details of the equipment are included in the following subsections.

2.2 TNO

The Netherlands Organisation for Applied Scientific Research (TNO), is The Netherlands' largest knowledge organization servicing companies, government bodies, and public organizations. The Climate, Air and Sustainability (CAS) department has a strong track record on air quality, data assimilation, and natural and anthropogenic emissions. The department investigates the emission and processing of anthropogenic pollutants in the atmosphere and their influence on the environment and climate change. For this, the department developed, tested and applied the modelling system LOTOS-EUROS (https://lotos-euros.tno.nl) which is capable of assimilation of in situ observations and satellite data using the ensemble Kalman filter approach. The LOTOS-EUROS regional air quality model is one of the 9 models in the Copernicus Regional air quality ensemble since 2008 and is operated as a joint effort between KNMI (Royal Netherlands Meteorological Institute) and TNO. The model is used for many air quality applications including air quality forecasts and analyses over the Netherlands and Europe in collaboration with RIVM and KNMI, emission (trend) estimates from combined use of the model with (satellite) observations, evaluation of added value of future satellite observations through Observing System Simulation Experiments, source apportionment of limit value exceedances, scenario studies and deposition studies. Within the department global and European anthropogenic emission inventories are constructed for air pollutants. TNO is responsible for the European anthropogenic emission data (within CAMS-81) used in the CAMS_50 project.

Participation in CAMS_61

Role and involvement: TNO will be the main contractor for CAMS_61 and WP leader of WPs 6100 (Management and coordination) and 6130 (Towards assimilation of observations from geostationary satellite sensors (Sentinel-4) to constrain concentrations and emissions of main pollutants). TNO will also contribute to WPs 6110 and 6120 with dedicated model runs, evaluation of model results focusing on wood burning emissions and the evaluation of the value of using analyses results (including emission updates) for improving forecasts
Activities main location: Utrecht, the Netherlands.
Provisional budget share: 489k€ +10k€ for engaging external advisors. (≈37,5% of the current contract budget)

Key team members

Prof. Dr. Martijn Schaap is a senior researcher on air quality at TNO and part time professor in atmospheric chemistry at the Free University of Berlin. He has nearly 20 years of experience on chemistry transport modelling. He coordinates the development of the LOTOS-EUROS air quality model in the regional modelling group at TNO. His research focuses on particulate matter and the nitrogen cycle. In the last years he has performed and coordinated projects on source apportionment and the assimilation of observations into the air quality model. He has coordinated the EU FP7 project EnerGEO and contributed to numerous other EU projects such as AIR4EU, MACC, CAMS, PASODOBLE, MEGAPOLI, MarcoPolo. In CAMS_61, Martijn will be the scientific manager (Prime investigator) of the project.

Dr. Arjo Segers is a leading expert on data assimilation. He has over 20 years of experience as a research scientist in the field of atmospheric chemistry modelling. During this period, he worked with and contributed to the regional LOTOS-EUROS model, as well as the global TM5 model and their various data assimilation systems. Arjo is strongly involved in CAMS_50 regarding the development of LOTOS-EUROS for the regional air quality services. In CAMS_61, Arjo will be the work package leader of WP6130. He will mainly work on the development of specific elements for the data assimilation of Sentinel 4 and 5p data, on the use of analysis results (including updated emission information) in the forecasts and on the investigation of the use of satellite observations for emission estimates.
Dr. Renske Timmermans is working in the field of atmospheric measurements, dynamics and chemistry for over 20 years as scientist and project leader. She is a leading expert on the use of OSSEs within the air quality community. She has been coordinator of several (inter)national projects for the past 10 years and has been work package leader in the EU project PASODOBLE. She is a member of the scientific committee of the ITM meeting on Air Pollution Modelling and its Application and of the Global and Regional Atmospheric Modelling (GLOREAM) workshops. She is coordinating the TNO activities in CAMS_50. In CAMS_61, she will be coordinating the TNO activities in WPs 6110 and 6120.
Peter Coenen MSc, graduated in 1986 at the Wageningen Agricultural University in the Netherlands in the field of air pollution, toxicology and industrial hygiene. He has a background in emission monitoring and technology consulting in industry. In 1997 he joined TNO as researcher and senior project manager. In this function he has senior expertise in the field of monitoring air pollutants and GHG, abatement technologies, (international) emission inventories (air and water) and reporting obligations and policies. His senior managerial expertise is deployed in complex international studies within the department of Climate, Air and Sustainability. In CAMS_61, he will contribute to the project management.
Dr. Hugo Denier van der Gon is senior scientist and emission inventory expert at TNO. He has over 20 years of experience in measuring and estimating emissions from anthropogenic sources. He coordinates a team at TNO that provides emission inventories to modelling groups and policy advising bodies such as the UNECE-EMEP, various EU-IPs, and Dutch institutes (RIVM, PBL). He is in charge of providing high resolution emission inventories under the MACC projects as well as CAMS_81. In CAMS_61, he will advise on emission inventories and their uncertainties in all WPs.

Key facilities, infrastructure and equipment

The research team has access to TNO's in-house high-performance computing server. The current server has been installed in 2019 and is equipped with 350 cores. A disk storage of at least 100 Tb is available for the team, which can be extended on request for data-intensive projects. The server is used for simulation and assimilation experiments with the LOTOS-EUROS model (*1*), and also performs the daily source apportionment runs for the TOPAS service (topas.tno.nl).

Involvement in projects relevant to CAMS_61

CAMS: The Copernicus Atmosphere Monitoring Service (CAMS) projects represent the atmospheric core service of the EU Copernicus programme. TNO is participating in CAMS_50, CAMS_71 , CAMS_81 and a CAMS_95 use case project. **CAMS_50**: TNO is active within CAMS_50 through the provision of re-analyses and forecasts of atmospheric pollution with the LOTOS-EUROS model. The focus of TNO within CAMS_50 is on the development of the model to improve its quality and on the inclusion of requested developments by the service contractor and ECMWF. **CAMS_81**: Within CAMS_81 TNO is providing the emission data on European anthropogenic and natural aerosol components used by all the Regional modelling teams. TNO has extensive knowledge on these emissions and the associated uncertainties. **CAMS_71**: Within CAMS-71 TNO provides information on source contributions to exceedances of air quality limit values based on the LOTOS-EUROS model and a source apportionment tool.
ESA-ISOTROP: Within the ESA-ISOTROP project (*2*) the benefit of future Sentinel 4 and 5p observations for air quality applications was investigated. By performing an OSSE TNO was responsible for

evaluating the impact of synthetic Sentinel 4 and Sentinel 5p observations of NO_2 and HCHO on analyses of NO_2, ozone and PM.

ESA GlobEmission: Within this project emission inversions of a variety of pollutants over different regions in the world were produced. TNO was responsible for emission inversions of NO_2 over Europe based on OMI tropospheric NO_2 columns.

EURODELTA (1 to 3): Within the framework of the EMEP Programme, EURODELTA projects examine the performance of air quality models in predicting recent, past and future air pollution in Europe using up-to-date chemistry transport models to measure robustness of scenario predictions. TNO has contributed with LOTOS-EUROS model runs and analysis of model results against observations in all the EURODELTA projects.

2.3 MET Norway

The Norwegian Meteorological Institute is the national meteorological service of Norway and represents Norway e.g. in ECMWF, EUMETSAT, EUMETNET, and WMO. The institute employs about 450 persons, among them 140 scientists doing research within numerical weather prediction, ocean modelling, remote sensing, air pollution, product development, instrumentation, climatology and climate research. For several decades, MET Norway has been hosting the **Meteorological Synthesizing Centre – West (MSC-W), supporting the European Monitoring and Evaluation Programme (EMEP**) under the UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP). EMEP provides the technical underpinning for air pollution policies within CLRTAP and for the European Union. MSC-W has been issuing annual status reports on air pollution for all EMEP countries since the 1990s, including detailed source-receptor relationships and country notes as a basis for international air quality policy agreements (e.g. Gothenburg protocol). As a part of the annual work for EMEP, MSC-W issues yearly evaluation reports (EMEP/MSC-W model versus observations) that scrutinize the performance of the model towards all available data (including chemical observations). Through its long-term involvement in EMEP and its support of the EU Thematic Strategy on air pollution, MET Norway has the capability and mission to provide services for air pollution assessment and in the understanding of the coupling of weather variability and air pollution. MET Norway also co-ordinates the international AeroCom initiative and develops and hosts the AeroCom web interface (aerocom.met.no) which offers access to multi-model evaluation of atmospheric aerosol and chemistry models. Considerable developments in tools and observation data understanding have been developed as part of the AeroCom work.

Participation in CAMS_61

Role and involvement: MET Norway will be responsible for the work package 6110 (In-depth assessment of the CAMS Regional Systems to identify future development needs) and contribute to the study of the coupling between forecasts and analyses, as well as to the preparations for the data assimilation of Sentinel-4 in WPs 6120 and 6130.

Activities location: Oslo, Norway.

Provisional budget share: 330k€ (≈24,8% of the current contract budget).

Key team members

Dr. Hilde Fagerli (female); has 20 years of experience in air pollution modelling. She has co-authored more than 40 peer-reviewed publications. Since 2009, she is head of the Division of Climate Modelling and Air Pollution at the Norwegian Meteorological Institute, managing more than 25 scientists. She is also project leader of the United Nations EMEP/MSC-W, working on the assessment of European transboundary pollution. Nationally, she coordinates MET Norway's contribution to the inter-agency cooperation to improve Norwegian local air pollution forecasts. In CAMS_61, Hilde Fagerli will lead Work Package 6110.

Prof. Dr. Michael Schulz (male); received his PhD at the University of Hamburg in 1993, and joined MET Norway in 2010, co-leading EMEP and NorESM model development and applications. He became an adjunct professor for meteorology at the University of Oslo in 2017. His research focuses on the understanding of the role of aerosols for climate change and air quality. He worked as PI in numerous German, French, Norwegian, ESA and EU funded projects, contributing to policy relevant assessment reports such as being a lead/contributing author of the 4th&5th IPCC assessment reports. In CAMS_61, Michael Schulz will co-lead WP6110 and supervise the use of the AeroCom tool for model evaluation and the testing of assimilation of Sentinel data.

Prof. Dr. David Simpson (male) has been employed at MET Norway since 1990. He is the lead developer of the EMEP MSC-W chemical transport model system, with special interests in biosphere-atmosphere exchange (BVOC and soil emissions, ozone deposition and N-exchange), ozone chemistry, and secondary organic aerosol formation. He has participated in numerous EU-projects, among them INFOS, MERLIN, NOFRETETE, CARBOSOL, EUCAARI, NitroEurope, PEGASOS, ECLAIRE, CAMS_50 and CAMS_81. He is a member of the Scientific Advisory group (SAG) on Reactive Gases, Global Atmosphere Watch (GAW) Programme of the World Meteorological Organisation (WMO) since 2017. In CAMS_61, David Simpson will be involved in the model developments to test in WP6110 but also supervise the analysis and contribute to the WP6110 reports.

Dr. Álvaro Valdebenito (male) received his Ph.D. in natural sciences in 2008 at the University of Hamburg / Max Planck Institute for Meteorology. He has been employed at MET Norway since 2008, where he has worked on the development of the EMEP MSC-W model and the implementation of air pollution forecasting and analysis capabilities. He has been strongly involved in GEMS, MACC I-III, CAMS_50, CAMS_71 (chemical weather forecasts, data assimilation and source receptor products for policy support), largely responsible for the operational implementation of the MACC forecast model chain. In CAMS_61, Alvaro Valdebenito will implement the assimilation of Sentinel data in the EMEP model in WP6130 and participate in the model experiments of WP6120 and dedicated runs of WP6110 Task 6113.

Key facilities, infrastructure and equipment

Hardware: As the national meteorological service of Norway, and as a governmental body, MET Norway has extensive and stable access to High Performance Computing resources. In CAMS_61 MET Norway will use one of the two supercomputing clusters, which includes 104 Compute nodes (52 owned by MET Norway) with a total number of 3328 cores. The global storage available is ca 1.0 Pbyte. For data storage, analysis and evaluation MET Norway uses a Post-Processing Infrastructure, which has been established in house, and which is divided into four parts: a) storage b) post-processing section composed of Login Nodes, c) visualization nodes d) data gateways, i.e. servers that will make the storage available to other servers or directly on a computer through the CIFS or NFS protocols. Currently we have 94 Tbytes of disk space.

Software: The scripts for air quality model evaluation have been developed, tested and maintained over many years within the frame of the LRTAP convention, but we will also use the well tested AeroCom-tool (aerocom.met.no), which is used for in-house evaluation of the CAMS_50 validated reanalyses as well. As our main software for air quality modelling, the EMEP MSC-W model is used (*3*).

Involvement in projects relevant to CAMS_61

EMEP: MET Norway has been hosting the scientific and technical centre EMEP MSC-W (under the UN CLRTAP) for many years, providing modelling results on acidifying and eutrophying air pollution, photochemical oxidants and particulate matter, as well as transboundary fluxes. The work under EMEP ensures that the EMEP air quality model is up-to-date and evaluated regularly. It has allowed MET Norway to gain outstanding experience within regional air quality modelling, emissions and observational data over more than three decades.

EURODELTA (1 to 3): within the framework of the EMEP Programme, EURODELTA projects examine the performance of air quality models in predicting recent, past and future air pollution in Europe using up-to-date chemistry transport models to measure robustness of scenario predictions. MET Norway has contributed with EMEP MSC-W model runs and analysis of model results against observations in all three EURODELTA projects.

CAMS_84: MET Norway regularly scores aerosol simulation performance from the global CAMS model C-IFS run by ECMWF by evaluating optical properties, PM and speciated concentrations at the surface. Further it evaluates above surface aerosol concentrations for the Regional Systems using various observations.

CAMS_50: MET Norway provides daily air quality forecasts, analyses and re-analyses with the EMEP model as a member of the regional air quality ensemble production in CAMS_50.

CAMS_81: MET Norway is involved in the provision of natural emissions (NOx emissions from soil, BVOC emissions from vegetation, DMS emissions from oceans, etc.), including recommendations on how to calculate these emissions dynamically in models. MET Norway also leads a work package on the providers-users interface.

CAMS_71: MET Norway is developing and maintaining the CAMS_71 source-receptor allocation product and its visualisation. It has been responsible for the web-based policy portal within CAMS_71. It further contributed to policy workshops and science work leading to the evolution of CAMS_71 policy products (https://policy.atmosphere.copernicus.eu/ DailySourceAllocation.html).

Local air pollution in Norway: MET Norway is responsible for local air pollution forecasts for all Norwegian regions and cities, using the uEMEP downscaling module, and operational since January 2019. The uEMEP module is developed, evaluated and improved at MET Norway.

VALS-5p: A project with ESA and the Norwegian Space Centre validating Sentinel-5p data and preparing for the use of Sentinel-5p data in the EMEP model (assimilation and evaluation). The project runs through 2019.

AeroCom: an open international initiative of scientists interested in the advancement of the understanding of the global aerosol and its impact on climate, but widely used as model evaluation tool and model/observation data base (hosted at MET Norway aerocom.met.no).

2.4 FMI

The Finnish Meteorological Institute (FMI) is designated by the Finnish government as national air quality expert with a mandate to produce information and forecasts on the state of the atmosphere and its characteristics, with the aim of promoting safety and serving various needs of the public, industry and commerce, as well as contributing to scientific ends. FMI makes observations of the physical state of the atmosphere, its chemical composition, and electromagnetic phenomena. FMI also develops and applies numerical models – from urban to global scales – in order to analyse and forecast various atmospheric physical and chemical processes. FMI employs about 550 people, about 150 of which are involved in research. The modelling teams have extensive experience in developing and implementing various numerical systems, from urban pollution models up to global stratospheric ozone studies.

Scientists from the Atmospheric Composition Research department of FMI will be involved in CAMS_61. The Atmospheric Composition research division has its tasks to evaluate the use of analyses for forecasts and develop and evaluate the data assimilation procedures.

FMI is involved in numerous international co-operative, research and assessment efforts. Current projects involve the following activities: monitoring of air quality and atmospheric composition (e.g., EMEP, HELCOM/EGAP, WMO/GAW, AMAP), research and development in air chemistry and aerosol physics (including in particular one National and two Nordic Centres of Excellence, ACCENT, EC/Environment), assessment and modelling of the dispersion, transformation and deposition of airborne pollutants from the local to the continental scale (AirQast, PAPILA, EUNADICS, MarcoPolo, PASODOBLE, HIALINE, TRANSPHORM, PESCADO, PEGASOS, CAIR4HEALTH, ACCENT, GEOmon, ESA

GlobEmis, EUMETSAT O3M SAF and others; contributed to the AQ assessments within IPCC, UN/ECE EMEP and IM, HELCOM, WMO/GAW, AMAP, GEOSS, etc.).

Participation in CAMS_61

Role and involvement: FMI will be responsible for the work package 6120 (coupling of regional forecasts and analysis) and contribute to the other two work packages 6110 and 6130.
Activities main location: Helsinki, Finland.
Provisional budget share: 250 k€ (≈18,8% of the current contract budget)

Key team members

Research Prof., Dr. Mikhail Sofiev has over 25 years of experience in atmospheric composition research and model development, coordinator of the SILAM team, deputy leader of the FMI AQR Modelling Group and Adjunct Professor at University of Helsinki. He has an extensive experience in development and application of air pollution models at various scales – from meso- to hemispheric scales – and for various compounds – acidifying, toxic, aerosol, and radioactive accidental releases – and in related fields: model verification, statistical methodology, data analysis, computer experiments, etc. He coordinated FMI modelling work in numerous international projects, coordinated Finnish national projects POLLEN, IS4FIRES, and ASTREX, was a leader of the PASODOBLE AQ sub-project DS-PUBLIC, etc. M. Sofiev is the author of 226 scientific publications; 129 of these have been published in refereed journals and books (h-index = 40, Google Scholar 4.07.2019). M. Sofiev is a member of the WMO Scientific Advisory Group on Applications, European Aerobiological society, European Academy of Allergology and Clinical Immunology, Board member of European Aeroallergen Network, Associate Editor of Atmospheric Environment, member of Finnish emergency preparedness team, and has contributed into policy advisory boards. His role in CAMS_61 will be the coordination of WP6120 (coupling of regional forecasts and analysis) and all SILAM team activities in the project.
Adj. Prof. Dr. A. Karppinen has over 30 years of experience in air quality research and model development. He is currently leading a research group (20 researchers) on Atmospheric Dispersion Modelling. The research group is working in 15 internationally funded (EU, ESA, ESF) and 10 nationally funded research and networking projects. He is the author of more than 300 scientific publications; 70 of these in refereed international journals (h-index 29). He has been working in 19 EU-funded projects, including vice-coordinator position in EU/MARQUIS project and scientific and technical coordination of EU/PESCaDO project. His role in CAMS_61 will be coordination of the modelling work, data assimilation and fusion techniques.
Dr. Rostislav Kouznetsov has over 15 years of experience in atmospheric physics, in observational and modelling aspects. He is currently one of the main SILAM developers, responsible for its operational applications, also strongly contributing to its scientific agenda, in particular, numerical aspects and physical parameterizations. R. Kouznetsov is an author of over 100 scientific publications, 40 of which are published in peer-reviewed journals, books and series (H-index 15). In CAMS_61, he will be responsible for the SILAM model activities falling under WP6120 and 6130 and will also contribute to model development and evaluation.
Dr. Andreas Uppstu is a new scientist in the SILAM team. His background is theoretical physics (modelling of physical and optical features of graphene), where he also made his PhD (5 papers in top-level journals in the field of theoretical solid-state physics). A. Uppstu joined the SILAM team two years ago and since then became the main developer of the data assimilation modules, in particular, EnKF. He is the primary SILAM developer for the needs of H2020 EUNADICS-AV and AirQast projects. In the current project, he will be responsible for the data assimilation of Sentinel 4 and 5p observations.

