

Interacting parts
of
GEMS, MERSEA and geoland

Data, products and infra-structure

(Inventory of common data user and common data requirements &
First Statement of tasks for industrial partners)

HALO
report

Draft

Version 1
Wednesday, 02 February 2005

Johannes Flemming
Tony Hollingsworth
and HALO partners





Content

1	Introduction.....	3
2	Definition of data and operability categories.....	5
2.1	Origin Categories.....	5
2.2	Operational constrains categories.....	6
3	Data covered by HALO.....	8
3.1	Identified data exchange.....	8
3.2	Un-accomplished data needs.....	11
3.3	Common data.....	13
3.3.1	Satellite instruments.....	13
3.3.2	Meteorological field data.....	14
3.3.3	In-situ data.....	15
4	System infra-structure.....	16
4.1	System architecture comparison.....	16
4.2	Net-working.....	18
4.3	Data volumes and reception times of satellite data.....	18
5	HALO Documentation.....	22

1 Introduction

The report gives a united summary of data requirements, production and system architecture of the integrated projects GEMS, MERSEA and geoland with respect to their interacting parts. The main purpose of the report is to provide a starting point for developing candidate solutions for a common approach to data acquisition, sharing and dissemination by the industrial partners in HALO. The HALO team at ECMWF will guide the work of the industrial partners in 2005.

Within HALO¹'s thematic analysis, individual reports about GEMS, MERSEA and geoland and ECMWF as well as a paper summarising the IPSI contributions to the carbon theme have been prepared. Based on available documentation, those reports aim at providing (i) an overview of the project and its data requirements, (ii) an inventory of products and in-put data and (iii) information on system architecture and data formats. The report on the carbon theme presents the contributions of the IP to the carbon cycle as to a common implementation of a global carbon data assimilation system, which was proposed in the IGOS²-report³ on the carbon theme. All reports are published on the HALO web-site⁴.

The first HALO workshop was held 16.-17. November 2004 in Reading. It was a forum for discussions among the HALO-partners and further GEMS stakeholders such as ESA and Eumetsat. Working group 1 and 2 focused on the interaction between MERSEA and GEMS and geoland and GEMS respectively. Working group 3 discussed future developments in the GMES framework. The presentations of the first workshop day and the minutes of the working group discussions are accessible on the HALO web-pages.

The present report is based on the individual and already published reports on the IPs. Focus is put on the interacting parts of the land, ocean and atmosphere segments in terms of data and infra-structure. Furthermore, the outcome of the HALO work shop was input to the contents of this report since the scientific assessment of potential links between the IPs, the definition of the HALO scope and infra-structure issues have been dealt with in the discussions.

The report starts in chapter 2 with definitions of data and operability categories. Chapter 3 list data and products covered within the HALO work. These include products to be exchanged, common data needs and un-accomplished data demands. The last chapter provides an overview of the planned data-infra structure of the IPs.

According to definition of the HALO work scope, only the global and continental or basin scale model and data activities with operational commitment will be considered. In particular these activities are:

¹ HALO is a GMES specific support action (SSA), which will be responsible for the harmonized coordination of the interaction between the atmosphere (GEMS), land (geoland) and ocean (MERSEA) integrated projects of the GMES backbone. The objective of HALO is to enable an efficient and coherent data exchange between the relevant IP's of GMES. HALO will provide agreed recommendations for the data architecture of GMES in its operational phase starting in 2008.

² Integrated Global Observing Strategy

³ <http://ioc.unesco.org/igospartners/Carbon.htm>

⁴ http://www.ecmwf.int/research/EU_projects/HALO/detail#Documents



- MERSEA's core models (Mercator, FOAM, MFS and TOPAZ)
- geoland's global observatories, mainly Observatory Natural Carbon (ONC) and Core Service Geophysical parameters (CSP)
- GEMS operational and data assimilation system at ECMWF from the GHG⁵, GRG⁶ and AER⁷-subprojects.

The main technical challenge will be to coordinate efficient links between GEMS and MERSEAS modelling activities as well as between GEMS and an operational follow up on geoland's CSP. Further, a smooth interaction of ONC and GEMS activities in the framework of ECMWF's model has to be ensured. Designing a global carbon data assimilation as proposed by IGOS from the contributions of the IPs could become a scientifically based framework for planning the link between the projects.

Besides the problems related to the direct product exchange, strategies have to be developed for common problems such as handling of large satellite data volumes from future and present satellites and operational access to in-situ data from many different providers.

⁵ Green house gases

⁶ Global reactive gases

⁷ Global aerosol

2 Definition of data and operability categories

The data and products can be categorised in terms of their origin, the associated operational constraints and the level of interaction or communality among the IP. The latter categorisation expresses different levels of involvement of the HALO-SSA in the setup of the operational systems of the IPs. The first two criteria are generic for any data and product classification. The classification according to these criteria is prerequisite for a proposal of an appropriate data infrastructure for data handling covered by HALO.

The proposed criteria and categories for data and products are:

Origin

- Observation (In-situ, Satellite)
- Model and assimilation products (global, regional)
- Service information products for users from model products (aggregated information or for special purposes such as oil spill forecast)

Operational constraints of provision

- Operational mode - real time (RT) and near real time (NRT)
- Off-line or Re-analysis mode
- Research mode

Interaction and communality - HALO aspect

- Direct product exchange between IPs
- Unaccomplished data needs from individual IPs
- Common data needs
- not to be covered by HALO

2.1 Origin Categories

All IPs use extensively satellite observations in order to monitor the status of the earth system. Satellite data are provided in different processing levels: One distinguishes (raw) **radiances**, which are more or less the measurements of the instruments after calibration and quality checking, and retrieved **products**, which are already direct data about the geophysical variables of interest. Raw radiances tend to have bigger data volumes but can be used for retrieval of more than one parameter. Derived products and raw radiances are assimilated in dynamical models describing processes. In many cases, data from one instrument can be used for retrieval of different parameters needed in the IPs. For example, VIS and NIR imagers such as Modis and

Meris can be used to retrieve atmospheric aerosol, ocean colour and vegetation status.

