

D9.1 - IS-ENES - ACADEMY OF ATHENS

REPORT ON EVALUATION DATASETS, METHODS, AND PROBLEMS FROM A SURVEY OF EUROPEAN ESM GROUPS

In order to document the evaluation process performed by different ESM groups, questionnaires were distributed and a database was created to store the responses of the modelers. The procedure is still in progress as all the questionnaires have not been returned yet. In this report, the creation, structure, and function of the database are explained. Preliminary results as they can currently be extracted from the database are presented.

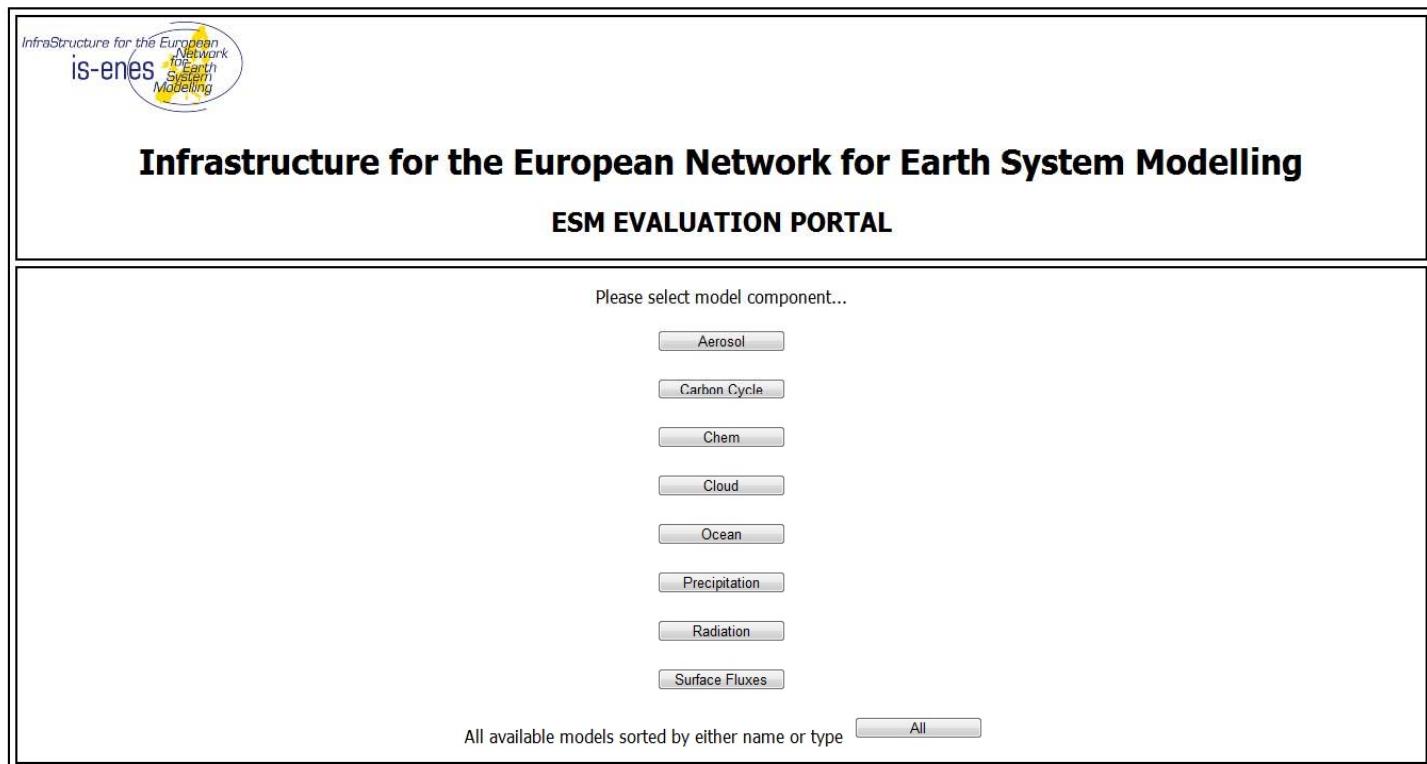
1. Evaluation Survey

In order to collect the needed data, a template questionnaire was created. The questionnaire consisted in two parts. The first part of the questionnaire was introductory. It was the same for all questionnaires and it consisted of three questions concerning identification of the model, contact details and general model information. The second part targeted the scientific content and was different for each of the following model components: aerosols, carbon cycle, atmospheric chemistry, clouds, ocean, precipitation, radiation and surface fluxes. The focus was given on the evaluation of the model results using observational data. The second part of the questionnaire consisted of eight questions with similar structure for each model component. The homogeneity of the questionnaires was intentional. It facilitates the categorization of the data and their entry in a common database as the number of the necessary tables in the database is minimized. It also allows the database to be more effective in the entry and extraction of data as well as in creating statistics not only by category but also by model component. A sample questionnaire can be downloaded from this [link](#).