Key facilities, infrastructure and equipment

The FMI supercomputer facility consists of two identical Cray XC40 systems with 172 compute nodes of 28 (56 hyperthreading) cores. Each node has 128GB RAM. Peak performance is 1035 GFLOP per node, or 178 TFLOP in total. It is accompanied with 960-TB fast-access storage and a tape archive of essentially unlimited capacity. The system is equipped with all necessary software for large-scale computations, data processing, archiving, visualization and dissemination.

Involvement in projects relevant to CAMS_61

CAMS_50: Regional air quality forecasting ENSEMBLE, one of the Regional models, responsible for pollen line development.

GLORIA (2017-2020): Global health risks related to atmospheric composition and weather. Finnish Academy. Long-term re-analysis of global/European AQ and pollen exposure.

BATMAN (2015-2018): Environmental impact assessment of airborne particulate matter: the effects of abatement and management strategies. Academy of Finland. High-resolution (1km) air quality assessment for Finland.

EUNADICS-AV (2016-2019): European Natural Airborne Disaster Information and Coordination System for Aviation. EU Horizon 2020. Responsible for the data assimilation developments WP.

NeGI NCOE (2014-2017): Ensemble-based methods for environmental monitoring and prediction. Responsible for atmospheric applications.

CarboNord (2014-2016): Impact of black carbon on air quality and climate in Northern Europe and Arctic. Nordic Ministry of Research. Project coordination and one of the models.

Ragweed (2011-2012): Assessing and controlling the spread and the effects of common ragweed in Europe. Responsible for the modelling part of the project.

2.5 INERIS

Founded in 1990, the National Institute for Industrial Environment and Risks (www.ineris.fr) is a public industrial and commercial institution acting under the **supervision of the Ministry in charge of environment and sustainable development**. INERIS leads research and expertise activities requested by public authorities and industrial operators. It comprises multi-disciplinary teams with approximately 550 people. INERIS is certified ISO9001 and builds up the whole of its work and seeks within this framework of quality assurance. A large part of the activity of the Institute is devoted to air quality: meteorological and modelling tools are developed and assessed to contribute to better understanding and management of atmospheric pollution. Expertise of INERIS relies on both local sources (industrial and urban) and transboundary air pollution.

Within this framework, INERIS brings its expertise to the Ministry in charge of the environment during the phases of negotiation and implementation of the Directives and protocols. It develops, in partnership with the National Scientific Research Centre (CNRS) the **CHIMERE model** (www.lmd.polytechnique.fr/chimere/) a chemistry-transport model that is a reference in the air quality community. INERIS ensures the continuous development of this model and the improvement of its capacities and performances through European research projects. **INERIS operates the PREv'air system** (www.prevair.org), the national air quality forecasting platform, since 2003 in cooperation with METEO FRANCE and CNRS. INERIS hosts the modelling platform in a supercomputing centre, is in charge of its maintenance, evolution, and exploitation.

INERIS is one of the three members of the **French Central Laboratory for Air Quality Monitoring (LCSQA)**. The LCSQA provides technical support to local approved associations in charge of air quality monitoring (AASQAs) and coordinates their activity.

Participation in CAMS_61

Role and involvement: INERIS will contribute to the work packages 6110, 6120 and 6130.

Activities main location: Verneuil en Halatte, France.

Provisional budget share: 123k€ (≈9,2% of the current contract budget).

Staff members and their role in CAMS_61

Augustin COLETTE is head of the Atmospheric Modelling and Environmental Mapping Unit of the French public Institute INERIS. He holds a Ph.D. in Atmospheric Sciences from Sorbonne Universités and worked in the past for UNESCO, Stanford University, Ecole Polytechnique and the private sector for Meteorological Risk Assessment. He has co-authored 75 reviewed articles in the field of atmospheric modelling. Augustin is chair of the Task Force on Measurement and Modelling in support of the UNECE Geneva Convention on Long Range Transboundary of Air Pollution, Scientific Advisor of the CAMS_50 Regional Production, Member of the Management Committee of the European Topic Centre on Air pollution, Transport, Noise and Industrial pollution of EEA and editor for the journal Geosciences Model Development.

Gaël DESCOMBES is a R&D engineer and has joined INERIS in 2018. He received his master in 1999. He has been working his first nine years in a private company for air quality modelling and meteorological risk assessment at local and regional scale. During the following nine years at NCAR, he has developed his skills in data assimilation working with different platforms (RTFDDA, GSI, WRFDA, DART and OOPS) for the Research Application Laboratory and the Mesoscale Microscale Meteorology Laboratory. He was involved in projects for nowcasting of clouds and solar energy using geostationary and polar orbiting satellites infrared data. He developed the background covariance error matrix for cloud optical depth and all sky radiances data assimilation. He was involved in projects at Atmospheric Chemistry Observations and Modelling laboratory to estimate actinic flux using satellite cloud retrievals. At INERIS, he is in charge of the development of an ensemble data assimilation system and involved in the National Project ARGONAUT (PollutAnts and Greenhouse Gases EmissiOns MoNitoring from SpAce at high ResolUTion led by the LISA laboratory, CNRS, Créteil-Paris) for inversion of the anthropogenic air pollutant and greenhouse gases emissions.

Frédérik MELEUX is a R&D engineer who joined INERIS in 2005. He received his Ph.D. in 2002 working on the links between local, regional and global air pollution with nested models. He has co-authored 40 reviewed papers in the field of air quality. At INERIS he is in charge of the coordination of technical and scientific activities in the framework of the Prev'Air operational system. He supervised a number of activities and projects related to data assimilation, to model evaluation and to forecasting system improvements. He is leading a CAMS_95 project and manages the reanalysis activities within CAMS_50.

Dr. Anthony Ung is R&D engineer who joined INERIS in 2007. His doctoral thesis dealt with air quality mapping using multi-sources data. At INERIS, he is involved in several national and European air quality projects such as PREV'AIR, CITEAIR2 and the Copernicus projects. His field of expertise is data assimilation, mapping and scripting tools for operational post-processing and validation of modelled data.

Key facilities, infrastructure and equipment

For its activities in air quality modelling, INERIS relies on two high performance computing clusters, one for the daily operational activities and the other one for development and operational activities with less time constraint. The first one is made of 400 CPUs with more than 100 To of rapid storage and is hosted at Meteo-France and benefits from a specific IT team dedicated to the supervision of HPC activities including all the system monitoring in 24/7. The second one is part of the large computing centre of CEA (French Alternative Energies and Atomic Energy Commission) made of several computers for a total of more than 500 TFlops with a large storage capacity associated.

Involvement in projects relevant to CAMS_61

INERIS has been involved in all the projects that prepared the implementation of the operational Copernicus Atmosphere services: **from GEMS (FP6) and PROMOTE (ESA service Elements) to the MACC suites**. It ran the CHIMERE model for the provision of daily air quality forecasts and analyses as a member of the Ensemble MACC air quality model. INERIS is involved in many other national and European projects and related to air quality monitoring and management. In particular:

CAMS_50, CAMS_71: INERIS has a coordination and technical role in two CAMS contracts highly related to the present ITT: CAMS_50 on Regional Production and CAMS_71 on Policy Support Services.

PREV'AIR: INERIS is with METEO FRANCE and the National Scientific Research Centre one of the founder members of the national air quality forecasting and mapping system PREV' AIR

(www.prevair.org). The institute hosts the modelling platform in a supercomputing centre, is in charge of its maintenance and evolution, and is responsible for its exploitation.
EURODELTA (1 to 3): within the framework of the EMEP program, EURODELTA examines the performances of air quality models in predicting recent, past and future air pollution in Europe using up-to-date chemistry transport models to measure robustness of scenario predictions. The scope of EURODELTA 3 coordinated by INERIS is to assess model performances against the EMEP field measurement campaigns. About 10 European chemistry transport models, among which CHIMERE, are involved.
ETC/ATNI: the European Topic Centre on Air pollution, Transport, Noise and Industrial pollution of the European Environment Agency. Since 2010, INERIS is one of the members of the Topic Centre providing expertise in the field of air quality modelling and mapping. Analysis and definition of relevant air quality monitoring strategies and assessment are also in its area of work.

2.6 BSC

The Barcelona Supercomputing Centre (BSC) is a research centre active at both national and international levels. The BSC combines unique high-performance computing facilities and in-house top research departments on computer, life, and Earth sciences, and in computational applications in science and engineering. The BSC is the main provider of public supercomputing services in Spain. The BSC represents Spain in international initiatives such as PRACE, the Research Data Association and the Big Data Value Association. The BSC has a total staff of about 700 employees. Established in 2006, the Earth Sciences Department of the BSC, hereafter ES-BSC, worked on atmospheric composition modelling. The designation of Professor Francisco J. Doblas-Reyes as Director of BSC-ES in 2014 initiated the merging of the Climate Forecast Unit of the Institut Catala de Ciencies del Clima, that has created a more efficient and competitive Department that holds a sufficient critical mass to compete with the top international research groups in environmental forecasting and Earth system services. ES-BSC is structured around four groups (Climate Prediction, Atmospheric Composition, Computational Earth Sciences, and Earth System Services), with more than 100 employees, including technical and support staff.
The department has a wide experience in running operational air quality forecasting systems (i.e. CALIOPE system, www.bsc.es/calliope) and delivering timely and high quality forecasts, observations, information and knowledge to users. BSC co-host two WMO Regional Centres on sand and dust forecasts (the Regional Specialized Meteorological Centre for Atmospheric Sand and the Dust Forecast and Sand and Dust Storm Warning Advisory and Assessment System Regional Centre, WMO SDS-WAS RC). Since October 2016, BSC hosts an AXA Chair on Sand and Dust Storms, which intends to support the two WMO SDS Regional Centres. In October 2018, the BSC started developing an ERC Consolidator Grant that focuses on dust mineralogy and its effects upon climate. The department develops the GHOST System (Globally Harmonised Observational Surface Treatment) (https://slideplayer.com/slide/17011109/), which harmonises observations from the EEA AQ e-reporting/EBAS/WMO-GAW networks in Europe.

Another major activity in the ES-BSC is the development of the Multiscale Online Nonhydrostatic Atmosphere Chemistry model (MONARCH, formerly known as NMMB/BSC-CTM)), in coordination with the works on CAMS_50.II to adapt MONARCH as a candidate model for the regional production suite, and the development of high-resolution anthropogenic emission models (i.e. HERMESv3).

Participation in CAMS_61
Involved in: WP6110: providing quality assurance/quality control in-situ observational datasets, evaluating reactive chemistry and mineral dust in southern Europe, and providing a dedicated run with the MONARCH model. WP6130: testing the Sentinel observational operator developed by TNO and generating the Sentinel-4 synthetic dataset.
Activities main location: Barcelona, Spain.
Provisional budget share: 128.5k€ (≈9,7% of total provisional budget).

Key team members

Dr. Oriol Jorba is co-group leader of the Atmospheric Composition Group of the BSC and researcher there since 2005. His research expertise includes high-resolution mesoscale meteorology and air quality, and atmospheric chemistry studies. He leads the developments of the multiscale chemical weather forecasting system NMMB-MONARCH at BSC. He has been member of the management committee of COST Actions and is part of the scientific committee of ITM on Air Pollution Modelling and its Application since 2012. He is a member of the International Cooperative for Aerosol Prediction initiative and a modeller in AQMEII. He coordinates the BSC activities to demonstrate the added value of MONARCH in the regional production for CAMS_50.II. He is also involved in CAMS_81 and CAMS_84. Under CAMS_61, he will provide scientific advice in each of the BSC tasks.

Dr. María Teresa Pay is a researcher of the Atmospheric Composition Group with 12 years of experience working with regional air quality models. She played a seminal role on the successful implementation of the first air quality forecast system at high resolution for Spain (CALIOPE). She is a EURODELTA member contributing to the Task Force on Measurements and Modelling and is part of the GLOREAM scientific committee since 2019. Under CAMS_50.II, she is in charge of the MONARCH evaluation. Under CAMS_61, she will be the coordinator of the BSC activities, and evaluate the modelled photochemical pollutants in southern Europe.

Dr. Dene Bowdalo is a researcher in the Atmospheric Composition Group since April 2018. Before that, he was employed as a Post-Doc at the University of York, applying novel spectral methods tailored for the identification of systematic biases for surface ozone in GEOS-Chem model. He is the main developer of the GHOST System. Under CAMS_50.II, he is in charge of the MONARCH developments on aerosols to fulfil the CAMS_50 requirements. Under CAMS_61, he will prepare the GHOST database for evaluation studies and will assess the models' performance in southern Europe.

Dr. Jerónimo Escribano is a researcher in the Atmospheric Composition Group working with the data assimilation team since June 2018. He was behind the success of the CAMS_43 during the first phase and he is involved in the second phase in charge of the radiance data assimilation aspects. Under CAMS_61, he will contribute to testing the Sentinel observation operator and adapt some of the strategies developed by TNO in the MONARCH system to produce the synthetic dataset of Sentinel-4.

Dr. Enza di Tomaso is a researcher in the Atmospheric Composition Group since 2013. At BSC, she is working in aerosol data assimilation and in radiance assimilation, implementing a scheme to ingest satellite observations into the BSC chemical transport model. Under CAMS_61, she will contribute testing the Sentinel observation operator developed by TNO and produce the synthetic dataset of Sentinel-4 with the MONARCH model.

Dr. Sara Basart is the scientist in charge of the WMO Sand and Dust Storm Warning Advisory and Assessment System Regional Centre and the Barcelona Dust Forecast Centre at BSC. She is the Lead Project Investigator of the DustClim and Chair of the COST Action InDust. She has experience in evaluating dust products in the CAMS_84, leading the part on dust evaluation. Under CAMS_61, she will continue with this task but using individual CAMS_50 regional models.

Key facilities, infrastructure and equipment

The BSC infrastructure and support consists of: a 13 PFlops supercomputer, a long-term storage, commodity computational facilities with both physical and virtualised environments, a solid project management team, and the outstanding collaborations with researchers in computing sciences. BSC hosts the Mare Nostrum IV which is one of the 7 Tier-0 PRACE systems currently available for European scientists. Its peak power is 11.15 Petaflops and has a total of 165,888 processors and a main memory of 390 Terabytes. EuroHPC has recently selected BSC as entity to host one of the largest European supercomputers. The new supercomputer will have peak performance of 200 petaflops and will start operations in 2021. The operational products of the Earth Sciences Department are located in a Big Data storage system of 3.4 PBytes that will be accessible through THREADS technology in the near future.

BSC will use the MONARCH regional model for the air quality modelling activities in CAMS_61.

Involvement in projects relevant to CAMS_61

- BSC is involved in different CAMS contracts: in **CAMS_84** in charge of dust evaluation; in **CAMS_81** leading the service evolution developing new temporal profiles; in **CAMS_50.II** demonstrating the benefit of including the in-house MONARCH model in the regional production service; in **CAMS_43** on the radiance data assimilation aspects; and in **CAMS_95** providing mineral dust products for aviation.
- The WMO Regional Centre Northern Africa-Middle East-Europe for the Sand and Dust Storm Warning Advisory and Assessment System (**SDS-WAS**) (http://sds-was.aemet.es/) and the Barcelona Dust Forecast Centre (http://dust.aemet.es/).
- The **DustClim** (Dust Storms Assessment for the development of user-oriented Climate Services in Northern Africa, the Middle East and Europe) project of the European Research Area for Climate Services (ERA4CS) will produce and deliver an advanced dust regional model reanalysis for Northern Africa, the Middle East and Europe covering the satellite era of quantitative aerosol information, and by developing dust-related services tailored to specific socioeconomic sectors (air quality, aviation and solar energy).

3 Quality of Resources to be Deployed

3.1 Consortium qualifications

The team built up and proposed for the development and implementation of the CAMS_61 activities includes very high-level experts with a strong experience in model development and application, model evaluation, data assimilation, satellite observations and the production of generic tools. Importantly, a major asset of the proposed consortium is that TNO and its subcontractors all develop their own chemical transport models and data assimilation capabilities, which are used within the CAMS regional air quality ensemble, ensuring the fulfilment of the requirement on model configurations requested in the ITT (without the need for long start up times in the project), the ability to perform dedicated model runs for extensive evaluation, and the presence of a strong incentive to deliver products for the improvement of the CAMS regional service. In addition, the team includes in-depth expertise on the most important drivers of air quality, i.e. meteorology and (anthropogenic and natural) emissions. Hence, all partners are well aware about the scientific and practical challenges facing the improvement of the quality of the CAMS regional products.