The data infra-structure within and in-between the IPs has to consider the operational transfer data volumes, the operational availability of satellite data in the GMES starting phase beyond 2007/8 both in terms of instruments in space and in terms of the commitment of satellite agencies to provide products in a timely manner.

All IPs use in situ data to validate the modelling and retrieval activities. Moreover, in-situ observations play a dominant role for ocean monitoring and assimilation since the status of the ocean state (temperature, salinity, currents, CO₂) can hardly be observed from satellites. Besides the validation, GEMS uses in-situ CO₂ data as input for inversion modelling to obtain carbon flux data. The challenge of the in-situ data is their collection from a large variety of data providers with different operational commitment. The Coriolis ocean in-situ data centre (IFREMER) will act as the dedicated in-situ collector and provider for the MERSEA - project.

Producing field data of the geophysical variables of the earth system by modelling and data assimilation is the main effort in all IP. Certain products should be exchanged between the IPs to better account for process in the interacting earth system, which have not been sufficiently covered yet. To foster this link from a scientific, organisational and technical point of view is the main task of HALO.

Providing information aggregated from the products of the IPs for end-user such as government authorities, international bodies, maritime agencies or private companies is a main objective of GMES. However, the link to the end users will not be covered by HALO since it deals only with product provision between the IPs. This inter-project data - exchange could be described in GEMS terms as an “(operational) service for users” but the use of this terminology will be avoided to prevent from unnecessary confusion.

2.2 Operational constraints categories

HALO wants to support the transition of the IPs to operational status. Operational commitment imposes time constraints of the data production, transfer and storage, which have to be considered by the infra-structure proposals.

One can distinguish between operational mode, re-analysis mode and research mode. The operational mode will provide products with a certain up-date frequency that reflects the temporal variability of the parameters and the capability of the observation system. The product is either provided in real time or with a certain fixed time delay. ONC, GEMS and MERSEA will provide products in operational mode.

Re-analysis mode means the product provision for periods in the past. The aim is to provide a consistent long-term data set prepared by the latest assimilation and model techniques which uses all available observation data. Observation data, which could not be collected in operational mode can be used in re-analysis. Re-analysis data help to understand the relevant process in many realisations and to explore trends in the data. ONC (geoland) and GEMS have a strong commitment to provide re-analysis of land atmosphere fluxes and global atmospheric composition. The envisaged period of the GEMS re-analysis is the Envisat epoch (1999-2007). The re-analysis mode does not have the strong time constraints for the data transfer but it requires resources for storing huge data sets and finally for disseminating them to different users.



At ECMWF, the ERA-40 Reanalysis project (FP 5) has built substantial software systems to undertake extended reanalyses of historical meteorological data (1945-2001) and provide access services to the validated results. These capabilities will be used to produce the GEMS and geoland/ONC reanalyses and to provide access to the outputs and to visualise the outputs.

Research modus data demands will be dealt with by HALO to support the transition to re-analysis or operational mode in the near future. HALO will support the request of unaccomplished data need by lobbying for their provision. Possible examples are urgently needed river run-off data for MERSEA's models and rain-gauge observation not available on GTS for geoland.

3 Data covered by HALO

The thematic analysis defined three data categories concerning communalities and interaction, which should be covered by HALO:

- Direct product exchange between IPs
- Unaccomplished data needs from individual IPs
- Common data needs

Figure 1 shows a schematic of the flow structure of the data categories and introduces example for each category. The data groups can now be further classified in terms of their origin and their associated operational constraints. The letter categorisation often alters during the development of the project. Data currently used in research-mode will be used in operational near-real time or re-analysis mode.

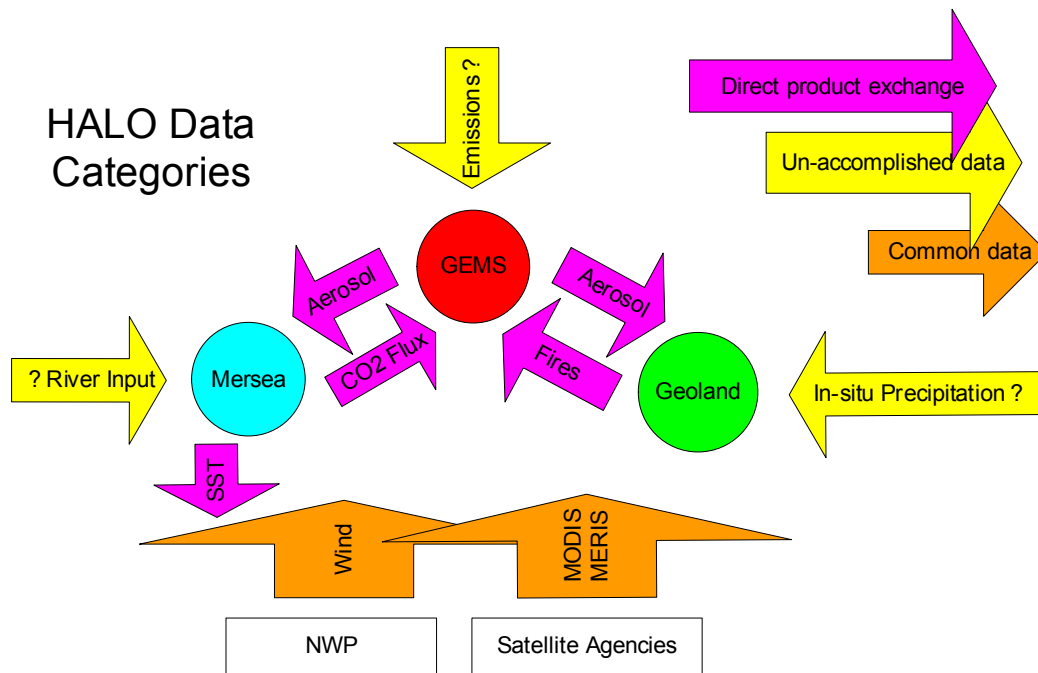


Figure 2 Flow structure of different data categories covered by HALO, including selected examples.