The questionnaires were sent to the five participant ESM groups: CMCC, CNRM, COSMOS, Hadley and IPSL. More specifically they were sent to the head of each group who distributed them to the suitable persons to answer for each model component. When the answered questionnaires started to return to AA the effort of representing the data contained in the answers in the form of matrices started. The effort was successful and furthermore it was found that the questions 2, 3 and 4 could be combined so that the answers can be presented in the form of one larger matrix, thus the questionnaires could be rewritten in a new form with the second part containing six questions. Namely question 2 refers to classifying the parameters that are evaluated into categories, question 3 refers to the databases that were used for the evaluation and question 4 refers to the temporal and spatial scales.

It is worth noting that the web application code was almost entirely written in the AA. The web design part must be completed in conjunction with the rest of the portal in which the application code will be integrated with. I should also be noted that the focus was given on the application code and not on the web design part, because this application will be integrated with the ISENES portal and its design will be used as a template.

2. Evaluation Portal



Please select model component...

Aerosol

Carbon Cycle

Chem

Cloud

Ocean

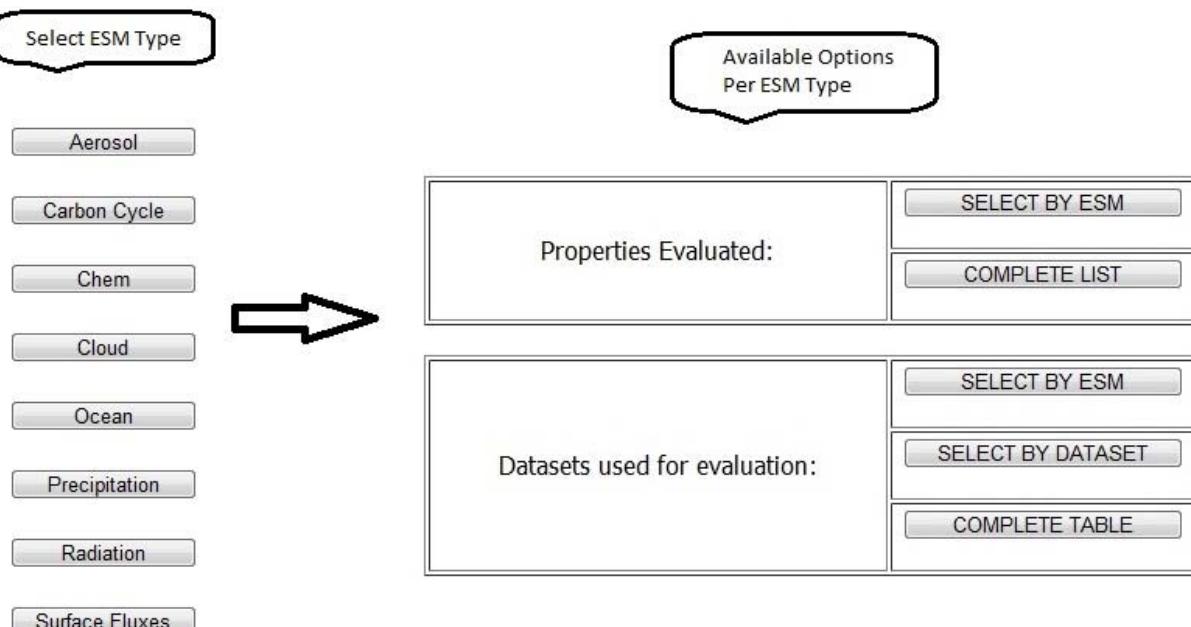
Precipitation

Radiation

Surface Fluxes

All available models sorted by either name or type

The main page is divided in two parts, the first one giving us the choice of model component.



After selecting the model component we would like to get information of, we can select to view either Parameter Based results or Dataset Based. If selecting Parameter based results we can either get a list of available models and the parameters they get evaluated against (a drop down menu gives access to some finer details per model), or we get a list of parameters and some simple statistics consisting of basic parameters per model component, number of available subparameters per parameter and the percentage of parameter usage from the available models of this component category (there is a drop down menu here also that returns all available subparameters and the associated models per parameter type)

The Dataset based results are divided into three categories. The first option is a list of available models and the datasets they use to evaluate various parameter and the second is the reversed one – returns a list of datasets and the models that are using them for evaluation with a percentage of dataset usage on the total number of models. The last option is a list that shows all associations of available models, evaluated parameters and datasets used, space and time variability and available comments of the specific combination. Empty cells meaning that the above cell is repeated. Also, there is the color change between ESMs for usability.