Model evaluation: The consortium is unique in the fact that, besides development and evaluation of their own models, it combines a vast experience on model intercomparison studies and dynamic and diagnostic model evaluation activities. MET Norway, TNO and INERIS have been coordinating (and BSC has participated in) the intercomparison studies aimed at benchmarking the EMEP model, currently in the framework of the TFMM (*4*–*6*). In addition, the consortium teams have coordinated and participated in European AQMEII phases I-III (TNO, FMI, INERIS) and intercomparison studies for China (FMI, TNO, MET Norway) and Africa (FMI, BSC), as well as a series of European studies (TRANSPHORM (FMI, MET Norway, TNO), COST-728 (FMI, MET Norway), series of CAMS pollen evaluations (FMI, MET Norway, TNO, INERIS), etc. BSC is responsible for the evaluation of dust over the Mediterranean area within CAMS_84 on Global and Regional validations and the evaluation of the desert dust forecast ensemble of the WMO SDS-WAS activities. The link with the CAMS-84 global validation team will allow understanding the consistency between the global and the regional aerosol components of CAMS.

The work package 6110 on model evaluation is coordinated by MET Norway. MET Norway hosts one of the technical centres of the EMEP program since several decades and has been developing the EMEP MSC-W model in the perspective of meeting policy objectives and scientific challenges. Therefore, the MET Norway team is acknowledged in Europe as pioneer in the production and interpretation of air pollution products such as overview and evaluation of the air pollution situations (including assessment of photo-oxidants, PM and its different chemical components, depositions). Furthermore, MET Norway also co-ordinates the international AeroCom initiative, and contributed to the last IPCC reports with model simulations and lead author contributions, in particular in the field of aerosols. Considerable developments in tools and observation data understanding had been developed as part of the AeroCom work.

Data assimilation*:* The partners in the consortium have long term experience in operating data assimilation systems for the provision of air quality analyses. The data assimilation systems used by the consortium cover a wide range of data assimilation techniques: Ensemble Kalman Filter (EnKf) (TNO, FMI), Ensemble Transform Kalman Filter (BSC), Ensemble Adjusted Kalman Filter (INERIS), 3D-var (MET Norway, FMI), 4D-var (FMI) and optimal interpolation (BSC, INERIS). This coverage of different techniques will ensure that the developed tools within the proposed work are tested and made suitable for application within the majority models in the CAMS regional air quality service. TNO, MET Norway, FMI and INERIS all have experience with the assimilation of ground-based observations in an operational set-up. In addition, all teams have experience with the (operational) assimilation of large set of satellite observations from a.o. OMI (NO_2, SO_2), MODIS (AOD), IASI (NH_3 and AOD) and even synthetic observations (NO_2 and HCHO) from Sentinel 4 and 5p (see (*2,7–9*) and track records in previous section). The active data assimilation techniques utilised by FMI and TNO allow the inversion of emission information, among which restoration of temporal and spatial patterns in emissions.

Assessment of new satellite products: TNO has a wide experience in the application of Observing System Simulation Experiments (OSSEs) for the assessment of the impact of future new satellite products for air quality applications. Within the ESA ISOTROP project TNO and FMI have been involved in the evaluation of the added value of Sentinel 4 and 5p products. This project involved the production of synthetic datasets and subsequent assimilation within regional model systems.

Production of generic and model agnostic tools: The teams in the consortium have experience in the production of code and tools for use by external parties**.** MET Norway, FMI, TNO and INERIS all develop their own open source models. FMI produces the modules for pollen emissions that are implemented by the CAMS regional systems. In addition, FMI has experience in the production of observation operators to be used by multiple models from the EUNADICS-AV project. TNO and MET Norway have built and share a TROPOMI pre-processor that can be used as a basis for the pre-processing tool to be developed in CAMS_61. Key personnel at INERIS have experience with the OOPS framework that will be used in CAMS_61 for the development of model-agnostic tools.

3.2 Key personnel qualifications

The Prime Investigator, Prof. Dr. Martijn Schaap, has about 20 years of experience in air quality modelling, data assimilation and management. He is one of the leading experts in Europe on the development and application of chemistry transport models for policy support and scientific applications. He coordinates the development of the LOTOS-EUROS air quality model at TNO. Under his lead a data assimilation capacity was developed for in-situ and remote sensing observations of AOD, PM10, SO_4, SO_2, NO_2 and NH_3. He has coordinated the EU FP7 project EnerGEO and contributed to numerous other EU or large international projects such as AIR4EU, MACC, CAMS, PASODOBLE, MEGAPOLI, MarcoPolo and ESA-GlobEmission. There is a strong emphasis on the evaluation and subsequent improvement of emission data in his research.

To manage the CAMS_61 project Martijn Schaap will be assisted by two senior scientists, Dr. Renske Timmermans and Dr. Arjo Segers, and a senior project manager, Peter Coenen. Renske Timmermans has over 20 years of experience in air quality monitoring and is the coordinator of the development activities for LOTOS-EUROS in the CAMS_50 project. She is in the scientific committee of GLOREAM which is a forum for European air quality modelling teams to discuss ongoing model developments and strategies for model improvements. She has coordinated many (inter) national projects and was work package leader in the EU project PASODOBLE. Renske is also one of the leading experts on air quality OSSEs, focusing on the data assimilation of future satellite products. Arjo Segers is seen as one of the world leading experts on data assimilation and through his work at different institutes and on different models, he is experienced in the use of different assimilation techniques. He has been developing the ensemble Kalman filter data assimilation in LOTOS-EUROS and applied this in a wide variety of applications. In CAMS_73 he is responsible for the emission inversion of CH_4 concentrations using data assimilation in the global TM5 model. Peter Coenen has over 20 years of experience in project management at TNO. His senior managerial expertise is deployed in complex international studies within the department of Climate, Air and Sustainability.

Dr. Hilde Fagerli (MET Norway) will be responsible for the WP6110 (In-depth assessment of the CAMS regional systems). She has a long experience in air quality model development, application and evaluation through her work for the UN LRTAP convention for almost 20 years, the last 10 years as the head of EMEP/MSC-W. Since 2009, she has been the leader of the Climate Modelling and Air Pollution

Division of MET Norway. She has been actively participating in the regular evaluation of the EMEP MSC-W model within the frame of the annual status reports to the UN LRTAP convention and multi-model inter comparisons exercises under EMEP/TFMM, the screening of EMEP measurement data, the use of Airbase data for model evaluation, and the continuous development of the EMEP MSC-W model with focus on chemistry of sulphur and nitrogen containing species. She has also experience in using the AeroCom tool for model evaluation.

Prof. Dr. Mikhail Sofiev will be responsible for the WP6120 (Coupling of regional forecasts and analyses). He has over 25 years of experience in atmospheric composition research and model development. He is the PI of the SILAM modelling team, coordinated the FMI modelling work in numerous international projects, and was a work package leader in the EU- PASODOBLE project. He is an expert in the data assimilation technologies and one of the key authors of variational algorithms in SILAM, both in their classical and extended forms. His latest interests concentrate on the fields of aerosol modelling, non-classical data assimilation and biogenic and natural emission modelling including inverse problems solution, all highly relevant for the current project.

Dr. Augustin COLETTE (INERIS), has about 15 years of expertise in the field of atmospheric science in support to policy and private users. He leads the Air Quality Modelling team of INERIS. He is Chair of the EMEP Task Force and Measurement and Modelling, leading him to lead multi-model inter-comparison projects such as Eurodelta and communicating the outcome to United Nations levels policy makers. Augustin COLETTE is currently part of the coordination of the CAMS_50 Regional Production Service. In that project he has the role of Scientific Advisor to coordinate the development activities of all regional models involved, therefore he benefits from a unique insight of the State of the Art of those models. He will have a crucial role in making sure the diagnostics performed in CAMS_61 will be transferred to CAMS_50 operational production systems.

Dr. Maria Teresa Pay (BSC) has over 12 year of experience in developing and evaluating of air quality models at regional scale. She played a seminal role on the successful implementation of the first air quality forecast system at high spatial and temporal resolution for Spain (CALIOPE system). She is a EURODELTA member contributing to the Task Force on Measurements and Modelling under the UNECE, and is part of the GLOREAM scientific committee since 2019. Under CAMS_50.II, she is in charge of the MONARCH evaluation. Under CAMS_61, she will be the coordinator of the BSC activities, and will be in charge of the evaluation of photochemical pollutants in southern Europe.

An overview of the qualifications of the key persons involved for each organization in the execution of the proposed project is provided in the table below.

Table 3-1: HR Profiles, the total effort in months fits the numbers in the pricing tables except for some small differences related to contributions from non-key personnel.

Title	Broad description of work in relation to Service	List of personnel who fit the profile and whose CVs are submitted with tender	Qualifications	Effort / engagement in months
TNO				
Contract manager and Senior scientist	Prime Investigator	Prof. Dr. M. Schaap	Professor and nearly 20 years of experience in air quality modelling, PI on EU & other (inter) national projects	**4.6**
Senior project manager	Service manager, leader of WP6100	P. Coenen	30 years of project management experience	**2**
Senior scientist/ project manager	Deputy coordinator, coordination of TNO activities and contribution to WPs 6110, 6120	Dr. R. Timmermans	PhD and 20 years of experience in air quality and satellites, 10 years of experience in project management, contributor to CAMS_50	**6**
Senior scientist	Leader of the WP6130	Dr. A. Segers	PhD and expert on data assimilation, contributor to CAMS_50	**12.5**
Senior scientist	Scientific advisor with relations to emissions	Dr. H. Denier van der Gon	PhD and expert on anthropogenic emissions	**0.5**
MET Norway				
Senior scientist / work package leader	Leader of the WP6110 In-depth assessment of the CAMS Regional Systems	Dr. Hilde Fagerli	PhD and 20 years of experience in air pollution modelling. Head of EMEP MSC-W (technical centre of the EMEP Programme)	**8**
Senior scientist	Contribution to WP6110 (including link to CAMS84 on above surface evaluation)	Prof. Dr. Michael Schulz	PI on several EU & Norwegian projects, experience in CAMS84, CAMS71 and AeroCom	**3**
Senior scientist	Contribution to WP6110	Dr. David Simpson	EMEP expert in model development and evaluation	**6**
Senior scientist	Chemical data assimilation model	Dr. Alvaro Valdebenito	Main technician for EMEP MSC-W in	**9**

	runs (WP6120 and WP6130)		CAMS regional production (and all precursor projects), expert in data assimilation	
FMI				
Senior scientist, team leader	WP 6120 leader, FMI team leader	Prof. Dr. M. Sofiev	Res. Prof. in AC dept. of FMI, Adj. Prof. of physics	**2.5**
Senior scientist	Data assimilation	Dr. R. Kouznetsov	PhD in atmospheric physics	**10**
Senior scientist	Modelling experiments	Dr. A. Uppstu	PhD in theoretical physics, experience in EnKF development	**10**
Senior scientist, team manager	Data assimilation and fusion	Adj. Prof. Dr. A. Karppinen	Dr. Sci. on Tech., Adj. Prof. of physics	**0.5**
INERIS				
Scientific team manager	Model evaluation	Augustin Colette	PhD Atmospheric Physics and Chemistry	**1**
Senior scientist	EAKF Assimilation	Gaël Descombes	MSc atmospheric physics	**2,8**
Senior scientist	Operational Forecasting system	Frédérik Meleux	PhD Atmospheric Physics and Chemistry	**1**
Scientist	Satellite observations	Anthony Ung	PhD Atmospheric Physics and Chemistry	**2,5**
BSC				
Scientist	Scientific supervisor of BSC activities	Oriol Jorba	Air quality modeler and main developer of the MONARCH model.; contributor to CAMS_50.II, CAMS_81 and CAMS_84	**0.5**
Scientist	Coordinator of BSC activities, involved in WP6110 evaluation of pollutants in southern Europe	María Teresa Pay	Air quality modeler expert; contributor to CAMS_50.II	**4.0**
Scientist	WP6110 Evaluation work lead	Dene Bowdalo	Air quality modeler expert and air quality data analyst; contributor to CAMS_50.II	**6.0**
Scientist	WP6110 mineral dust evaluation lead	Sara Basart	Dust evaluation expert; main contributor to	**1.7**

			CAMS_84 phase 1 and 2	
Scientists	Test observational operator and production of synthetic S4 dataset in WP6130	Enza Di Tomaso	Aerosol data assimilation expert	**0.8**
Scientists	Test observational operator and production of synthetic S4 dataset in WP6130	Jeronimo Escribano	Aerosol data assimilation experts; main contributor to CAMS_43 phase 1 and 2	**4**

3.3 CV's of Key Personnel

The CV's of all key personnel are provided in Annex 1 (separate document).

4 Technical Solution Proposed

4.1 Introduction

Air pollution is associated with adverse effects on human health through population exposure to particulate matter, nitrogen oxides and ozone, loss of biodiversity through acidification and eutrophication, decreased crop yields as well as climate change through interactions of short-lived climate forcers with the earth's radiation balance and carbon and nitrogen cycles. Furthermore, atmospheric composition may affect solar power generation, transport safety and UV exposure. The Copernicus Atmosphere Monitoring Service aims to support policymakers, business and citizens with robust atmospheric environmental information to address these issues.

Within CAMS the regional air quality service provides daily near real-time European air quality analyses as well as 96-hour forecasts with a multi-model ensemble system. Annually, reanalyses providing consistent annual datasets of European air quality are provided, supporting in particular policy applications. In addition, a subset of the modelling systems is used to provide policy relevant information on the origin of air pollution in European cities. To support stakeholders with robust information, the quality of the information needs to be quantified and improved where possible. To better meet the requirements and expectations from the users, the quality of the regional air quality service should be continuously increased by improving the model descriptions and their input data as well as through increasing use of data assimilation. Within this project we aim to provide the underpinning information and experience to guide the development and upgrade of the regional systems delivering to the CAMS regional production.

4.2 Understanding and goals

Air pollutant concentrations are determined by a delicate balance between emission strengths, physical and chemical conversions, transport, and wet and dry depositions. Chemistry Transport Models (CTMs) are complex computer models aiming to comprehensively describe the fate of the relevant pollutants and their interaction. In Europe several CTMs have been developed over the past decades, most of which are currently participating in the CAMS regional ensemble. The ensemble approach is motivated by the recognition within the EURODELTA model intercomparison studies, and within CAMS, that the ensemble median almost always outperforms the individual model systems. Figure 2 provides an example of the model evaluation of surface NO_2 forecasts for the last three months. It can be seen here that the ensemble shows lower RMSE than any of the individual models.

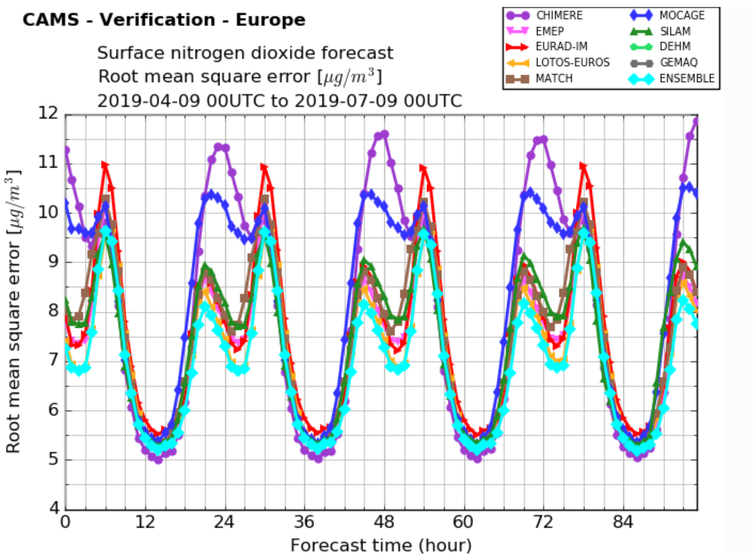


Figure 2. Verification results (root mean squared error) for NO_2 forecasts for CAMS regional air quality service models and ensemble over a three-month period. ***(source: http://macc-raq.copernicus-atmosphere.eu/)***

The different regional models are supposed to use the same input data regarding emissions, meteorological data and boundary conditions. The spread in model performance and the variable performance of each model in comparison to the others illustrates that the process descriptions in the different modelling systems have their strengths and weaknesses. The common behaviour of the models in some cases, illustrates that there are also common challenges to tackle.

Verification results of the regional air quality service illustrate that in some cases the regional model ensemble displays systematic errors with large differences to observations. In some other cases the spread between the models is large. This is illustrated in Figure 3, where during the first half of April 2019 all model forecasts seemed to underestimate the observed PM10 levels while between 22 and 29 April 2019 the models showed a large spread in bias. The latter has been identified as a dust episode. By analysing the archived model output from CAMS_50 Regional Systems in detail it is possible to identify periods and regions of large errors and/or large model spread, also considering elements like daily and weekly profiles, the forecast length, as well as vertical and horizontal patterns. However, in many cases it will be difficult to identify the exact causes of the varying model performance from these archived results only. For example, the underestimation in the first half of April could be related to errors in the input data (e.g. emissions or emission time profiles), or due to difficulties in representing the meteorological situation during this period. Also, the error might be noticeable throughout the entire domain or only in certain parts of Europe. For the large spread between the models in the second part of April, possible reasons could be differences in deposition process descriptions, in vertical mixing or in the implementation of boundary conditions. An in-depth evaluation of these situations is required to identify the exact cause of the systematic errors and the large spread.

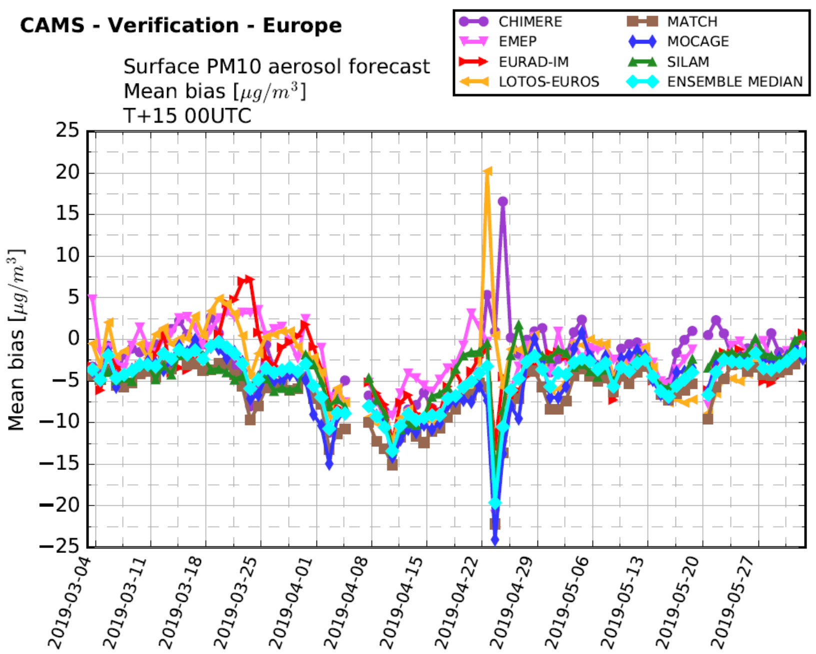


Figure 3. Mean bias of individual models and ensemble with observations of PM10 [µg/m3] for the period of March-May 2019. (source: http://macc-raq.copernicus-atmosphere.eu/index.php?category=verification)

Intercomparison studies and especially those in which the analysis was performed by the model developers themselves (e.g. EURODELTA and AQMEII) have turned out to be a very efficient way to identify and solve issues in individual systems as the researchers involved learned from each other's experiences. Hence, **we aim to provide an in-depth assessment of the inter-model differences between the Regional Systems and their performance against observations**. To be able to diagnose the reasons behind the model output differences we feel it is of utmost importance to have more detailed information available then currently reported operationally by the CAMS regional systems. In addition to analysing the standard model output, we propose to organize an open and inclusive model intercomparison for 2018 in which parameters required for diagnostic evaluation are requested. Based on the analysis of the results we will **provide a priority list for general and Regional System-specific Research & Development activities to upgrade the quality of the regional systems in the short and mid-term future**.