3.1 Identified data exchange

Coordinating envisaged direct product exchange between IPs in operational or re-analysis mode has the highest priority within the HALO activities. Besides the common need of meteorological forcing data for all IPs, no data transfer or interaction has been explicitly planned at the proposal stage of all HALO IPs.

The thematic analysis of the HALO partners has identified a list of products, which could be exchanged in the near future. Model enhancement and comprehensive

validation as well as quality assurance are still needed to use the additional information in operational or re-analysis mode.

An example of the inter-dependencies of the IPs is their contribution⁸ to the carbon cycle. ONC and GEMS carbon models could become building blocks of a Global Carbon Data Assimilation system as proposed in IGOS Carbon theme report. MERSEA could provide the missing ocean part once their ecological modelling activities have been validated. Both GEMS and ONC will be part of ECMWFs operational model. The overall structure of the proposed system is depicted in figure 2.

The exchange between the IPs will be mainly between GEMS and geoland (ONC and CSP) as well as between GEMS and MERSEA. Since both ONC and GEMS models will be hosted at ECMWF, no external link will have to be established. MERSEAs connection for receiving NWP data is already well established and need no support from HALO. However, MERSEAs global SST forecast could be utilised in Numerical weather prediction centres, especially for the purpose of monthly and seasonal forecast.

In order to organise product exchange between the IPs one has to know their temporal and spatial resolution, the production delay and, most important, the date when the production will start. These meta-data are provided in Table 1 for the products from envisaged operational production. Table 2 contains this information and additional the temporal coverage for data from re-analysis mode production. The operational mode production will start in 2006 for ONC and MERSEA and possibly a year later for GEMS. ONC and CSP re-analysis periods last till 2003/4 whereas GEMS will produce re-analysis mode data till the year 2007.

The following products have been identified for direct exchange between the IPs. The list specifies the IPs, the product and its purpose in the project receiving the data.

GEMS -> geoland

- Global aerosol forecast, re-analysis (AER)
 - for atmospheric corrections of CSP retrievals
 - for improved CSP radiation and albedo products
- Atmospheric CO₂ fields from GEMS GHG
 - for constraining biospheric fluxes from ONC

geoland -> GEMS

- Biospheric - land CO₂ flux from geoland (ONC)
 - for CO₂ modelling and assimilation in GEMS-GHG
 - for first guess/constrain for GEMS-GHG inversion modelling
- Burnt areas (CSP, so far only on continental scale)
 - for aerosol emissions in GEMS-AER (Global fire data still needed)
- Precipitations (CSP)

⁸ see http://www.ecmwf.int/research/EU_projects/HALO/pdf/paper_HALO_carbon_draft.pdf



- for wet deposition from Aerosol and Gases
- Soil moisture (CSP/ONC), vegetation status (ONC) and snow mask
 - for wind blown dust aerosol emission in (AER)
 - for dry deposition in (GRG, AER)
 - for vegetation emissions (GHG, GRG, AER)
 - for surface correction in atmospheric retrievals

MERSEA -> GEMS

- Maritime CO2 fluxes
 - for CO2-modelling and assimilation in GEMS GHG
- Wave status
 - for emissions of sea spray aerosol and gas-air exchange coefficients

GEMS -> MERSEA

- Global aerosol data (forecast, re-analysis) from GEMS-AER
 - for atmospheric corrections in retrieval in MERSEA
 - for PAR input calculations in MERSEA
 - iron input for ecological ocean modelling in MERSEA

MERSEA -> NWP

- SST data for NWP centres
 - for monthly and seasonal forecast

PRODUCT	SPATIAL RESOLUTION	TEMPORAL RESOLUTION	PROVISION DELAY	PRODUCTION START
geoland - ONC Terrestrial CO2 flux Root zone soil moisture	0.5 degree in tiles	daily	1 -2 week	2006
MERSEA Mercator model fields Core models fields	1/12° (ecological model ¼°.)	weekly	none	2006 (runs already operational)
GEMS GHG field GRG fields Aerosol fields	50 km	4 times a day	none	2006
GEMS CO2-fluxes	?	?	?	2008

Table 1 Global operational mode products for envisaged exchange

PRODUCT	SPATIAL RESOLUTION	TEMPORAL RESOLUTION	TEMPORAL COVERAGE	PRODUCTION START-END
CSP				
Precipitation	1 degree	1 day	1997-2003	2004-2005
Soil Moisture	50 km	10 days	1992-2004	2004-2005
Burnt areas	1 km (Africa, Eurasia)	1 month	2000-2003	2005-2006
SW radiation	50 km	4 times a day	2000	2006
ONC				
Carbon flux	0.5 degree in tiles	daily	1997-2003	2006-2007
Soil moisture	ditto	ditto	ditto	ditto
Vegetation	ditto	ditto	ditto	ditto
GEMS				
Aerosol fields	50km	4 times day	1999-2007	2006-2008
GHG fields	50km	ditto	ditto	ditto
GRG fields	50km	ditto	ditto	ditto
CO2-fluxes	?	?	ditto	2007

Table 2 Global Re-analysis mode products for envisaged exchange. CSP LW and SW products are available for Africa and Europe.