The quality of analyses and forecasts could also be improved through data assimilation. The current CAMS regional (re-)analyses are mostly based on the assimilation of in-situ observations provided by the EEA. The assimilation of in-situ ground-based observations shows a clear improvement in the modelled results with values closer to the (independent) observations. In addition, three of the models contributing to the CAMS regional service assimilate satellite observations (NO_2 and SO_2 columns from OMI and GOME-2 and CO columns from MOPITT and IASI). However, due to the timeliness of European air quality observations and the timeliness required by the users (CAMS forecasts need to be available before 09UTC), the CAMS regional analyses and forecasts are currently uncoupled for nearly all models involved (EMEP forecasts are the only exception being initialised at D-1, 00 UTC with the last hour output of the analysis of D-2).

There is limited literature on the impact of optimizing initial conditions on air quality forecast performance but it has been shown that the impact tends to disappear with time (e.g. *9–15*). From the available cases, it seems that in the regions and for the species strongly impacted by emissions this window is the shortest whereas long-living species and/or clean areas are characterised by longer relaxation time. Here, **we aim to thoroughly evaluate the benefit of basing regional forecasts on analyses of surface observations**. Specific attention will be given to identify for which conditions the impact is largest and optimize the assimilation strategy accordingly. The few available studies indicate that optimizing initial conditions positively affects forecasts for a limited time frame up to about 12-36 hours (*10,11,14,15*). An example of the model relaxation after initialization of SO_x compounds is shown in Figure 4 for a randomly picked episode in Europe (*12*). To preserve the impact of data assimilation beyond this time horizon **we will in explore the potential for inheriting parameter estimates into the forecasts**. This is anticipated to be more effective for reactive short-lived gases such as NO_x, whereas updating non-reactive (particulate) compounds should benefit from optimizing the initial conditions.

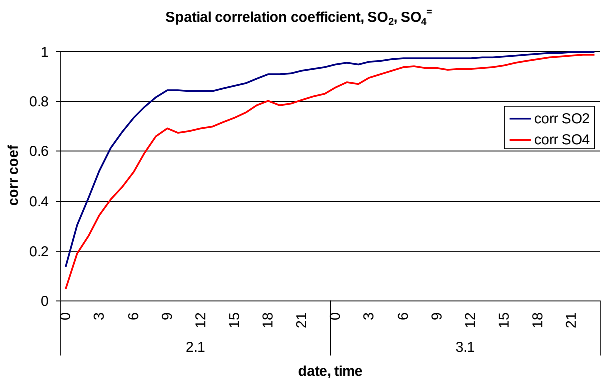


Figure 4. Evolution of the spatial correlation coefficient between a run with initialisation of SOx compounds and a run without. t=0 represents the end of pre-computation and start of the unconstrained 2-day run of the model. Adopted from (*12*).

Apart for the surface observations analysed regularly by the CAMS models, satellite data plays an increasing role. Assimilation of satellite data in addition to surface observations is not always efficient for improving modelled concentrations at the surface. Besides the mismatch between the observed columnar values and surface concentrations, one of the reasons for this is the fact that the satellite observations used so far are mostly from polar-orbiting satellites that deliver daily observations only for overpass time of the satellite and during clear sky conditions. With the upcoming geostationary Sentinel-4 mission the temporal coverage of available satellite observations for air quality will improve considerably. This satellite will provide observations with a spatial resolution similar to the currently operational polar orbiting Sentinel-5p mission, but with European coverage at an hourly resolution. An OSSE study performed within the ESA ISOTROP project showed that the Sentinel-4 NO_2 observations with hourly temporal resolution have an added value for the NO_2 analyses throughout the entire day, while the daily Sentinel-5p NO_2 observations have a slightly lower impact that lasts up to 3-6 hours after overpass (Figure 5).

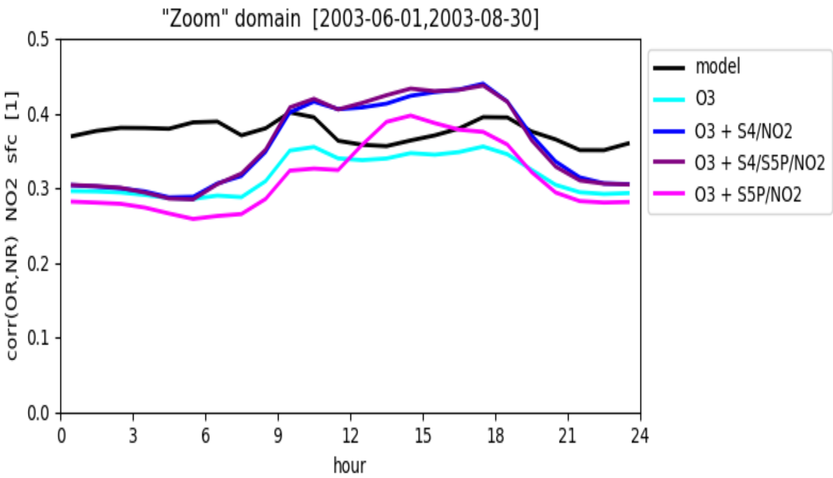


Figure 5. Correlation between modelled surface NO_2 concentration and assumed true concentrations for a three-month summer period; without assimilation (black line), and after assimilation of observations (coloured , ground-based O_3 only (cyan), ground-based O_3 + Sentinel 4 NO_2 (blue), ground-based O_3 + Sentinel 5P NO_2 (pink), or ground-based O_3 + Sentinel 4 and Sentinel 5p NO_2 (purple)).

The arrival of new high-resolution datasets does however also provide challenges:

- As with other satellite products, they are not directly comparable to the surface air quality variables of interest, for example due to the difference between columnar values from satellites in comparison to surface concentrations. To overcome this issue, the modelled values need to be converted to a value as observed by the satellite taking into account the spatial and vertical coverage of the satellite data product as well as the sensitivity of the instrument to different altitudes in the atmosphere.
- Given the high temporal and spatial resolution of new instruments the amount of data is very large. To allow storage and use of the data in operational systems, the data volume needs to be decreased in a smart way without losing relevant information.
- The uncertainties defined for the model, the surface observations, and the satellite observations should be in balance with the differences between simulations and observations. With the massive amount of Sentinel 4 observations it becomes more challenging to maintain this balance at every location and every hour during the assimilation.

All modelling teams will need to deal with these challenges when willing to incorporate the new satellite data in their data assimilation systems. However, many of the activities required to overcome these issues are independent of the model and/or data assimilation system used. To boost the data assimilation developments in the CAMS regional production we will **deliver modular "model-agnostic" elements for performing data assimilation of future Sentinel-4 observations in the CAMS Regional Systems.** The model-agnostic tools will be tested through the assimilation of currently available Sentinel-5p observations in some of the CAMS regional systems in the project. This will provide **guidelines for the configuration of assimilations including both surface and of Sentinel-4-type of observations**. As an additional benefit this will **also explore the use of Sentinel 5p to improve initial conditions and forecast performance for the CAMS regional systems.**

The new satellite instruments will also allow an advancement in the use of satellite observations for emission estimates. Emission estimation through data assimilation is an important process that may constitute one of the most promising perspectives for improving air quality forecasts. The availability of hourly observations will most likely improve the capabilities for retrieving information on temporally varying emissions. **We will evaluate the potential of using satellite observations, specifically from the Sentinel 4 mission, for emission estimates and propose strategies for developing this type of applications.** Here, a close cooperation between satellite emission inversion experts and emission experts with experience on bottom-up emission estimates is deemed vital.

In the next sections we will describe the specific activities planned to reach the above-mentioned goals. The regional model systems and data assimilation techniques that will be used in the proposed project are listed in Section 4.6 Table 4-2. Through being part of the CAMS regional service, **the considered models and assimilation systems all fulfil the general requirements set in the ITT.**

4.3 Work package 6110: In-depth assessment of the CAMS Regional Systems to identify future development needs

CAMS is operating an ensemble of nine regional model systems to provide its regional air quality services (two further candidate models are planned to join this ensemble in 2022). All these model systems are driven by ECMWF's high-resolution operational meteorological forecasts and they use the same anthropogenic emissions, fire emissions and chemical boundary conditions. The differences in results should therefore solely come from different formulations of resolved and sub grid scale transport processes, physical and chemical processes, as well as removal processes in the various model systems. This work package is to investigate the reasons behind common and model-specific deviations from observations, related to model formulations, and the treatment of physical and chemical processes, such as:

- gas-phase chemistry, microphysics of aerosols, the number of VOC and PM species represented in the models;
- wet scavenging and dry deposition schemes; e.g. dry deposition in some models makes use of canopy resistance and conductance, with explicit calculation of stomatal conductance, while others use less sophisticated schemes;
- read-in of emissions and other input data, e.g. some models use VOC speciation directly from the CAMS emission data set, while other models split VOC emissions based on other data sources;
- injections heights and temporal profiles of emissions used in the models;
- emissions that are calculated online in models, e.g. biogenic volatile organic compounds (BVOC), NO_x from soil and NO_y from lightning, etc.;
- the way of using chemical boundary conditions from the CAMS Global Product that differs among the models;

- surface layer thickness, boundary layer height, and vertical structure;
- method of putting out concentrations at observation height, e.g. as concentration in lowest layer or inferred from the lowest layer concentrations and dry deposition velocities;
- turbulent mixing schemes of models, e.g. how K_z is computed and whether convection is included (and if so, how).

The work will be organized in three tasks: Task 6111 for the acquisition of data and documentation, Task 6112 for the evaluation of CAMS_50 operational output, and Task 6113 to perform dedicated model experiments and an extended analysis to consolidate the list of recommended developments, which will be the main output from this work package. Importantly, WP6110 will also organize three meetings with ECMWF and the CAMS_50 development teams, where the model evaluation and the suggested model developments will be discussed in detail. The meetings will also assist the planning of (and encourage the participation in) the dedicated model experiments.

4.3.1 Task 6111: Acquisition of model data, observational data, and model documentation

In this task modelling data from the CAMS_50 model database hosted at Meteo France will be acquired for the period specified in the call (09/2018 to 09/2019), along with all available measurement data. Archived forecast data from the 7 Regional Systems for NO, NO_2, O_3, CO, SO_2, $PM_{2.5}$, PM_{10} are available from the beginning of that period but, since June 2019, also include output of dust aerosols and secondary inorganic aerosols (as well as output for 2 additional operational models). Output data will be collected in all 8 vertical levels, reaching up to 5km above ground.

In June 2019, the CAMS Regional model systems were upgraded, including the move from using TNO-MACC-III to CAMS-REG-AP_v2.2.1 anthropogenic emissions. In addition, the individual Regional Systems were upgraded, including updates of model specific routines and bug fixes in the individual systems. For the proposed assessment, it is important to know exactly which changes were made in that upgrade; thus the 'Request for Change' documents provided by the CAMS_50 model teams need to be scrutinized. Furthermore, detailed descriptions of the individual models are needed to better understand the different behaviour of the different models. Such descriptions are being made within CAMS_50, and we expect that they will be available for the work under this contract.

The evaluation to be done in this work package will, to a large extent be based on comparisons of the Regional Systems to the EEA AQ e-reporting/EMEP observations, going beyond the observational data sets extracted in CAMS_50. For quality assurance, the observational data for the evaluation period 09/2018 to 09/2019 will be screened using the GHOST System (Globally Harmonized Observational Surface Treatment) at BSC. The central concept of GHOST is the standardisation of the data/metadata from all major public reporting networks which provide atmospheric surface measurements. Each processed measurement is additionally associated with quality assurance/classification flags, which pertain to a plethora of documented quality control checks/metadata groupings. All quality control checks and classification groupings are identified with codes, so that every team knows exactly how the measurements have been subset. The idea here is that every team will use exactly the same data set, and cross-team/cross-experiment evaluation efforts can be directly compared.

By summer 2020, chemically speciated data of PM will become available for the year 2018 from the EMEP network. These data are quality controlled within EMEP in a process where EMEP MSC-W/MET Norway is participating, giving first-hand information about the quality of the measurement data and the problems associated with them. These data will complement the data from e.g. the EEA AQ e-reporting in the extended analysis of Task 6113. Furthermore, a joint EMEP/ACTRIS/COLOSSAL intensive measurement period (IMP) was conducted for the winter season 2017-2018 and provides a harmonized European-wide data set of EBC_{ff} (equivalent black carbon from fossil fuel) and EBC_{bb} (equivalent black carbon from biomass burning) applicable for model evaluation of the dedicated runs for 2018. Out of the 57 sites contributing to the IMP, 27 sites also measured levoglucosan, a tracer for wood burning.

4.3.2 Task 6112: Evaluation of the operational CAMS regional forecast data with extended statistical analyses

An overall assessment of the performance of the ensemble and the individual components compared to observations will be made, where performances will be analysed separately for different regions, all seasons, and for all chemical components within the period 9/2018 to 9/2019 specified in the call (Phase 1 evaluation, see Figure 6). The temporal profile (hour of the day, day of the week) of the ensemble and the individual models will be compared to corresponding observations to identify significant systematic differences in their diurnal and day-of-week profiles.

Obviously, the analysis of surface concentrations will also build on the evaluation done regularly by CAMS_50 in the quarterly NRT production reports. From there we have already learned that, e.g. ozone is overestimated by most of the models, while NO_2 and PM are mostly underestimated. Disagreements vary with seasons (e.g. correlation in PM being better and its underestimation worse during winter than during summer, and the positive bias in ozone tends to be lower in summer), but also with time of the day and with forecast day (e.g. the underestimation of NO_2 usually peaks around noon, and correlation of all species degrades with forecast day, while other scores (bias and RMSE) degrade much less. These observations can be used to give hints for possible model improvements.

The analysis to be performed in this task will extend the existing evaluations in CAMS_50 to a more detailed analysis which consists of looking at models' performances during high pollution events, different regions of Europe, comparison of frequency distributions of concentrations, comparisons for specific PM chemical components, etc.

We will be using tools that have been developed and maintained for regular model evaluation within the EMEP programme over many years. In addition, we will use and extend the AeroCom tools behind the AeroCom web interface (aerocom.met.no). These tools are aimed at multi-model evaluation of atmospheric aerosol and chemistry models interpreting different database extracts from Aeronet, Skynet, EBAS, EMEP, ACTRIS, GAW, Earlinet, NDACC, EIONET and satellite data.

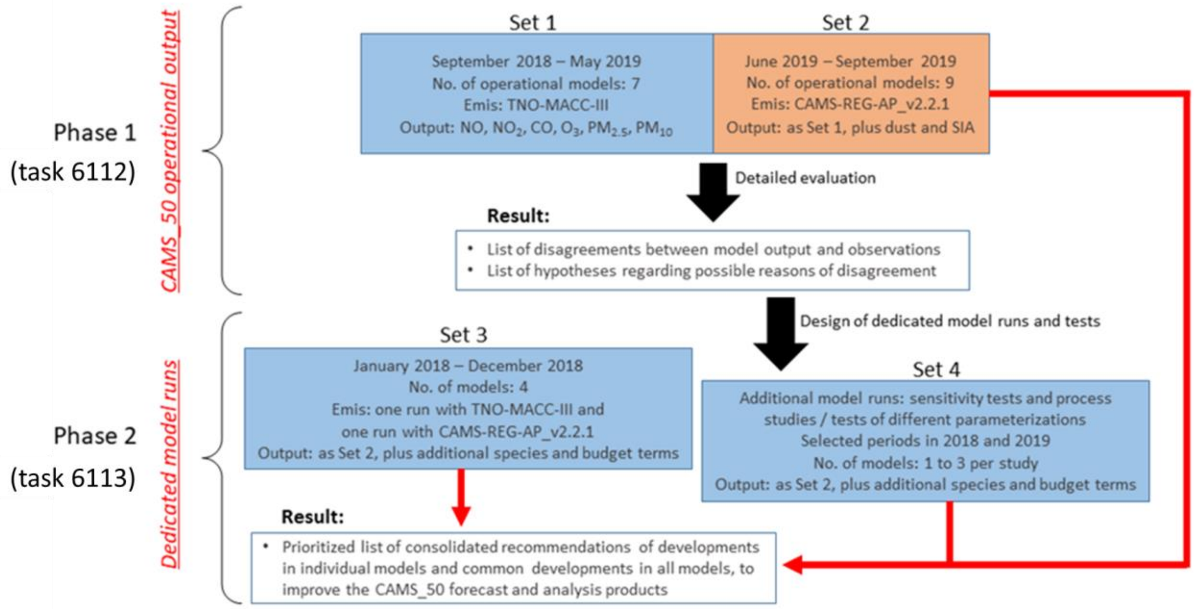


Figure 6 Schematic of the evaluation in Tasks 6112 and 6113 of Work Package 6110. The details of Set 4 will depend on the outcome of the evaluation of Sets 1 and 2, and thus cannot be planned in detail yet. Typically, between one and three CAMS_61 models will participate in the simulations for Set 4.

The analysis of model results versus observations will consist in comparing output from individual model systems to observations for different species, in different regions, altitudes, seasons, etc., and focus on identifying situations where:

- the regional model ensemble is systematically different from the observations; common problems of the Regional Systems will be identified, and common developments will be proposed to resolve them;
- either the spread of the ensemble is particularly large, or for which one specific Regional System is significantly different from the other Regional Systems; this analysis will identify challenges in specific Regional Systems and will be used to identify model specific developments.

All the available pollutants will be considered (NO, NO_2, O_3, CO, SO_2, $PM_{2.5}$, PM_{10}, dust aerosols and secondary inorganic aerosols), as well as different lead times of the forecasts.

In addition to the overall assessment, the evaluation will pay specific attention to events with high air pollution concentrations. These episodes can have severe health effects and receive considerable public attention. In addition, episodes magnify and emphasize important processes (e.g. chemistry, meteorological conditions, sources of emissions), and enable a better understanding of underlying problems in the model systems or input data. Therefore, the performance of the CAMS ensemble and the specific models during events with high air pollution will be investigated, for instance:

Reactive pollutants in Southern Europe. It is well established that high O_3 concentrations and photochemical activity are particularly pronounced in southern/Mediterranean countries due to the meteorological conditions favouring its formation. The coupling between synoptic and mesoscale processes governing O_3 and secondary aerosols in the Mediterranean Basin needs special analysis in order to understand current model performances.

Ozone episodes in Central Europe. Very high ozone concentrations are experienced also in Central Europe every summer, in connection with stagnant high-pressure systems, high temperatures and intense sunlight. It is important to assess how well the CAMS_50 models capture these ozone episodes, and when and where major discrepancies occur.

Dust Aerosol contribution in Southern Europe. Mineral dust events will be identified and evaluated combining the information from a set of observational databases that include column-load and surface information (i.e. AERONET, MODIS and the ratio $PM_{2.5}/PM_{10}$). The link with the CAMS_84 global validation team will allow to understand the consistency between the global and the regional aerosol components of CAMS.

PM episodes in late winter/early spring in Central Europe. Central-Western Europe and especially the BeNeLux area often experience high levels of $PM_{2.5}$ during late winter/early spring when increasing ammonia emissions from agriculture react with HNO_3 to form ammonium nitrate. Although it is not possible to distinguish ammonium nitrate from other SIA in PM (due to lack of chemical speciation in the operational CAMS_50 output), this first assessment will evaluate the performance of the ensemble and individual models for these episodes.

PM and NO_2 episodes in winter. Major PM and NO_2 pollution episodes often occur during winter, caused by a combination of stagnant air, inversions, and enhanced use of wood burning for residential heating (or high loads of road dust resulting from the use of studded tires in the Nordic countries).

The results from the analyses in Task 6112 will be used to provide an overview of situations for which the regional ensemble is significantly and systematically different from the observations and situations for which the spread of the ensemble is large, or for which one specific Regional System is consistently different from the other Regional Systems. The analysis will be used, to the extent possible, to explain (or to make hypotheses of) the reasons behind the deviations identified.

The variance of the ensemble and the temporal auto-correlation for the main pollutants, in the vicinity and away from large pollutant emission regions, will be investigated based on methods described in (*16*–*19*). Vertical information for relevant pollutants will be used as well, to complement the surface evaluation. Given that the models have different vertical resolution, the vertical auto-correlation error can be diagnosed using the ensemble as a reference and provide important information also for the assimilation of satellite products done in WP6130. The analysis of spatial length scales will also be extended horizontally to assess the influence of boundary conditions. We will apply first a spectral decomposition analysis of the bias, variance, correlation and mean square error to identify the most important time scales (long-term, synoptic, diurnal or hourly) that can then be related to the time scale of driving forces in the AQ forecast (meteorology, emissions, boundary conditions). Subsequently, a statistical selection and classification of the same set of precursors will help us identify and quantify the sensitivity of the different models. In this case, a basic statistical model is fitted to each member of the CAMS_50 ensemble to predict concentrations of a pollutant based on a restraint set of precursors as in (*20*). This will allow to investigate the sensitivity to parameters from the physics or the chemistry for each model during selected episodes. Parameters identified as being most important will receive more attention in Task 6113 and in the resulting list of recommendations for future development.

In CAMS_84, the concentrations above surface for the CAMS Regional Systems are evaluated in quarterly reports for O_3, NO_2, aerosols (PM_{10}/$PM_{2.5}$) and CO. The analysis in Task 6112 and also 6113 will complement the evaluation done in CAMS_84 in the sense that it will provide more information about possible reasons behind systematic differences between the models and the observations. Two of the partners in the proposed project are partners of CAMS_84 (MET Norway and BSC) and will ensure that we can draw upon and extend the results already existing in CAMS_84.

4.3.3 Task 6113: Diagnostic evaluation based on dedicated model runs

Based on the data available from the operational CAMS_50 production it will be possible to identify situations for which the ensemble or specific Regional Systems deviate systematically from observations. For some components, e.g. NO_2 and O_3, we expect that it will be possible to make recommendations for development based on Task 6112, whilst for PM additional analyses are necessary in order to be more specific on the developments needed. PM consists of a range of different chemical components, where for instance overestimation of one component can be compensated by underestimation of another.

Furthermore, quantifying the impact of the emission update implemented in June 2019 is complicated by simultaneous updates in model descriptions as well as different meteorological conditions prior to and after the CAMS_50 product update. The proper way to evaluate the emission update is to rerun a predefined year with both emission data sets.

Therefore, we propose to perform a set of consistent model reruns (also to be coordinated with WP 6120), where the Regional Systems use their configurations as of today. These reruns will be used to assess in more detail the reasons behind the systematic differences documented in Task 6112. An essential feature of the reruns will be additional model output, e.g. chemically speciated PM, all relevant precursors of PM and O_3, as well as dust, sea salt, AOD, EC_{ff} and EC_{bb} (elemental carbon split into fossil fuel and biomass burning origins), OC_{ff} and OC_{bb}, and other primary particulate matter. Where relevant, the output will also include concentrations in precipitation, dry deposition, effective dry deposition velocities and more. Another important feature will be that the same models will be run twice (at least), with different parameterizations, in order to isolate the effect/benefit of the change in parameterization or model setup. The model reruns and extended analyses include:

Model simulations with either TNO-MACC-III or CAMS-REG-AP_v2.2.1 anthropogenic emissions. A similar test was done in CAMS_50 for two months of 2016, but here it will be done for the entire year of 2018 (Set 3 in Figure 6) by the EMEP, LOTOS-EUROS, SILAM and CHIMERE models. The models will harmonize, to the extent possible, the way in which the emission data are used, e.g. in CAMS_50 some of the models used their own VOC split, while others used the one from CAMS-REG-AP_v2.2.1, resulting in a larger effect of the emission update.

Selected sensitivity tests for the entire year of 2018 or for shorter periods (Set 4 in Figure 6) will be conducted and analysed, in particular for episodes that in the previous task were found to be reproduced inadequately. The tests will be designed to verify or falsify the hypotheses brought forward based on the evaluation in Task 6112 regarding the reasons for the discrepancies (missing emission sources, resolution, model formulations, etc.). The tests will focus on:

Assessment of PM composition. The different chemical components of PM (SO_4^{2-}, NO_3^-, NH_4^+, dust, sea salt, EC, OC) from the model runs will be compared to corresponding observations from the EMEP network to evaluate the capability of the models to reproduce the chemical composition of PM.

Assessment of wood burning emissions. EC_{ff} and EC_{bb} will be compared to observations from the EMEP/ACTRIS/COLOSSAL Intensive Measurement Period (IMP), which provide information on the origin of the observed BC to evaluate the effect of underestimation of wood burning emissions on modelled winter PM. Runs will also be done with scientific emission data sets (containing, e.g. condensable VOC emissions for all countries), therefore providing scope for collaboration with the Task Force on Measurement and Modelling of EMEP, which is launching a model intercomparison project for that IMP (Eurodelta-Carb).

Budget assessment of PM species. To understand why some PM components are under- or overestimated, the 'budget' of the different PM components and their precursors will be analysed and compared to measurements of sulphur and nitrogen containing species. For instance, an underestimation of SO_4^{2-} could be related to too efficient washout of SO_4^{2-}, too high dry deposition of SO_4^{2-} or SO_2, or to too slow conversion of SO_2 to SO_4^{2-}. From the EMEP network, gaseous precursors, particle concentrations, concentrations in precipitation, and wet deposition are available and will be compared to the model output.

Impact of boundary conditions. Although the proposed project will not duplicate the evaluation of the global model (which is done in other contracts), effects of using its output as boundary conditions in the regional systems will be assessed. For example, boundary conditions of PM and CO have occasionally caused overestimations within the regional domain. Such problems can be caused either by biases in the global model or by inconsistent use of the boundary conditions. E.g. when a component is overestimated by a certain percentage in the regional models and we know from CAMS_84 that the global model overestimates that component a test will be run by selected regional models to see whether a corresponding scale-down of the boundary values helps remove the bias.

This above list is not exhaustive. Experiments will be added focussing on all aspects mentioned in the beginning of section 4.3, as deemed necessary based on the outcome of Task 6112 and the first experiments of Task 6113.

Several of the CAMS_50 models are represented in this consortium and will contribute to these tests. However, the experiments of Task 6113 will be open for all other CAMS_50 models as well and be announced in due time at appropriate CAMS meetings or by e-mail communication.

Based on the analyses in Task 6112 and Task 6113, three detailed reports on the assessment of each Regional System and inter-model differences will be delivered. The first report (after 6 months) will be based solely on the results from Task 6112, whilst report 2 (after 12 months) and report 3 (after 18 months) will also include the results from Task 6113. In addition, the assessment of the Regional Systems will make use of the lessons learned from WPs 6120 and 6130, where data assimilation of in-situ observations and Sentinel-5p and parameter estimation may provide further understanding of the deviations between the Regional Systems and observations.

The results will be presented and discussed at three meetings with ECMWF and the teams of the CAMS Regional Systems (milestones M1.2.1, M1.2.2 and M1.3.3). The first and third meeting (month 6 and 18) will be organized back-to-back with CAMS_50 meetings but can also be extended video conferences (though CAMS_61 partners will meet face-to-face in any case). The second meeting will most likely be back-to-back with a CAMS General Assembly (fall 2020). Also, representatives from CAMS_84 will be invited to these meetings.

Taking into account the feedback from each of the CAMS_50 Regional Systems' development teams and relevant items from the CAMS User Requirement Data Base (URDB), a prioritized list of recommended model developments will be provided. These recommended activities will target individual model systems as well as cross-cutting activities that are applicable to all model systems. The recommendations might extend beyond direct model developments and concern, for instance, the development of emission inventories. However, as requested in the ITT, the list will focus on developments that can be implemented within a 6 to 12-month period.

As a by-product of the data screening and of our detailed model evaluations in this work package, we will also propose a one-year benchmark set of observations that can be used for future evaluations of regional models (and developments) in CAMS.

4.4 Work package 6120: Coupling of regional forecasts and analyses

This WP will concentrate on maximising the benefits of assimilation of the model initial state for the forecast. As mentioned in the introduction, this approach is challenged in several literature references, which point out at the limited "memory" of the air quality modelling systems regarding their initial state. This fact is the main challenge of the current WP.

Addressing the above challenge, the **overall goal** of WP 6120 is to **evaluate the benefit of basing regional forecasts on analyses of surface observations** (responding to the Tender Specification item 3). Specific tasks aiming at the overall goal are three-fold:

4.4.1 Task 6121: Analysis of existing material on analysis usage for air quality models initialization.

A literature review will be conducted in order to collect and analyse the existing practices in the field of air quality data assimilation coupling to forecasts.

Attention will be primarily given to the "standard" approach of creating the analysis fields and then initiating the forecasts from them. We shall systematize the available literature on the problem, identify and analyse the existing approaches to the initial-value assimilation in the air quality problems. All such works reporting improvements of the forecasts in the short range, generally agree that the effect eventually disappears – but diverge in the estimated longevity of the improvement. These differences and the reasons behind will be highlighted and used in Tasks 6122 and 6123 for in-depth analysis.

This literature review will, first, highlight the species, conditions and regions where the initial-state assimilation is the most-efficient and has the longest impact on the following forecast. The second target will be the species, conditions and regions where the initial state is quickly forgotten, and the system evolution is governed by other factors, such as emission or boundary conditions. Having these two extremes in hands we will identify the reasons behind the specific system behaviour, cross-check the stability of the conclusions across different models, finally formulating the recommendations on the best practice for CAMS regional models. These recommendations will be further evaluated in the Tasks 6122 and 6123.

In addition to the "standard approach" we will also review literature on more novel approaches where not only the initialisation of the forecasts is based on assimilation results but also other parameters such as the emissions. The results of this part of the literature review will form the basis of task 6123.

4.4.2 Task 6122: Multi-model assessment of efficiency of analysis-based initialization

We shall perform a series of model experiments to assess the improvement in the forecasting skills against independent air quality observations, when an analysis is used as initial conditions for the forecast. The setup of the experiments will be based on the experience of the consortium members but also account for the recommendations of the literature review.

Within this task, three CAMS-50 operational models (SILAM, EMEP, LOTOS-EUROS) will be run through the full year of 2018 several times using the following setups:

(i) "standard" forecasting mode with previous-day forecast at +24 hours used to initialise the simulations; this is the most common way for the CAMS_50 models, not including the analysis into the forecasting chain. This year-long hindcast will create the reference dataset for the improvement.

(ii) analysis mode to generate the 00h, 06h, 12h and 18h analysis fields for each day of the year. Analyses will be generated using the same data assimilation techniques as in CAMS_50 for each model, i.e. 3D-Var for SILAM and EMEP and EnKF for LOTOS-EUROS using ground-based observations of O_3, PM, NO_2, SO_2 and CO.

(iii) forecasting mode initiated by the 00h, 06h, 12h and 18h analysis to initiate the forecast. From the analyses generated within the step (ii) , we shall pick the following species, separately for each run:

a. NOx and O_3
b. CO
c. SO_2
d. PM2.5 and PM10
e. all analysed species

The outcome of these simulations will be compared paying attention to different lead times (see also Task 6123), the spatial distribution of the effects in polluted / clean regions and cross-species impact via chemical links. Specific attention will be given to the impact on forecasted episodes with exceedances of limit values.

The episodes with particularly strong / weak effect of the analysis-based initialization will be reviewed through the performance of additional sensitivity tests with the Ensemble Adjusted Kalman Filter (EAKF) of the CHIMERE data assimilation system, aiming at the mechanisms behind and possibilities to improve the efficiency of the analysis usage for the air quality forecast improvement.

The conclusions will be compared with those suggested in the literature and summarised by the Task 6121. The resulting set of recommendations will be reported as guidelines for usage of analysis fields in the CAMS regional forecasts.

4.4.3 Task 6123: Analysis of episodic longevity and extent of forecast improvement for different assimilation strategies

We shall evaluate the longevity of the forecast improvement when analysis is used as initial conditions and compare the effect with the effect of using assimilation of other variables, such as emission or model parameters, on the forecasts.

For these episodes, the effect of the analysis-based model initialization will be compared with other emerging methods of the air quality forecast improvements: via assimilation of model parameters and/or emission fluxes. In these cases, the assimilation of observations leads to updated parameter/emission estimates, the updated parameters or emissions could then be prolonged into the forecasts. The episodes will be recomputed with the SILAM and LOTOS-EUROS model using 4D-Var and EnKF data assimilation techniques directed to the emission fluxes and/or model parameters rather than (together with) the initial conditions. The primary targets will be the real-time corrections to the anthropogenic emission inventories for NO_x, SO_2, and primary PM, as well as to the biogenic VOC emission model parameterization, which will be adjusted following the NO_x and O_3 assimilation. The results of this work will be taken into account in Task 6134 on the R&D plan for emission estimates.

4.5 Work package 6130: Towards assimilation of observations from geostationary satellite sensors (Sentinel-4) to constrain concentrations and emissions of main pollutants

Within the CAMS regional service, daily analyses as well as full-year reanalyses are delivered based on a large variety of assimilation methods. Many systems use three-dimensional variation (3D-Var): EMEP, EURAD, MATCH, MOCAGE, and SILAM. Outside the operational context some systems have four-dimensional variation (4D-Var) available too, for example EURAD and SILAM. Optimal Interpolation (OI) is used pre-operational by DEMH and GEM-AQ. Ensemble Kalman Filter and related methods are used by LOTOS-EUROS and (pre-operationally) by MONARCH, while the CHIMERE system will soon replace its current Kriging based approach by an Ensemble Adjusted Kalman Filter (EAKF).

As mentioned in section 4.2 the future Sentinel-4 (S4) mission is expected to become an important source of data for improvement of CAMS regional products. However the data assimilation of Sentinel-4 data will be challenging, since this geo-stationary mission does not only provide observations at high spatial resolution, but also at high temporal resolution (hourly during day time) for the entire European domain (Figure 7) leading to large volumes of data. At mid-summer conditions, a number of 175,000 pixels per hour could be reached, with a total up to 2,75 million for a day. Clouds and low observation zenith angles will reduce the number of useful pixels, but the volume will remain large.

Another challenge for assimilation of the satellite data in addition to surface data is to prevent a degradation of simulation skill at the surface by the analysis. Such degradation could occur if the surface and satellite observations (often columnar values) point in different directions for adjustment of the concentrations (e.g. (*2*)). Joined assimilation of surface and satellite data therefore requires careful configuration of vertical correlations, model uncertainties (especially related to the vertical profile), and observation representation errors. These are especially important to avoid that the large amount of satellite data overrules all information from the surface network. The representation errors

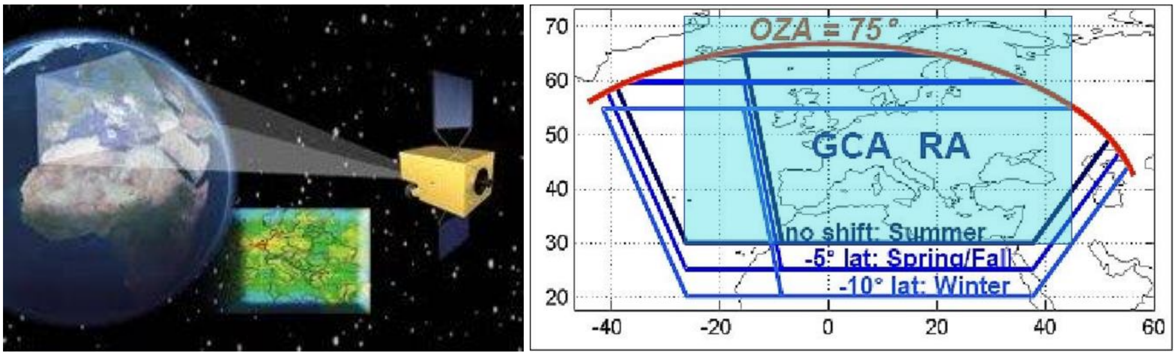


Figure 7 Artist impression of Sentinel-4 at geo-stationary position observing Europe. Shaded area on the right shows extend of the domain used by CAMS Regional Ensemble, which for a large part overlaps with the Sentinel-4 Geographic Coverage Area (GCA). The scan window is optimized to observe the smaller Reference Area (RA) when illuminated.

assigned to the two observation types should therefore be in balance with each other in case the observations are representative for the same part of the atmosphere.