IGOS_P / International Global Carbon Observation Strategy Development of Carbon Cycle Data assimilation systems

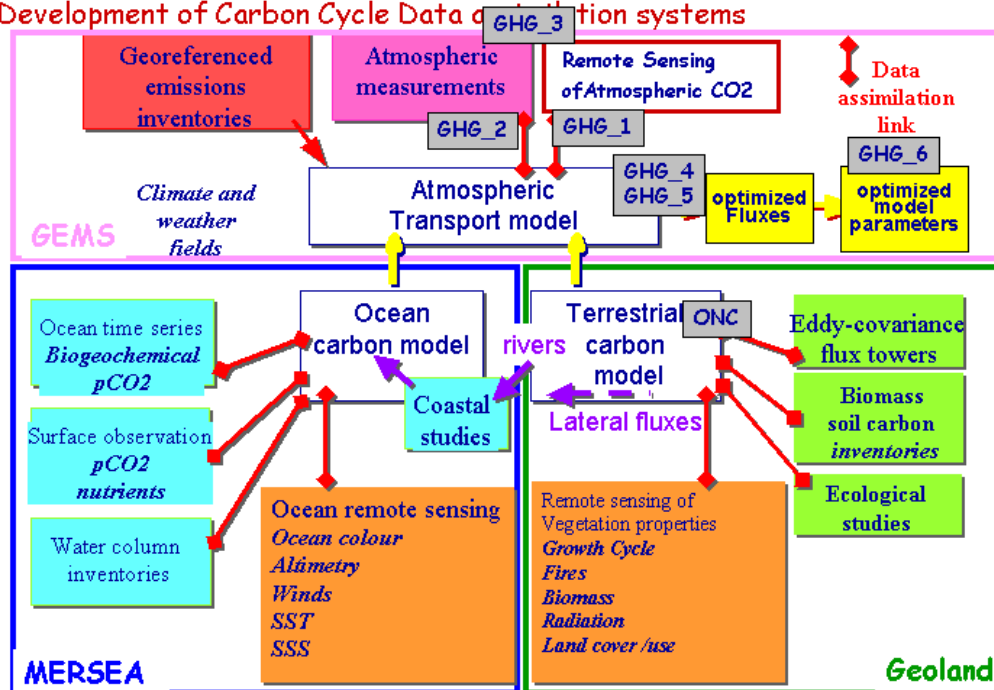


Figure 3 Schematic of Carbon Cycle data assimilation system proposed by IGOS-P annotated by specific GEMS work package activities (GHG 1 to 6) and by the ONC activity of GEOLAND.

3.2 Un-accomplished data needs

Unaccomplished data needs concern data, which can not be provided by one of the other IPs but are needed for operational, re-analysis or research purpose. HALO will support the request for accomplishing these demands by lobbying for their provision.

This can either mean exploring links to existing activities capable to provide the required data or to recommend their initiation. In some cases, such as rain gauge observation and air quality observation, only the coordinated data access to already existing observation networks has to be ensured.

The cooperation with other FP 5 & 6 projects or GEMS activities can help to match the data request. For example, the Carbo-ocean project (FP6) could provide maritime and atmospheric in-situ CO₂ observations and a GEMS water IP could provide river run-off data for MERSEA.

geoland unaccomplished demands:

- Operational access to in-situ precipitation data from climatological and hydrological network (not available on GTS) in Europe and other parts of the world
- Operational access precipitation radar data (CSP)
- Snow depth in-situ observations
- Global vegetation product from EO (LAI, FAPAR, FVC) for ONCs operational mode (CSP only in re-analysis mode for the 1998-2003 period)
- Global soil moisture product for ONCs operational mode (CSP only in re-analysis mode for the 1992-2000, 2000)
- Snow mask data - operational mode

MERSEA unaccomplished demands:

- River run-off data for freshwater and substance input
- Iron content extracted from Aerosol data (size and age of aerosol is important, still an research issue)
- High resolution atmospheric flux (stress, heat) data
- Ocean wave status for air - sea exchange & coupling (still an research issue)
- More in-situ observations of biogeochemical parameters of the ocean

GEMS unaccomplished demands:

- Surface land and ocean status products:
 - Emission for GEMS - GRG, GHG and AER data from fires (fire product and emissions parameterisation needed)
 - Aerosol emission GEMS - AER data from wind blown (soil status, i.e. soil moisture and soil type and emissions parameterisation needed)
 - Aerosol emissions from sea spray (fires) - (ocean wave status and emissions parameterisation needed)
 - Surface properties for dry deposition
 - Precipitation product for wet deposition of Aerosol and reactive gases
- In-situ observation for GHG (CO₂) , GRG , AER, RAQ (operational access and in research mode for validation)

3.3 Common data

Common data needs are a strong indirect link between the IP because a shared solution for problems associated to certain data could be beneficial for all IP. Meteorological forcing data and commonly exploited satellite instruments are the most important entities in this category. CO₂ in-situ measurements are a further common data demand. Satellite data and meteorological forcing data tend to be of large volumes but the present links for transfer and data handling seem to be well established.

Although IP use often the same instruments, the approach of the IPs towards satellite data processing differs: MERSEA has parameter- specific retrieval centres which provide satellite products for assimilation to the modelling centres. GEMS fosters the assimilation of direct (raw) radiance data in its operational model at ECMWF without the intermediate retrieval. The Core Service Geophysical Parameters in geoland acts a single data and retrieval centre for the demands of the three global observatories.

A common strategy, possibly coordinated by HALO, might be helpful to tackle upcoming challenges such as future availability of sensors, increasing data volumes due to high-resolution instruments or retrievals using best available information on atmospheric correction. HALO may work out an agreed position towards the instrumentation of future missions such as ESAs Sentinel satellites class for consultation with satellite agencies or the GMES advisory committee.

Proposals for a common practicable data infrastructure may be needed to cope with the future challenges and to design a common operational GMES system. However, instead of proposing one single system architecture, pragmatic candidate solutions for specific problems might be a more realistic approach.

3.3.1 *Satellite instruments*

Table 3 is a list of instruments onboard current and future satellite missions, which are used or could be used in at least two of the IPs GEMS, MERSEA and geoland. It contains the main application of the instruments in the respective projects.

The list is a compilation of the IP specific instrument tables of the reports on the IP. With the exception of IR atmospheric sounders for GEMS (e.g. AIRS, CRIS), most of the typical EO-instruments are used in more than one IP. A large overlap appears between geoland and MERSEA since both strongly rely on imaging instruments in the VIS - NIR and MW range.

Given that Envisat, SPOT, Auqua and Terra will reach the end of their nominal lifetime by 2008, all IPs face the common problem of a lack of continuity in Earth observations. A common approach for the use of operational meteorological satellites of the MSG, MetOp, MTG, NPP and NPOESS series might be beneficial for a transition to operational status. Eumetsat EO data are provided via RMDCN/GTS and/or the Eumetcast facilities.

The Sentinel satellites from ESA are foreseen to provide the missing continuity of EO data for GMES. User consultations are taking place and funding will be decided in 2006. The sentinels (launch from 2008) will rely on proven instruments and technology, which allows a quick transition from planning state to launch. The currently proposes missions are:

1. SAR family
2. Superspectral imaging family
3. Ocean monitoring family embarking a wide swath multi-spectral sensor as well as an altimeter
4. A geostationary family for atmospheric composition and trans-boundary pollution detection
5. An atmospheric composition monitoring family in low Earth orbit

INSTRUMENT	TYP	MISSION	EXPECTED LIFETIME	GEMS	MERSEA	GEOLAND
ATSR	IR - MW radiometer	ERS-2	1995-2003	AER	SST	Burned areas
AATSR	IR - MW radiometer	Envisat	2002-2008	AER	SST	Burnted areas
ASAR	MW active, SAR	Envisat	2002-2008		Wind, ice	HR imaging
SAR	MW active, SAR	ERS-2	1995-2003		Wind, ice	HR imaging
SCAT	MW active, Scatterometer	ERS-2	1995-2003		Wind	Soil moisture
AMSR-E	MW passive	Aqua	1999-2008		SST, ice	Vegetation, Snow
MIRAS	MW passive (L-Band)	SMOS	2007-2009		Salinity	Soil moisture
MERIS	VIS - IR radiometer	Envisat	2002-2008	AER	Ocean Col.	Vegetation, Burnt areas
AVHRR/3	VIS - IR radiometer	NOAA-14-17 MetOp	1997-2003 2005-2020		SST	Vegetation, Imaging, LW-radiation
VEGETATION	VIS - NIR radiometer	Spot 4 Spot 5	1998-2005 2002-2010		Ocean Col.	Vegetation
Seawifs	VIS radiometer	Orbview-2	1997-2002	AER	Ocean Col.	Vegetation
VIIRS	VIS-IR radiometer	NPP NPOESS	2006-2012 2010-2020	AER	SST	Vegetation
MODIS	VIS-IR radiometer	Aqua Terra	2002-2008 1999-2006	AER	Ocean Col. SST	Vegetation
SERVI	VIS-IR radiometer	MSG geostationary	2004-2010, 2020	AER, GRG	SST	Vegetation, Radiation

Table 3 Satellite instrument commonly used in the GEMS, MERSEA and geoland.

3.3.2 Meteorological field data

Global meteorological field data are provided by meteorological services. They are a grid-like representation of the atmospheric state and include information from satellite observations and in-situ observation as well as from the model formulation.

Meteorological field data are provided either

- Real time forecast and analysis (ECMWF, meteo-france, UK-metoffice)

- re-analysis (ECMWF)

Real time forecast and analysis are needed by MERSEA ocean models, by GEMS and by ONC. Real time data might be necessary for atmospheric correction of retrieval activities within MERSEA and an operational follow up of geolands CSP. The meteorological data set from ECMWF could include the global data about atmospheric composition, in particular Aerosol, provided by GEMS and the land surface data (LAI, CO₂ flux, soil moisture) from ONC.

In the case of ONC and GEMS, the modelling activities are already part of the meteorological model and there is no need for a designing a link. MERSEAS meteorological forcing data for the core models are provided by ECMWF, meteo-france and UK met-office. This link is already well established and does not need further involvement of HALO according to present assessment.

A further data stream is re-analysis data, which are used for re-analysis mode activities and retrievals. Many CSP retrieval rely on Era-40 data from ECMWF. The FP 5 ERA-40 project produced retrospectively meteorological data from September 1957 to August 2002.

ONC and GEMS re-analysis commitment do not need efforts for the data supply due to their incorporation in ECMWF facilities .

The ERA-40 Reanalysis project (<http://www.ecmwf.int/research/era/>) has built substantial software systems to provide access services to the validated results. These capabilities will be used to produce the GEMS and ONC reanalyses and to provide access to the outputs and to visualise the outputs.

3.3.3 *In-situ data*

The main problem concerning in-situ data is the granularity of their provision. In-situ data of atmospheric composition, the carbon and water cycle come from many different research networks and projects or local authorities, which are not prepared for an operational provision. The HALO report on GEMS lists about 40 different sources for in-situ data used in GEMS. At the current state, only MERSEA operationally assimilates geophysical in-situ data being pre-processed by the Coriolis data centre at IFREMER. The other projects and MERSEA geochemical activities need in-situ data mainly for validation. However, validation will become an operational task as part of the operational services.

FP projects for carbon in-situ data are:

- CARBOOCEAN (FP 6): <http://www.pangaea.de/Projects/Carbo-Ocean/>
- CarboEurope: (FP 5) A cluster of projects to understand and quantify the carbon balance of Europe <http://www.bgc-jena.mpg.de/public/carboeur/>

First attempts to collect European air pollution in organised manner are

- Airbase, European Air quality data
<http://etc-acc.eionet.eu.int/databases/AirBaseXML.html>
- Smogwarners: selected ozone observations online
<http://www.aeat.co.uk/netcen/airqual/forecast/smogwarners/>

4 System infra-structure

This chapter provides a brief overview and comparison of the system architecture of the three IPs. According to definition of the HALO work scope, only the global and continental / basic scale model and data activities with operational commitment will be considered. In particular these activities are:

- MERSEA's core models (Mercator, FOAM, MFS and TOPAZ)
- geolands global observatories, but mainly Observatory Natural Carbon (ONC) and Core Service Geophysical parameters (CSP)
- GEMS operational model and data assimilation system at ECMWF from the GHG, GRG and AER-subprojects.