Work package 6130 will set the first step towards analysis of the future Sentinel-4 data by the assimilation systems in the CAMS regional ensemble. A generic observation operator will be developed, tested and released that will facilitate the various operations in the assimilation related to the use of this data. Efficient handling of the large data volumes involved with Sentinel-4 will be a key aspect. In addition, we will also investigate the potential of using observations from geostationary satellites (specifically Sentinel-4 data) for emission parameter estimation as part of the data assimilation process. Emission parameters form one of the most important model uncertainties, and data assimilation could therefore help to reduce this uncertainty leading to improved air quality analyses. A Research and Development plan will be written that describes the current state of emission estimations and provide guidelines for further development to allow implementation, specifically for Sentinel-4 data, in the CAMS regional service.

The work is assigned to 4 tasks that are described below.

4.5.1 Task 6131: Generic observation operator for satellite data

To be able to assimilate observations within a CTM, an observation operator should be defined that simulates the observations from the 3D fields of trace gas and aerosol concentrations (e.g. (*21*)). In case of satellite data, the observation operator simulates what would be retrieved from the instrument in case that the true atmosphere would be exactly equal to a model state. This involves for example a horizontal mapping to pixel footprints, vertical mapping to retrieval layers, and application of the satellite averaging kernel which includes the sensitivity of the instrument to different altitudes.

In this task, a model-agnostic observation operator will be defined and tested for use by the CAMS Regional systems. The operator should in particular be able to digest data from the future Sentinel-4 mission but will be generic enough to also handle data from for example the current Sentinel-5p mission. Where possible the development will be based on existing generic observation operators, for example the operator for lidar- and aircraft-type datasets developed by FMI in the H2020 EUNADICS-AV project.

The first step in this task is the implementation of a pre-processor for satellite data. The pre-processor will be able to download, select, and convert satellite data into a common format and manageable volumes. The pre-processor will provide the following features:

- Selection on longitude-latitude range, cloudiness, quality flags, and variables of interest. This could greatly reduce the data volume.

- First-glance inspection of the data by creating sample plots and summary tables.
- An option to perform aggregation or thinning of observations. This is especially useful for noise reduction or avoiding over-sampling when satellite pixels have a much smaller footprint than the model grid cell (which is not the case for S4/S5p products and current operational regional CAMS resolution).

The second step in this task is the development of a generic observation operator. It will consist of a source code that is able to read the satellite data stored in the common format produced by the pre-processor, and perform the following actions with this data:

- Simulation of retrievals from a model state. This includes remapping to pixel footprints and levels, and convolution with the averaging kernel to account for the sensitivity of the instrument to different altitudes.
- Actions needed for the assimilation systems present in the CAMS-RAQ ensemble, including computation of innovations (difference between simulations and observations), additions of forcing terms to a state (which is needed for variational methods) and support of observation screening (flagging of observations to leave them out of the analysis).

The implementation will support memory distributed states as used by regional air quality models that employ parallel computing based on domain decomposition. This includes memory distribution of the observations and an option to swap between different memory distribution.

The observation operator will be first defined generically in terms of classes and methods. This will follow the abstract building blocks used in OOPS (Object Oriented Prediction System), the framework used by ECMWF for data assimilation applications (*22*). OOPS defines for example abstract classes for states, observations, covariance matrices, and observation operators. This will allow implementation in different programming languages when needed. In the second step an actual implementation will be made in a modern version of Fortran (the native language of the models in the CAMS Regional Ensemble) which supports object-oriented features such as classes and class-bounded procedures.

The assimilation systems available within the project team include both variational methods (3D and 4D) as well as sequential methods. This ensures that the specification of the operator will support both types of assimilations. The requirements for the observation operator will be further detailed based on feedback from CAMS regional service providers that are not included in the project team. During the implementation phase, each of the five CAMS_61 teams will implement the generic observation operator in their assimilation system, to test the functionally and to ensure the applicability of the developed code to the majority of CAMS regional systems. Run times spent on using the operator with a set of sample data will be collected and where necessary used to guide improvements that reduce the computational costs. Evaluation whether the assimilation is actually beneficial for the analysis or the forecast is not part of this test but will be done for a selection of the models in Task 6132.

The source code of the observation operator will be made available together with the pre-processor and a set of sample data.

4.5.2 Task 6132: Assimilation of Sentinel-5p observations

To further prepare for the assimilation of future Sentinel-4 observations, three of the CAMS_61 systems will use the developed observation operator to assimilate data from the Sentinel-5p mission (TROPOMI instrument) which is expected to be representative for the expected accuracy and spatial resolution (and therefore data density) of Sentinel-4. The assimilation experiments will be configured similar to the experiments described in WP6122, but for a short time range only (1-3 months) and with the addition of Sentinel-5p products.

The currently available Sentinel-5p products[1] largely overlap with the planned Sentinel-4 products[2]. The products most relevant for air quality are NO_2 tropospheric columns and SO_2 and HCHO total columns. Information on tropospheric ozone is available in the form of O_3 profiles, but these have a limited amount of information on the troposphere, and a dedicated tropospheric O_3 column product is only available for tropical latitudes, which is outside the target domain. Aerosol Optical Depth (AOD), which will be provided by Sentinel-4, is currently not available (and not planned) as product of Sentinel-5p. The Sentinel-5p products include UV Aerosol Index (UVAI), but this product has not been assimilated yet by any of the air quality models in the ensemble. Based on these considerations, the species that will be assimilated are NO_2, SO_2, and HCHO.

The main challenge of this task is to optimize the configuration of the assimilation systems for the combination of satellite and surface data. This includes:

- Establishing vertical correlations that allow changes in the concentration profile that agree with both surface and satellite observations.
- Parameterization of observation representation errors, which are a combination of instrumental and retrieval errors supplied with the satellite product and the difference between the model grid and the satellite footprints.
- Observation screening criteria to exclude satellite observations that are too far beyond the uncertainty range allowed by the covariances.
- Assessment of the impact of the assimilation in analysis and forecast series.

The assimilation of S5p data will be performed by the following 3 systems, each using a different assimilation technique:

- LOTOS-EUROS, using EnKF assimilation; initial focus will be on SO_2 observations;
- EMEP using 3D-var, with initial focus on NO_2;
- SILAM using 4D-var, with initial focus on HCHO.

After exchange of experiences between the assimilation teams, multi-species (NO_2, SO_2, and HCHO) assimilation will be tested to evaluate the ability and impact of simultaneous assimilation of these satellite products.

4.5.3 Task 6133: Experiment with synthetic Sentinel-4 data

The developments and experiments from Tasks 6131 and 6132 will be extended to the use of Sentinel-4 data. Since Sentinel-4 will only be launched in 2023, a synthetic data set will be produced that represents the data density, spatial and temporal resolution, and the expected accuracy of the instrument. This set will be assimilated by the three assimilation systems employed in Task 6132 to evaluate the computational costs of analysing such a large data set. It is therefore necessary that the synthetic set has a realistic number of valid observations, but it is not necessary that it is fully representative for the future retrievals.

The synthetic data set will be created using simulations by the MONARCH model and Sentinel-5p observations. The (online) model will simulate concentration profiles for NO_2, SO_2, and HCHO on Sentinel-4 and Sentinel-5 footprints, as well as cloud properties and other relevant variables. Comparison of the simulations with actual Sentinel-5 data should show whether the simulations have the same characteristics as the observed values; eventually SEVIRI cloud fields will be used instead in case cloud properties are too different. Lookup tables will be generated to be able to assign retrieval

[1] http://www.tropomi.eu/data-products/level-2-products
[2] https://sentinel.esa.int/web/sentinel/missions/sentinel-4/data-products

variables such as averaging kernels and observation errors to the Sentinel-4 footprints, and convolution with the simulated tracer profiles this will then provide the synthetic data set.

The LOTOS-EUROS, EMEP, and SILAM systems that assimilated Sentinel-5p data in Task 6132 will assimilate the synthetic Sentinel-4 data set. The assimilations should lead to adjusted concentration fields that are closer to the MONARCH simulations. Evaluation of the computing times will provide insight in the ability of the assimilation systems to handle the large data volumes related to S4.

4.5.4 Task 6134: Research and Development plan for emission estimates

As will be studied in Task 6123, the forecast skill of a regional air quality model may be improved if it incorporates parameter estimates from a preceding assimilation. The assimilation reduces the uncertainty in these parameters and provides a most likely value that brings the simulations in best agreement with the observations. An important class of uncertain model parameters is formed by emissions, and estimating these through data assimilation is an important process that can improve air quality analyses.

To which extend emission parameters can be estimated is strongly determined by the type and the spatial and temporal availability of the observations. Polar orbiting satellites provide daily observations at overpass time only, which limits the potential to estimate hourly varying emissions. The future Sentinel-4 mission will however deliver hourly observations between sunrise and sunset, and can potentially provide information on emission parameters during a large part of the day.

The goal of this task is to provide a Research & Development plan on how to estimate emission parameters in the CAMS regional system as part of the assimilation process, taking into account the results of the work performed under Task 6123. The R&D plan will include a literature review and will be done in collaboration with the global service provider, who is investigating emission estimation for the CAMS global forecast system. In particular it will focus on the spatial and temporal variation of the emission parameter that can be estimated, and how this is related to the available observations. Also, the computational costs of the assimilation methods that are used will be considered. This will provide an assessment of the potential of emission parameter estimations, in particular using the future S4 observations. The CAMS Global and Regional Emission Providers (CAMS_81) will be consulted to assess what kind of emission parameters are considered as most uncertain. Based on this priority can be assigned to the viable emission parameters. The Global and Regional Service Providers of CAMS will then be consulted to investigate which of these applications are in reach of the current assimilation systems, or within a few years, and what kind of long-term developments are needed.

4.6 Summary of equipment and models

Table 4-1: Equipment (including hardware and software) to be used for provision of the Service

Equipment	Describe Relevant Function	List each work package for which equipment will be used	Owned / To be Purchased / To be Leased
TNO			
High Performance Computer servers and storage facilities	Downloading and storage of model (input) data and (satellite) observations. Computing resource for running LOTOS-EUROS model and its assimilation system	6110, 6120, 6130	Owned
LOTOS-EUROS	- Dutch national air quality model developed by TNO (in collaboration with the RIVM and KNMI) and run by the Institute to support the Dutch Ministry of Environment as well as foreign governments in supervising air quality management strategies. - Basis of the SMOGPROG system (Dutch air quality forecasting platform) operationally implemented since 2009. - Evaluation of the impact of various future emission reduction scenarios developed in WP 7120 and 7150 - In CAMS_61 provision of model runs including analyses including Sentinel data	6110, 6120, 6130	Owned
MET Norway			
MET Norway's High-Performance Computer and storage facilities.	Provision of EMEP MSC-W model runs Storage of model data to be analyzed and observational data for evaluation and assimilation.	6110,6120,6130	Owned
EMEP chemical transport model	-Reference European model initially developed to support the work of the Convention on Long range Transboundary Air pollution (CLRTAP) of the UNECE. - European air quality assessments and trends are elaborated annually by MET Norway (which is one of the technical and scientific centers working for the Convention) with the EMEP model to support the implementation of UNECE emission reduction strategies. - In CAMS61, the EMEP model will be used to for dedicated model runs, to study best practices for initializing forecasts from the analyses and for testing assimilation of satellite data.	6110,6120,6130	Owned

IT cluster and Software environment for running MET Norway's data processing jobs	Operational steering of model and post processing software for data handling and visualisation	6110,6120,6130	Owned
FMI			
High-performance computing based on FMI own supercomputer	Performing the SILAM computations, data assimilation and analysis of the results	6110, 6120, 6130	Owned
SILAM	SILAM is the Finnish national air quality and emergency preparedness model developed in FMI and used for decision support in cases of emergency, as well as for daily environmental policy procedure: public information, episode analysis and forecast, emission abatement, government consultancy services, atmospheric composition research, etc. For CAMS-61, it will perform hindcast and reanalysis of 2018 and selected episodes	6110, 6120, 6130	Owned
INERIS			
Research High Performance Computer servers and storage facilities (CCRT/TGCC, France)	Storage and computing resources for the development of diagnostic and assimilation systems	6110, 6120, 6130	Leased
Pre-Operational High-Performance Computer servers and storage facilities (INERIS partition hosted by Météo-France)	Testing assimilation systems	6110, 6120, 6130	Leased
Computers with office suite	Reporting activities & presentations	6110, 6120, 6130	owned
CHIMERE	- French national air quality model developed by INERIS (in collaboration with the National Research center) and run by the Institute to support the French Ministry of Ecology in supervising air quality management strategies. - Operational capacities for providing daily air quality forecasts and analyses and yearly re-analyses - Used to evaluate the impact of various future emission reduction scenarios and to elaborate national air quality plans, as well as in the negotiations, which set at the European level national emission ceilings (Directive NEC and Gothenburg Protocol of the CLRTAP). - In CAMS_61 used for model runs including analyses	6110,6120,6130	

BSC			
20 nodes (960 cores) in Marenostrum4 supercomputer, Intel Xeon Platinum 8160, 20TB storage in GPFS HPC disk	HPC resources to run the MONARCH model and test the Sentinel observational operator	6110, 6130	Owned
100 Tb Storage in GPFS archive	Data archiving	6110, 6130	Owned
CTE-POWER, cluster based on IBM Power9 processors using GPFS HPC storage	In-situ observations processing, and model output postprocessing	6110, 6130	Owned
MONARCH model	Software code for air quality modelling/forecast	6110, 6130	Owned
GHOST database	Observational database management system for air quality observations with standardised quality assurance/quality control methods	6110	Owned

The regional model systems and data assimilation techniques that are used in the proposed project are listed in Table 4-2. They all fulfil the general requirements set in the ITT.

Table 4-2 Overview of regional modelling systems and assimilation techniques used within this ITT

MODEL	INSTITUTE	ASSIMILATION TECHNIQUE(S)
LOTOS-EUROS	TNO	EnKF
EMEP	MET Norway	3D-VAR
SILAM	FMI	3D-var, 4D-VAR, EnKF
MONARCH	BSC	EAKF
CHIMERE	INERIS	OI, ETKF

4.7 References

Below we have listed all references cited in Sections 1 to 4.

1. Manders, A. M. M. *et al.* Curriculum vitae of the LOTOS-EUROS (v2.0) chemistry transport model. *Geosci. Model Dev.* **10,** (2017).

2. Timmermans, R. *et al.* Impact of synthetic spaceborne NO2 observations from the Sentinel-4 and Sentinel-5p platforms on tropospheric NO2 analyses. *Atmos. Chem. Phys. Discuss.* 1–40 (2019). doi:10.5194/acp-2018-1360

3. Simpson, D. *et al.* The EMEP MSC-W chemical transport model; technical description. *Atmos. Chem. Phys.* **12,** 7825–7865 (2012).

4. Schaap, M. *et al.* Performance of European chemistry transport models as function of horizontal resolution. *Atmos. Environ.* **112,** 90–105 (2015).

5. Bessagnet, B. *et al.* Presentation of the EURODELTA III intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology. *Atmos. Chem. Phys.* **16,** 12667–12701 (2016).

6. Colette, A. *et al.* EURODELTA-Trends, a multi-model experiment of air quality hindcast in Europe over 1990-2010. *Geosci. Model Dev.* **10,** (2017).

7. Curier, R. L. *et al.* Synergistic use of lotos-euros model and OMI NO2 tropospheric column to evaluate the nox emission trends over Europe. in *34th International Symposium on Remote Sensing of Environment - The GEOSS Era: Towards Operational Environmental Monitoring* (2011).

8. Schaap, M. & Dammers, E. Assessing the ammonia emission distribution across Germany using remote sensing observations. *Atmsopheric Environ.* **submitted,** (2018).

9. Bocquet, M. *et al.* Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models. *Atmos. Chem. Phys.* **15,** 5325–5358 (2015).

10. Elbern, H., Strunk, A., Schmidt, H. & Talagrand, O. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos. Chem. Phys.* **7,** 3749–3769 (2007).

11. Timmermans, R. M. A. *et al.* The Added Value of a Proposed Satellite Imager for Ground Level Particulate Matter Analyses and Forecasts. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2,** 271–283 (2009).

12. Vira, J. & Sofiev, M. On variational data assimilation for estimating the model initial conditions and emission fluxes for short-term forecasting of SOx concentrations. *Atmos. Environ.* **46,** 318–328 (2012).

13. Vira, J. & Sofiev, M. Assimilation of surface NO2 and O3 observations into the SILAM chemistry transport model. *Geosci. Model Dev.* **8,** (2015).

14. Sofiev, M. On possibilities of assimilation of near-real-time pollen data by atmospheric composition models. *Aerobiologia (Bologna).* **1,** (2019).

15. Menut, L., Bessagnet, B., Menut, L. & Bessagnet, B. What Can We Expect from Data Assimilation for Air Quality Forecast? Part I: Quantification with Academic Test Cases. *J. Atmos. Ocean. Technol.* **36,** 269–279 (2019).

16. Pannekoucke, O., Berre, L. & Desroziers, G. Background-error correlation length-scale estimates and their sampling statistics. *Q. J. R. Meteorol. Soc.* **134,** 497–508 (2008).

17. Descombes, G., Auligné, T., Vandenberghe, F., Barker, D. M. & Barré, J. Generalized background error covariance matrix model (GEN_BE v2.0). *Geosci. Model Dev* **8,** 669–696 (2015).

18. Solazzo, E., Hogrefe, C., Colette, A., Garcia-Vivanco, M. & Galmarini, S. Advanced error diagnostics of the CMAQ and Chimere modelling systems within the AQMEII3 model evaluation framework. *Atmos. Chem. Phys.* **17,** 10435–10465 (2017).

19. Solazzo, E. *et al.* Evaluation and error apportionment of an ensemble of atmospheric chemistry transport modeling systems: multivariable temporal and spatial breakdown. *Atmos. Chem. Phys.* **17,** 3001–3054 (2017).