ONC, MERSEA's core models and GEMS production system are dynamic models and data assimilation system on the global scale or ocean scale with temporal resolution of hours to one day⁹. GEMS and Merses core models provide boundary condition for regional scale application within in their IPs.

4.1 System architecture comparison

A schematic of the structure and transfer lines of the interacting parts of the three IP is given in Figure 4. The schematic concentrates on the data flow from the input-data providers, to model centres and, therefore, does not show the links to the end-user. Input data is provided by satellite agencies, weather services and in-situ observation systems. The raw in-put data is process and retrieved by data centres within the IP, symbolized by a star-like shape, and passed over to the model centres, symbolized by circles.

All systems largely use and assimilate satellite observations but only MERSEA's operational modelling and assimilation activities rely on in-situ observation as well, which have to be transferred in a timely manner too. All IPs use in-situ observation for validation purposes, and GEMS uses in-situ observations together with the assimilated atmospheric CO₂-observations as starting point for inversion models to estimate carbon fluxes.

The strategy for the retrievals differs between the IP. MERSEA appointed dedicated centres for retrieval of altimetry, SST, ocean colour or sea ice. They transfer their products for assimilation to the four model centres in a crossover type topology.

GEMS fosters an approach for direct assimilation of radiance data into ECMWFs model without the intermediate retrieval step of the product retrieval. However, in the cases of certain trace gases and aerosol, retrieved products, provided from satellite agencies, will be assimilated in the GEMS model. Besides the assimilation in a dynamic atmospheric model, GEMS includes efforts for independent retrieval based on neural networks.

geoland establishes on dedicated centre (core service) for serving the input-data needs of the global observatories. The core service provides the data for different periods

⁹ Certain products, e.g. ONC carbon fluxes, might be issued in an aggregated form in a lower temporal resolution. Nevertheless, the model operates in 3-hour time steps.

and areas to the observatory in a star-like network topology. The predominate data sources are satellite data, in situ data are used to validate the products.

According to present planning, geoland CSPs operates only in the ‘re-analysis mode’, which means data provision for specific years before 2004. The operational mode of the ONC land-vegetation-carbon-model will be ensured by embedding it in ECMWF integrated model. ONCs demand for the operational provision of a global satellite vegetation product has to be satisfied in the operational phase.

GEMS sub-projects GHG, GRG and AER and geoland’s ONC include extensive re-analysis production for previous periods, in particular the Envisat-epoch. The main challenge in the re-analysis mode is the provision of the large in-pu data sets (meteorological forcing and satellite data) and archiving capabilities for the output.

The meteorological forcing data are needed for MERSEA-models are supplied from either ECMWF or UK-met-office in the case of the Foam model. The data is transferred via GTS.

The challenge for HALO will be to propose a sufficient link between GEMS and MERSEAs modelling activities as well as between GEMS and a operational follow up on geolands CSP. Further a smooth interaction of ONC and GEMS activities in the framework of ECMWFs model has to be ensured.

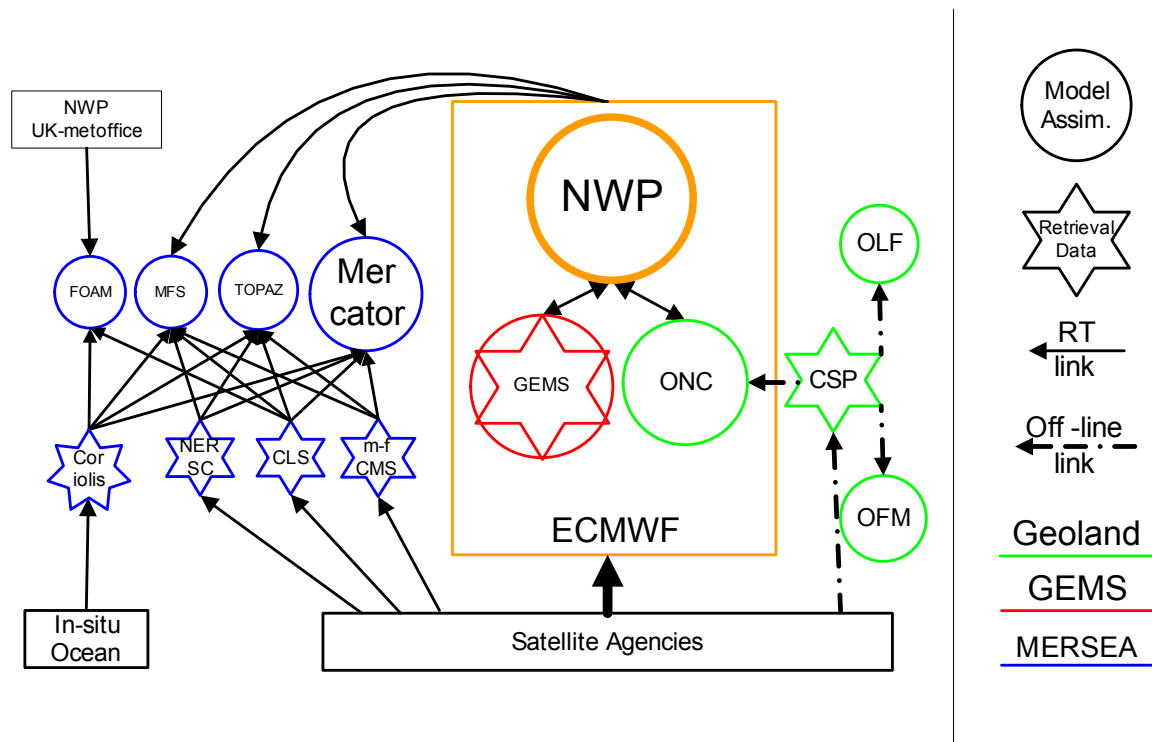


Figure 4 System structure and data flow of the operational model activities in MERSEA, geoland and GEMS.

4.2 Net-working

According to present planning, data transfer in the three IP are based on the following networks:

- GTS/RMDCN
- Standard internet/ FTP
- Dedicated Networks such as EumetCast or GEANT

GTS (Global telecommunication system) is the network of the meteorological community. The access to GTS is restricted and governed by WMO rules. RMDCN, which is part of GTS, is the European Regional Meteorological Data Communications Network between European meteorological centres including ECMWF, EUMETSAT and CIMSS/NESDIS. The GTS consists of an integrated network of point-to-point circuits, and multi-point circuits which interconnect meteorological telecommunication centres. Data on GTS are encoded in GRIB (fields) or BUFR (observations) format.

MERSEA access to meteorological filed data and-in situ observations relies on GTS. The links are provided by meteorological centres such as Meteo-France, UK-Metoffice or Met.no which often host the ocean modelling activities. GEMS and ONC, both hosted at ECMWF, will have access to GTS too.

Standard ftp connection are used for all kind of data exchange not covered by GTS or dedicated networks. geoland links between CSP and the global observatories uses mainly this connection type. ECMWF receives about 50% of its satellite data via this network from data sources such as the research platforms run by ESA and NASA.

Data transfer via Internet using the Opendap protocol (dods) allows a flexible way to exchange data from various data centres. Opendap will be used in MERSEA . It has been developed for the needs of the ocean community.

Data formats based on NetCDF are the most common way to decode the data sets. Live Access Servers using OPeNDAP are widely used in the Ocean community. They provide a good possibility to visualise gridded data from different sources. Their application could be expanded in order to access field data of all three HALO IPs. Opendaps capabilities for real time data transfer has to be investigated.

The utilisation of dedicated networks such EumetCast for EUMETSAT- products or the European high speed research network GEANT has to be explored further in the course of the HALO work.

4.3 Data volumes and reception times of satellite data

The main challenge for operational production might become the acquisition of satellite data in a timely manner. This paragraph, compiled from an ECMWF perspective only, communicates anticipated and current satellite data volume numbers for the GEMS project as well as typical reception times of selected satellite data. At the current stage, no information about equivalent numbers from MERSEA and geoland is available. It can be assumed that their demands will be of similar or smaller size.



Figure 3 shows in summary the estimated daily transfer rates of satellite radiance data needed for GEMS. Figure 4 shows the transfers summed up for retrieved products used in the GEMS subprojects.

Much of the high volume sounding data required for GEMS (IASI, 1500 Mb/day) is already being received in real-time, or will be received in real-time at ECMWF. The main challenge will be the transfer of MODIS (2 x 600 MB/day) and MERIS radiance data. Most of the satellite data to be used in GEMS will come from satellite agencies (ESA, EUMETSAT, NASA) and in the case of some products from research centres (KNMI) and universities (Universitaet of Bremen, Heidelberg).

The operational production in real or near real time mode requires data reception within a given time window. The end of this time windows is called cut off time since data arriving later can not be used anymore in the model and assimilation run. The reception of the satellite data is controlled by the dissemination performance of the provider and the network capabilities. Satellites products tend to be more slowly distributed since time for processing is needed. Figure 5 shows average arrival times at ECMWF of AIRS, MSG, ASAR and Meteosat data for the 0 UTC model run with a time window from 21 to 3 UTC (cut off). In the given example about 90 % of the MSG (Meteosat second generation) and Meteosat data, 80% of the ASAR (Envisat) and 50% of the AIRS (Aqua) data reach the centre in time. The cut off time of 3 hrs is typical for numerical weather prediction but the GEMS production might have longer cut off times.

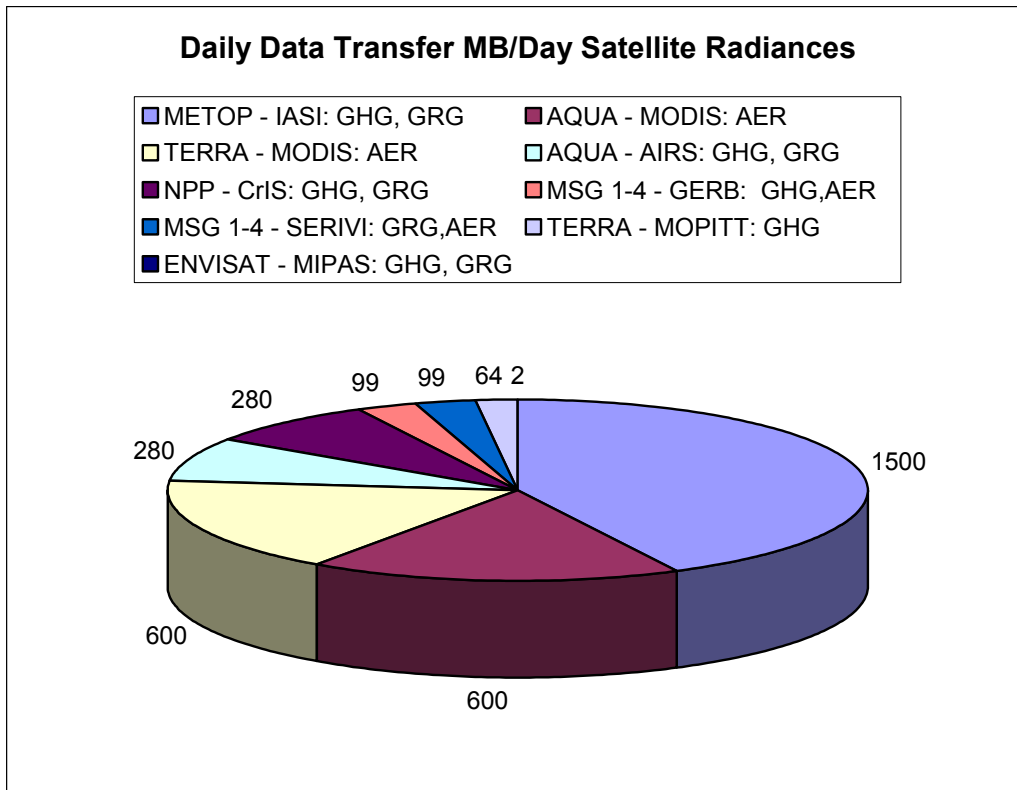


Figure 5 Estimated daily transfer in MB/day of satellite radiance data for GEMS.