20. Otero, N. *et al.* A multi-model comparison of meteorological drivers of surface ozone over Europe. *Atmos. Chem. Phys.* **18,** 12269–12288 (2018).

21. Benedetti, A. *et al.* Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation. *J. Geophys. Res. Atmos.* **114,** (2009).

22. Bonavita, M. *et al. A Strategy for Data Assimilation, ECMWF Technical Memoranda 800*. (2017). doi:10.21957/tx1epjd2p

5 Management and implementation plan

5.1 Introduction

TNO, as ECMWF's contractor, will be responsible for all management aspects of the CAMS_61 project (see WP6110 description). This includes the establishment of the physical and organisational structure to carry out the activities requested in this ITT.

Applicable documents will be:

- The CAMS_61 tender and then CAMS_61 contract should this tender be selected.
- The quality standards for research and service of TNO will be applied.

The main management structure is described in the following sections and is defined as Work package 6100 Management and coordination.

5.2 Organigram

The proposed management structure for CAMS_61 is outlined in Figure 8.

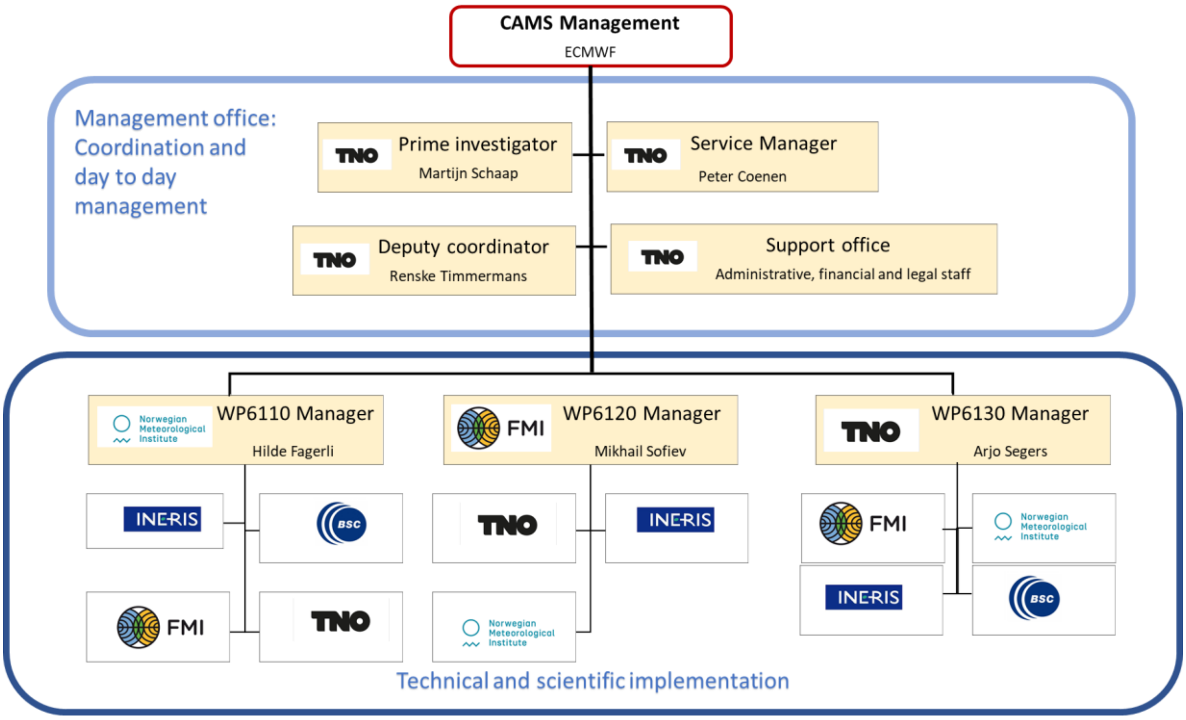


Figure 8. The CAMS_61 Management structure

TNO agrees to use ECMWF templates for deliverables and reports. To ensure efficient reporting on progress (content wise and financial) TNO will demand the work package leaders to report to TNO in ECMWF templates. It is expected that the frequency of changes to templates will be low (not more than once a year).

5.3 Management team

The overall coordination and project management of CAMS_61 is the responsibility of TNO.

The CAMS_61 management team will be composed of 7 members, reflecting the different work packages (leads), the different institutes (subcontractors) and from TNO the prime investigator and service manager (chair).

The management team is responsible for maintaining an effective and efficient working relationship and coordination between subcontractors within WPs 6100 to 6130 during the project.

Peter Coenen will be the service manager (and lead of WP 6110) and professor Martijn Schaap is the envisaged prime investigator. They will be assisted by Renske Timmermans, senior scientist and deputy coordinator and Arjo Segers, senior scientist. Renske will together with the service manager and prime investigator work on the different management tasks, especially related to scientific aspects (scientific management of subcontractors and quality assurance of deliverables). Arjo Segers will be work package leader for WP6130 on the preparation for assimilation of geostationary satellite observations.

The CAMS_61 management team will be complemented with the work package leaders Dr. Hilde Fagerli (MET Norway, WP6110 on the in-depth evaluation of CAMS regional systems) and Prof. Dr. M Sofiev (FMI, WP6120 on the coupling of analyses and forecasts) and with Dr. Maria Teresa Pay (BSC) and Dr. Augustin Colette who will represent the subcontractors who do not lead a work package in the management team.

The CAMS_61 Management team will maintain strong interactions with other relevant CAMS projects, such as CAMS_50 (regional services), CAMS_81 (emission services), CAMS_84 (validation services), CAMS_42 (developments for the global reactive gases) and CAMS_43 (developments for the global aerosols).

5.3.1 The Service manager

The service manager will be the primary contact for contractual delivery and performance aspects.
His responsibilities include the following tasks:

- organise coordination between the subcontractors and the WPs (incl. internal meetings);
- establish together with the prime investigator efficient management and reporting systems;
- assist WP leaders in coordinating their WPs (in particular for WP 6130);
- organise the CAMS_61 milestone meetings;
- follow and maintain the CAMS_61 Annual implementation plans (with prime investigator);
- coordinate and control the overall Service activities (incl. reporting activities across the WPs);
- coordinate and control the drafting of commitments and contracts with partner organisations and their execution;
- report to CAMS Management (Service-Level Boards);
- resource planning and trackingr
- manage risks issues;
- manage WP 6100 (Management and coordination) and associated deliverables.

Although the number of partners and work packages is limited, we know from experience that the managerial and administrative burden for such a project is considerable. The project manager will therefore be assisted by Renske Timmermans and experienced TNO administrative, financial and legal staff. This staff worked on earlier CAMS projects and is familiar with the CAMS requirements. This support staff will assist the service manager in :

- internal and external planning;
- handling administrative and budgetary issues;
- handling contract and subcontract monitoring and the financial reporting to ECMWF;

- personal data management (contact details for the responsible person can be supplied if necessary).

For scientific issues the service manager will liaise with the prime investigator.

5.3.2 The Prime investigator

The prime investigator will be coordinating and leading the scientific activities in CAMS_61. This includes the following tasks:

- represent the consortium during monthly teleconference meetings and 6-monthly review meetings with ECMWF to discuss the service provision and other topics;
- validate the scientific choices of the Service, in line with recommendations from ECMWF, from user needs and from the other CAMS services related to model development and validation;
- coordinate and monitor progress of the development activities among the partners;
- follow and maintain the CAMS_61 Annual implementation plans (with service manager);
- will establish together with the service manager efficient management and reporting systems;
- report to CAMS Management (Service-Level Boards);
- monitor the KPIs regarding the quality of the deliverables and propose actions to improve results, should this be necessary, or new KPIs;
- quality assurance and control (QA/QC) of the final (non-TNO) reports and deliverables; for the TNO reports the final QA/QC will be delegated to one of the other WP managers.

Besides these activities directly related to the services the prime investigator, with his extensive network in and outside the CAMS and air quality community, will also contribute and be responsible for external relations of the project. He will support ECMWF with their communication activities related to the CAMS_61 activities (e.g. website news items).

The prime investigator will liaise closely with the service manager and the administrative staff who will assist him in his day to day work.

To safeguard the landing of the outcomes of the project in the broader scientific community, the appointment of one or two external advisors is envisaged who will be engaged in the review of the main deliverables of the project and participate in the annual meetings. In addition, an external advisor will be appointed to ensure the inclusion of correct specifications for the future Sentinel-4 and use of current Sentinel-5p observations. Financial provisions for the advisors are included in the budget tables accompanying this proposal.

5.3.3 The work package managers

The work packages will be led by individual work package managers. Each work package will have its own specific day to-day management and will report on a monthly basis via web conferencing or email, focusing on important issues and changes. The work package manager will have the following tasks:

- responsible for the coordination of their own WP and for the associated deliverables (including QA/QC);
- participate in the management team;
- responsible for the planning and (financial) reporting to the CAMS_61 management team;
- validate the technical choices of the project, in line with recommendations and constraints from ECMWF and the envisaged advisors on these matters;
- propose and follow the KPIs related to their WP.

Within each work package the various tasks are allocated to task leaders. Distributing task management will allow to share management activities and to motivate all the contributors as well, giving them responsibilities in the achievement of the project goals. Each task leader will be responsible for the products and deliverables associated with their task, and for the milestone completion.

5.4 Management procedures

5.4.1 Project management tools

The anticipated tools to facilitate information sharing are as follow:

- a secured SharePoint site supplied by TNO and to be used by all partners; it will be the reference for all meetings reports, implementation plans and other documentation and will allow interactive work on documents;
- monthly CAMS_61 meetings, mainly through web conferencing (Skype for business of Blue Jeans sessions) allowing presentations to be shown and discussed;
- easy to use project management tooling to be used by TNO (and its partners) to keep track of all CAMS_61 activities and prepare reports and other related matters for communication towards ECMWF.

5.4.2 Reporting and meetings

Formal quarterly and annual reports will be prepared and submitted by TNO to ECMWF in conformity with Clause 2.3 of the Framework Agreement. They will amongst others provide information on the performed activities during the previous periods, review progress in deliverables, compliance with the milestones and solutions to fix potential deviations from the implementation plan.

Progress and results will also be formally reported to ECMWF through the following meetings organized by ECMWF:

- monthly teleconference (or videoconference) meetings hosted by ECMWF to discuss CAMS service provision, service evolution and other topics;
- six-monthly project review meetings organized by ECMWF (linked to Payment milestones);
- annual CAMS General Assemblies within EU member states; these will be attended by team members representing the different activities in the project.

In addition to the meetings organised by ECMWF, the following meetings will be organised by the consortium:

- kick off meeting;
- annual internal face to face meetings, preferably linked to the annual CAMS general assemblies;
- monthly internal videoconferences to discuss and follow progress of the project;
- three meetings with ECMWF and the CAMS_50 development teams, where the model evaluation and the suggested model developments from WP 6110 will be discussed in detail;
- one meeting with ECMWF to agree on the testing protocols in WP6120 and WP6130 (combination of milestones M2.1.1 and M3.0.3).

5.4.3 Subcontractors Management

TNO will apply a carefully-devised subcontractor management strategy, which already was implemented in the bid phase. The four subcontractors have signed pre-bid agreements with TNO outlining their responsibilities in the project (as stated in this proposal) and save guarding capacity to perform the CAMS_61 project. After approval of the tender the subcontractors will sign subcontracts with TNO in which ECMWF terms and conditions also will be included. This safeguards that the different organisations are bound to the required ECMWF service level.

The key features of the subcontract management are specified below:

- Legally-binding subcontracts with clear provisions concerning deliverables, milestones, performance obligations and budgetary constraints.
- A cohesive work environment conducive to open discussion between TNO and the subcontractors.
- Easy communication channels.
- Appropriate information allowing subcontractors to have an overview of the overall CAMS_61 activities and interactions with other CAMS services, and where their contributions fit.
- Smooth payment flows and assistance to subcontractors concerning invoicing and payment matters. This matter should not be overlooked since finances are one of the key considerations for the subcontractor.
- Subcontractor performance regularly monitored through regular meetings and reports (as much as possible tuned to the ECMWF templates) to ensure activities are being completed according to plan.
- Back-up contact person assigned by each subcontractor for all relevant domains.

This will ensure that TNO (and the management team) stays up-to-date on progress made and any potential roadblock.

5.4.4 Conflict resolution

The management team will facilitate a cooperative working environment where the specific teams will be able to discuss healthy conflicts in depth and express different points of view. Additionally, most of the teams have long worked together through a.o. CAMS_50 and the precursor projects, which should mitigate the risk of a critical conflict situation occurring and adversely affecting the success of CAMS_61.

Should such situation arise, it should be stressed that dispute provisions will be included in all subcontracts, both in terms of formal dispute settlement and in terms of performance continuity during pendency of dispute. However, TNO stands firm on the need to firstly explore all possible avenues and help the teams concerned find win-win solutions before escalading a conflict this way.

Concerning potential conflicts with other CAMS services connected to CAMS_61, we will seek to work together on good terms provided the activities fall within the scope of the ITT and within budgetary constraints.

5.5 Other aspects

5.5.1 Geographical and gender balance

The consortium includes two Nordic teams, one Dutch, one French and one Spanish team covering the different regions from North to South across the European Union.

The consortium exhibits a good female to-male ratio at the various work levels. The coordination of activities at two out of 5 consortium partners and the deputy coordination of the full project will be fulfilled by female personnel. Gender-biased practices or language will not be permitted in the Service day-to-day operations. Lastly, all partner institutes strive to maintain a supportive environment for a gender-balanced workforce.

5.5.2 Outreach towards users

Although outreach towards users does not lie at the heart of the ITT, limited communication actions will be performed by CAMS_61 Management team, by means of presentations during the CAMS General assemblies, press articles and news releases through the TNO websites for instance. Proper acknowledgement of EU Copernicus funding will be displayed on these occasions and in scientific publications resulting from CAMS_61.

5.5.3 Pre-existing technologies

The services and products developed by TNO and its sub-contractors are based on pre-existing tools they develop for a long time, generally for other purposes, and that are adapted or completed by new functionalities to achieve CAMS_61 goals. Pre-existing technologies and assets generally refer, in our cases to the chemistry transport models we develop and use in this framework.

Pre-existing technologies			
Title	**Type**	**Description**	**Owned by**
LOTOS-EUROS	Numerical model	Chemistry transport model	TNO
EMEP	Numerical model	Chemistry transport model	MET Norway
CHIMERE	Numerical model	Chemistry transport model	INERIS
SILAM	Numerical model	Chemistry transport model	FMI
MONARCH	Numerical model	Chemistry transport model	BSC

5.5.4 Custody of the deliverables

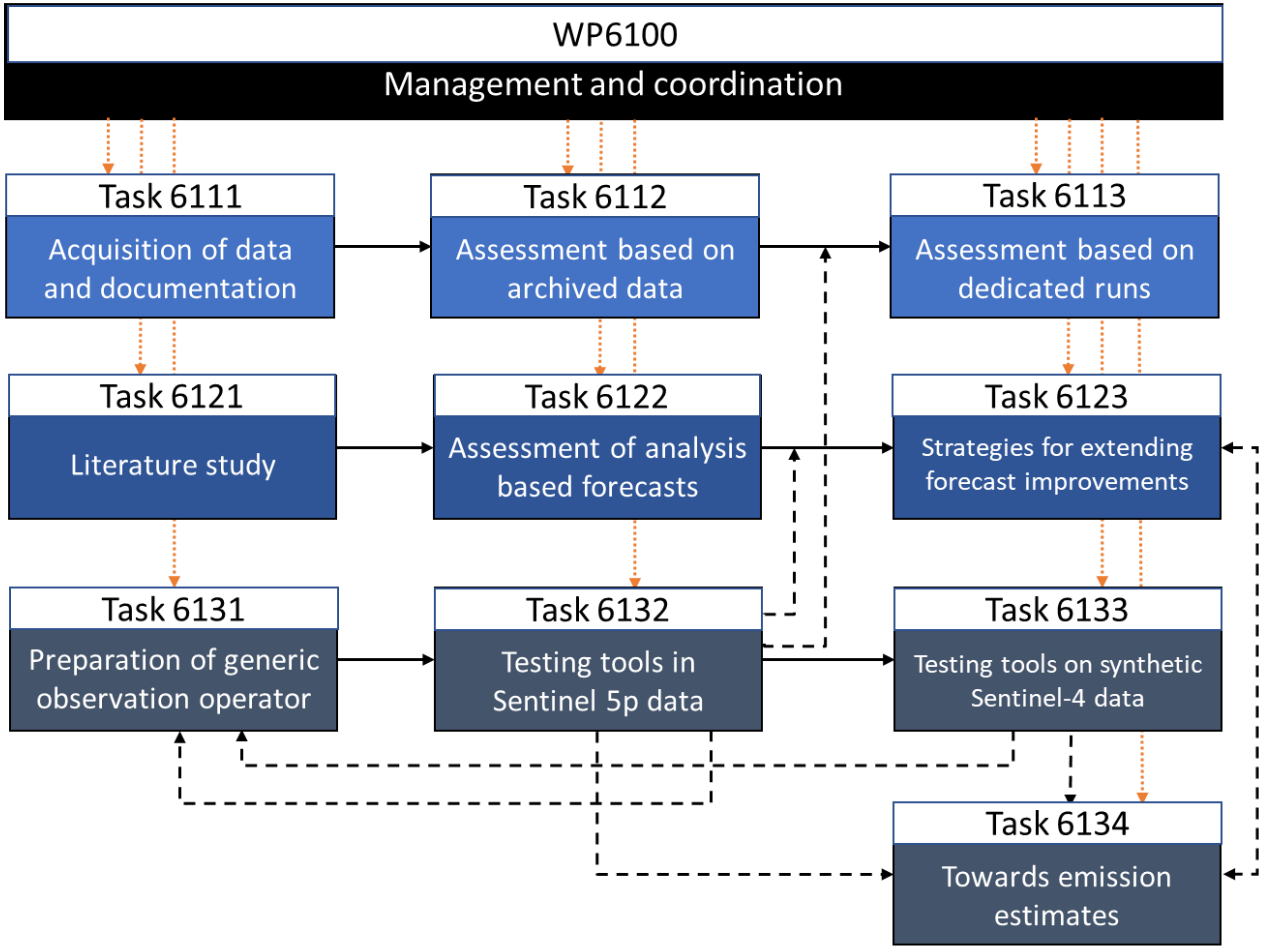
TNO will take care of the custody of the deliverables in the form of a secured electronic archive during the project. This archive can be made accessible to ECMWF. After closure of the project the content of this archive will be transferred to ECMWF.

5.6 Gantt chart and PERT chart

The Gantt chart below illustrates the CAMS_61 time line of each work package and tasks and the milestone and deliverable dates, D= Deliverable, M=Milestone, ♦= milestone meeting:

	Year	2019		2020												2021											
	Calender month	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
	Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
WP/Task	**Title**																										
WP 6100	**Management**																										
	Overall coordination																										
	Reporting																										
	Kick-off Meeting	♦																									
	CAMS general assembly + Internal face to face meeting											♦												♦			
	progress review meetings with ECMWF						♦						♦						♦						♦		
	teleconference meetings with ECMWF	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦
	Internal project teleconference meetings	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦	♦
WP6110	**In-Depth assessment regional forecasts**																										
Task 1	Acquisition of data and documentation					M															D						
Task 2	Assessment based on archived forecasts						D																				
Task 3	Assessment based on dedicated model runs								M				D			M			D								
	Meetings with ECMWF and development teams						♦						♦						♦								
WP6120	**Coupling of regional forecasts and analyses**																										
Task 1	Literature study			D																							
Task 2	Assessment of efficiency of analysis-based initialisation												D														
Task 3	Strategies for extent of forecast improvement																				D						
	Meeting with ECMWF to agree on testing protocol			♦																							
WP6130																											
Task 1	Generic observation operator for satellite data				D		M												D								
Task 2	Assimilation of Sentinel 5p observations												D/M														
Task 3	Tests with Synthetic Sentinel 4 observations												M						D								
Task 4	Towards emission estimates												M						D								
	Meeting with ECMWF to agree on testing protocol			♦																							

The Pert Chart below illustrates the interdependencies between the coordination work package steering all activities and the tasks in the three scientific work packages. The tasks will be mainly performed consecutively within the WPs and thus are dependent on the previous task. Strong relationships exist between all the work packages, as these have joined objectives regarding improvement of model and data assimilation performances and consider the same modelling and data assimilation systems.



5.7 Work package description

Table 5-1. WP6100

Work package #	WP6100	**Start/End date**	**M1/M26**
Work package title	**Management and coordination**		
Participants (person months)	TNO (5.68)		
Other main direct cost elements	Travel (4k), Reservation for fee imbursement of scientific advisory (10k)		

Main objectives
This work package is related to the overall management and coordination of all the activities in the project. The main objective is to implement and maintain all the physical and organisational structure required to reach the defined deliverables and milestones within budget and time. This includes the reporting both internally as well as towards ECMWF.

Description of activities

Task 6101 Overall coordination and links with other relevant CAMS projects (Lead: TNO)

- Follow the service contract implementation and performance monitoring (including monthly internal web conferences)
- Drafting annual implementation plans in collaborations with all project partners
- Maintaining necessary links with other relevant CAMS projects
- Resource planning and tracking
- Risk management
- KPI management
- Handling of contractual matters

Task 6102 Reporting (Lead: TNO)

- Reporting to ECMWF on progress, any major existing or anticipated issues
- Facilitating and reviewing the delivery of reports
- Financial reporting

Deliverables and Milestones
The deliverables and milestones are provided in Annex 2. List of deliverables and milestones

Table 5-2: WP6110

Work package #	WP6110	**Start/End date**	**M1/M20**
Work package title	**In-depth assessment of the CAMS Regional Systems to identify future development needs**		
Participants (person months)	MET Norway (15), TNO (4.28), FMI (2), BSC (12), INERIS (4.25)		
Other main direct cost elements	Travel (23k) and computing costs (10.4k), other (1.1k)		

Main objectives

The main objective of this work package is to identify model developments that are needed to improve the quality of the CAMS Regional Ensemble and the individual Regional Systems.

Based on an in-depth evaluation of the CAMS Regional Systems forecasts, an overview of situations where individual models or the Ensemble deviate significantly from observations will be provided, including possible reasons for the discrepancies.

Dedicated model runs and sensitivity studies will be performed to allow for a more robust identification of the reasons behind model discrepancies and possible ways of resolving them.

Based on our understanding as to which processes that are not well modelled by one or all of the Regional Systems, we will provide a prioritized list of individual model developments as well as developments that are common to all models.

Description of activities

Task 6111: Acquisition of model data, observational data, and model documentation (Lead: Metno with contributions from BSC)

In this task modelling data from the CAMS_50 model database hosted at Meteo France will be acquired along with available measurement data for the evaluation period specified in the call (09/2018 to 09/2019). Quality-controlled EEA AQ e-Reporting/EBAS/WMO-GAW observational data will be screened using the GHOST System to assure a set of high-quality measurement data, harmonised across species and networks in a consistent way. Measurement data will be complemented by chemically speciated PM measurements from the EMEP network and data from the EMEP/ACTRIS/COLOSSAL intensive measurement period 2017/2018.

Detailed descriptions of the individual models will be requested from CAMS_50 and studied in order to be able to better understand the behaviour by the different models. The CAMS_50 Regional Systems were upgraded in June 2019, i.e. within the evaluation period specified in the call. Therefore the 'Request for Change' documents provided by the CAMS_50 model teams will be retrieved and scrutinized as well.

Task 6112: Evaluation of the operational CAMS regional forecast data with extended statistical analyses (Lead: Metno with contributions from TNO, FMI, BSC and INERIS)

The analysis of the difference between the CAMS Regional Systems and air quality observations will be performed for all the pollutants available from the archived forecast data for the period 09/2018 to 09/2019 and will consider several aspects such as temporal profiles (diurnal, day-of-week), regional and seasonal variations, different meteorological conditions and different lead times of the forecast. The analysis will focus on main regulatory pollutants (ozone, NO_2 , PM_{10} and $PM_{2.5}$), but all other pollutants provided by the

Regional Systems from CAMS_50 will be covered as well. Episodes of high air pollution (PM, ozone and NO_2) concentrations in different regions of Europe will be given special attention. Episodes magnify and emphasize important physical and chemical processes and therefore better reveal underlying problems in the model systems or input data. The evaluation will be complemented by a correlation analysis, also extending to the vertical, to find specific sensitivities of a model to different physical and chemical processes in order to give hints on which processes should receive special attention.

Task 6112 will result in a list of main disagreements between models and observations and between individual models and the ensemble, as well as a report on possible reasons behind these differences, and suggestions on which developments should be made in models. The evaluation in Task 6112 will also serve as a basis for the design of dedicated model experiments in Task 6113.

Task 6113: Diagnostic evaluation based on dedicated model runs (Lead: Metno with contributions from TNO, FMI, BSC and INERIS)

In order to assess the reasons behind the systematic differences documented in Task 6112 in more detail and to be more specific on the model developments needed, we will perform a set of model reruns with considerably enhanced output (e.g. full chemical speciation of PM, AOD, dry and wet deposition of chemical components, etc.). The specifications of these runs will depend on the outcome of Task 6112, but will target different aspects, such as the impact of the changes in anthropogenic emissions (as specified in the call), usage of boundary conditions and other input data, as well as various parameterizations in models, which have been identified as particularly uncertain and important in Task 6112.

Several models of CAMS_61 will participate in these runs, but the exercise will also be open for all other CAMS_50 models and will be announced at appropriate CAMS meetings. As a by-product of our detailed model assessment, we will propose a one-year benchmark set of observations that can be used in future evaluations of model developments in CAMS.

In addition, the assessment of the Regional Systems will take into account the lessons learned from WP6130, where data assimilation of Sentinel-5p and parameter estimation may provide further understanding of the deviations between Regional Systems and observations.

The results from Task 6112 and Task 6113 will be presented at three meetings with ECMWF and the teams of the CAMS Regional Systems. Taking into account the feedback from each of the CAMS_50 Regional Systems development teams and relevant items from the CAMS User Requirement Data Base (URDB), a prioritized list of recommended model developments will be provided (with a maximum estimated duration of 6 to 12 months). These recommended activities will target individual model systems as well as cross-cutting activities applicable to all model systems.

Deliverables and Milestones
The deliverables and milestones are provided in Annex 2. List of deliverables and milestones

Table 5-3: WP6120

Work package #	WP6120	**Start/End date**	**M1/M26**
Work package title	**Coupling of regional forecasts and analyses**		
Participants (person months)	FMI (15.7), TNO (5.05), MET Norway (3), INERIS (2)		
Other main direct cost elements	Travel (9k) and computing costs (7k)		

Main objectives

The overall goal of this work package is to evaluate the benefit of basing regional forecasts on analyses of surface observations (responding to the Tender Specification item 3).

To achieve this goal, the following specific objectives have been formulated:

- to identify and analyse the existing approaches to the initial-value assimilation in the air quality problems: a literature review
- to assess the improvement in forecast skill against independent air quality observations, when an analysis is used as initial conditions for the forecast
- to evaluate the longevity of the forecast improvement when analysis is used as initial conditions and compare the effect with using novel approaches based on assimilation of other variables, such as emission or model parameters, and inheritance of updated parameters/emissions into the forecasts.

Description of activities

Task 6121. Analysis of existing material on analysis usage for air quality models initialization. (Lead: TNO, contributions from FMI, MET Norway, INERIS)

A literature review will be conducted in order to collect and analyse the existing practices in the field of coupling forecasts to analyses.

Task 6122. Multi-model assessment of efficiency of analysis-based initialization (Lead: FMI, contributions from TNO, MET Norway and INERIS)

Within this task, three CAMS-50 operational models (SILAM, EMEP, LOTOS-EUROS) will be run through the full year of 2018 several times using several setups.

- Original set-up for forecasts without coupling to analyses
- Analyses with data assimilation of ground-based observations of O_3, PM, CO, NO_2 and SO_2
- Forecasts initialised with the analyses at 00h, 06h, 12h, and 18h

The outcome of these simulations will be compared paying attention to spatial distribution of the effects in polluted / clean regions and cross-species impact via chemical links.

The episodes with particularly strong / weak effect of the analysis-based initialization will be reviewed with the CHIMERE EAKF data assimilation system aiming at the mechanisms behind and possibilities to improve the efficiency of the analysis usage

Task 6123. Analysis of episodic longevity and extent of forecast improvement for different assimilation strategies (Lead: FMI, contributions from TNO)

The episodes will particularly strong/weak effect of the analysis-based initialization will be recomputed with the SILAM and LOTOS-EUROS model using 4D-Var and/or EnKF data assimilation techniques directed to the emission fluxes and/or model parameters rather than the initial conditions. The forecasts will then be based on inherited parameter/emission estimates.

Deliverables and Milestones
The deliverables and milestones are provided in Annex 2. List of deliverables and milestones

Table 5-4: WP6130

Work package 3	WP6130	**Start/End date**	**M1/M20**
Work package title	**Towards assimilation of observations from geostationary satellite sensors (Sentinel-4) to constrain concentrations and emissions of main pollutants**		
Participants (person months)	TNO (12), FMI (5.3), MET Norway (8), BSC (5), INERIS (1.1)		
Other main direct cost elements	Travel (performed under other WPs) and computing costs (3k)		

Main objectives

The main objective is to develop, implement, and release a generic observation operator that will facilitate the assimilation of future Sentinel-4 observations with the air quality models of the CAMS ensemble. The source code of the operator will be released together with a pre-processor to convert satellite data into a common format.

The observation operator will be used to assimilate Sentinel-5p observations in three different assimilation systems, in addition to assimilation of surface observations. The assimilation configurations will be adapted to allow analysis of data from both sources, and impact on both analysis and forecasts will be examined.

A set of synthetic Senitel-4 data will be created to analyse the performance of the assimilation systems when assimilating data sets with the volume expected for this future mission.

A Research & Development plan on using satellite data for estimation of emission parameters within the CAMS services will be created. In particular this will focus on the potential use of Sentinel-4 data for estimating this type of parameters.

Description of activities

Task 6130 Work package management (Lead TNO, contributions from MET Norway, FMI, BSC, INERIS)

Management of reporting related to this work package, including production of development plan, progress report, and final report.

Task 6131 - Develop, implement, and test observation operator suitable for future S4 data sets (Lead TNO, contributions from MET Norway, FMI, BSC, INERIS)

A generic observation operator will be implemented to facilitate the assimilation of future Sentel-4 observations. The operator is implemented in a recent Fortran dialect (2003 or higher). A pre-processing tool is implemented in the Python language to convert, select, and eventually aggregate the satellite data into a common format. All teams will test the functionality of the operator in their assimilation system, to evaluate whether it contains all methods needed for the assimilation system, and whether it is efficient in handling large volumes of satellite data. The observation operator, the pre-processor, and a set of sample data will be released to the teams of the CAMS regional ensemble.

Task 6132 - Assimilation of S5p data in operational context (Lead TNO, contributions from MET Norway and FMI)

For a selected period of 1-3 months within 2018, observations from the Sentinel-5p mission will be assimilated in three operational analysis/forecasts systems. Each system first focusses on a single observed species: EMEP/MSC-W on NO_2, LOTOS-EUROS on SO_2, and SILAM on HCHO. Assimilation configurations will be optimized for the combined analysis of satellite and surface observations. After exchange of experiences, each of the systems will also assimilate species.

Task 6133 - Assimilation test with synthetic Sentinel-4 data (Lead BSC, contributions TNO, MET Norway, FMI)

The MONARCH model will be used to create a set of synthetic Sentinel-4 data, for the same period as used

for the Sentinel-5p assimilation. Auxiliary variables will be obtained from comparable Sentinel-5p data. The synthetic data set will be assimilated by the three assimilation system to evaluate performance. **Task 6134 Research and Development plan on emission parameter estimates from assimilation (Lead FMI, contributions from TNO, MET Norway, BSC, INERIS)** A Research and Development plan will be written on estimating emission parameters in the CAMS regional and global assimilation systems. A literature review will be done on current available methods, with special focus on the use of satellite data, in particular on satellites that share characteristics with future Sentinel-4. CAMS stakeholders on emissions (including the global and regional emission projects, and the global and regional modelling projects) will be consulted to include needs and limitations from the users in the plan.

Deliverables and Milestones
The deliverables and milestones are provided in Annex 2. List of deliverables and milestones

5.8 Key Performance Indicators

Table 5-5: Key Performance Indicators

KPI #	KPI Title	Performance Target and Unit of Measure	Frequency of Delivery	Explanations / Comments
KPI_61.0.1	% of deliverables delivered on time or with short delay (< 2 weeks)	90%	As indicated in Volume IIIC table of deliverables and milestones	
KPI 61.1.2	Number of developments that are proven to lead to improvements in the models participating in WP6110	3 development recommendations	Once per recommendation	Dedicated model runs will be performed to evaluate and test developments. At least three developments will be recommended that have proven to improve the models
KPI 61.1.3	Improvement of scores from the evaluation of the archived data in CAMS_50 (month 6) to the evaluation of the dedicated model runs in Task 6113 (month 12 to 18)	RMSE, bias and correlation should improve for more than half of the species	Month 18	The performance of the models participating in the CAMS_50 ensemble and in the dedicated model runs in WP 6110 should improve through suggested developments and the outcome of the other WPs.
KPI 63.1.1	Applicability of developed tools in task 6131 for CAMS regional systems in at least three different models	Three different models, analysis maps for three species based on Sentinel 5p assimilation and the developed tools	Once, Month 18	Five models will technically implement the developed tools. Three models will implement and test the developed tools on Sentinel 5p and synthetic Sentinel 4 data.

5.9 Risk management

Table 5-6: Risk Register for each Work package

Work package:					
Risk Name	**Description**	**Likelihood**	**Impact**	**Response Strategy**	**Period**
Additional data at the surface (e.g. from EMEP network)	Risk of difficulties in getting the ground based observational data.	3	3	Avoid and reduce: Management team shall interact directly with EMEP/CCC.	M1-M20
Information about the Regional Systems	Risk of not having enough detailed information about the different Regional Systems to be able to understand deviations	3	3	Avoid and reduce: The model development teams will be consulted directly with detailed questions	M1-M20
Unexpected unavailability of key personnel	One of the key personnel is not available to perform the proposed tasks	3	3	Avoid: For all key persons in Table 3-1, back-ups are available or will be set in place at each institute	M1-M20
Failure in the HPC infrastructures	The planned computational activities cannot be per	2	4	Reduce: Back-up systems or access to other HPC infrastructures are available	M1-M20
Failure of delivery	Risk of no delivery of model reruns for 2018 from one or several teams in time due to unexpected circumstances that may result in problems performing in depth assessment	2	3	Reduce and accept: The teams concerned will be requested to deliver the missing data later, but no later than M12	M1-M20
Failure from one of the partners to deliver	Risk of a partner not being able to perform agreed tasks within the project due to e.g. capacity issues, financial issues, computational issues, that may result partner withdrawing from project	1	3	Reduce and accept: The participating partners have overlapping expertise and capabilities, in case the risk takes place, another partner will be able to take over the responsibilities	M1-M20

Table 5-7: Guidance Table for Risk Register

Entry	*Guidance*
Risk Name	Title to identify the risk
Risk Description	High level description of the risk scenario and consequences Please use the following structure: Risk of [event]…due to [cause]…that may result in [consequence]
Risk Likelihood	A numeric value denoting the estimate of the probability that the residual risk will occur. The possible values are: 5 – very likely (> 70% prob of occurrence) 4 – likely (between 50% and 70% prob of occurrence) 3 – possible (between 20% and 50% prob of occurrence) 2 – unlikely (between 5% and 20% prob of occurrence) 1 – remote (< 5% prob of occurrence)
Risk Impact	A numeric value denoting the severity of the impact of the residual risk (should it occur). The possible values are: 5 – catastrophic (Critical impact impeding the achievement of the strategic objectives) 4 – damaging (Damaging impact impeding the achievement of the strategic objectives) 3 – significant (Significant impact affecting achievement of operational objectives) 2 – moderate (Moderate impact on the achievement of an operational objective) 1 – low (Minor impact on the global performance)
Risk Response Strategy	The available strategies to deal with the identified risks are: Avoid: risk avoidance, working around those conditions or activities which introduce the risks; Reduce: risk mitigation or reduction through the proactive implementation of risk reduction activities; Accept: acceptance of the risk; in these cases, contingency plans can also be defined in case the risk occurs; Transfer/share: transfer or share a risk with other entities e.g. through subcontracting, insurances etc.

Annex 1. CV's of key personnel

Annex 2. List of deliverables and milestones