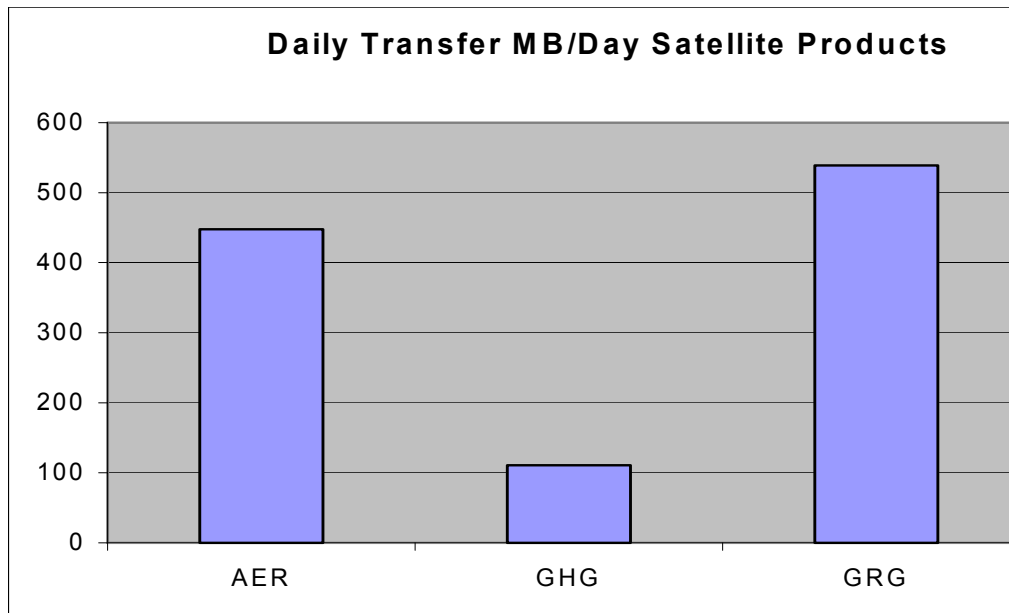


Figure 6 Estimated daily transfer in MB/day of satellite products need in GEMS subproject Aerosol (AER), Greenhouse gases (GHG) and Global Reactive Gases (GRG) .

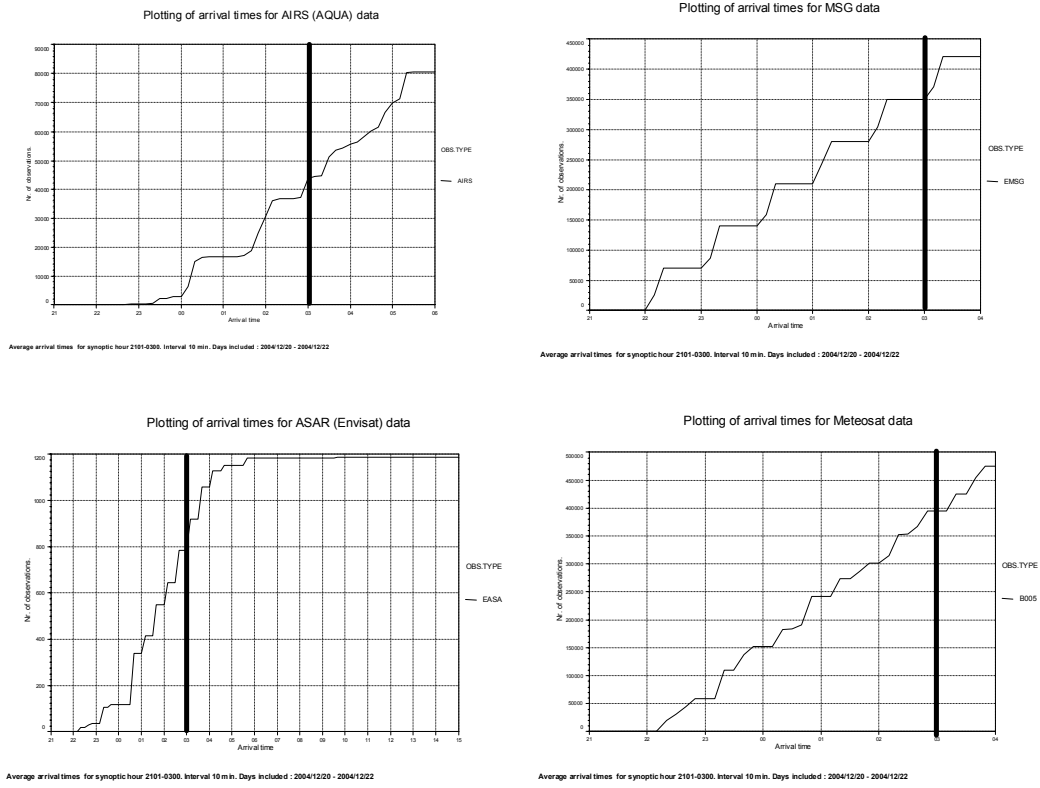


Figure 7 Arrival times for satellite data at ECMWF for the 0 UTC cycle from AIRS, MSG, ASAR and Meteosat. The time window lasts from 21 to 3 UTC (cut off). Data arriving after cut off could not be used in this cycle.



5 HALO Documentation

HALO Reports at

http://www.ecmwf.int/research/EU_projects/HALO/details.html#Documents

- Draft report: ECMWF data and products for GMES
- Draft report: MERSEA data and products for GMES
- Draft report: Geoland data and products for GMES
- Draft report: GEMS data and products for GMES
- Draft paper: The Carbon theme in GMES

HALO Workshop documents:

- Presentations:

http://www.ecmwf.int/research/EU_projects/HALO/workshop.html#Presentations

- Working group discussion summaries

http://www.ecmwf.int/research/EU_projects/HALO/pdf/workshop_WG_report.pdf