ESM details

Model Name:	ECHAM
Model Type:	AEROSOL
Contact:	Ulrike Lohmann / Isa, ETH Zurich, Institute for Atmospheric and Climate Science / ETH Zurich, Centre for Climate System Modelling (C2SM)
General Info:	Under construction
Main Properties:	We use the two-moment aerosol scheme HAM that predicts the aerosol mixing state in addition to the aerosol mass and number concentrations (Ster et al., 2005). The size-distribution is represented by a superposition of log-normal modes including the major global aerosol compounds sulfate, BC, organic carbon, sea salt and mineral dust. HAM is coupled to the two-moment cloud microphysics scheme (Lohmann et al., 2007). For certain applications we also use HAM coupled to the MOZART chemistry module, which includes a rather detailed description of both tropospheric and stratospheric chemistry.
Participation in MIPs:	Aerocom
Analysis Tools Used:	IDL (click to website) CDO (click to website) GrADS (click to website)

	Model Name	Parameter Evaluated	Dataset	Spatial Resolution	Temporal Resolution	Comments
1	ECHAM	Aerosol Optical Properties / AOD	-	GLOBAL	-	Satellite Data
2		Direct & Indirect Forcing	-	GLOBAL	-	Indirect Forcing / Satellite data
3		Aerosol Budgets	AERONET	STATION	-	In-situ data
4			POLARCAT	-	-	-
5		Aerosol Composition	AERONET	STATION	-	In-situ data
6			POLARCAT	-	-	-
7		Aerosol Deposition	-	-	-	Quantitative Information. From European/American AQ networks

In every result, the ESM name is a hyperlink to a more detailed page. We get basic ESM information (contact person, main properties, Analysis tools used) and the Model Name/Parameter Evaluated/Dataset/Space-time/Comment combination for it. It almost works as a full representation of the questionnaires used for data gathering.

All available models sorted by either name or type	<input type="button" value="All"/>
All available datasets (+statistics)	<input type="button" value="All"/>

All available Datasets and statistics

(on total of 15 Inserted/processed models)

13 Datasets are not shown as they are not associated with any model - see end of page.

69 Datasets are available in total.

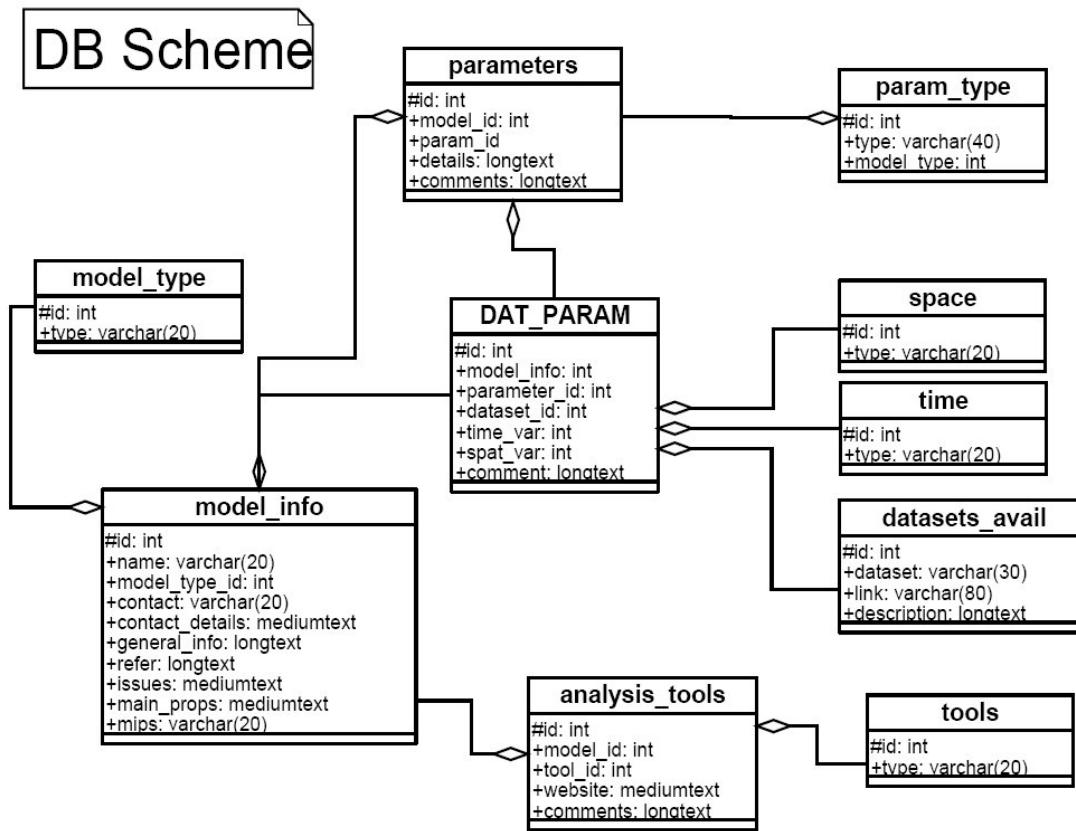
	ESM Type	Datasets	#
1	AEROSOL	EMEP, IMPROVE, AMSTERDAM ISLAND, CAPE GRIM, University of Miami Network, AERONET, MODIS & MISR, CASTNET, EANET, AEROCE, POLARCAT	11 out of 69
2	CARBON CYCLE	NCEP-DOE AMIP-II, ISLSCP climatology, NOAA/ESRL/GMD, Houghton et al., GLODAP DB, SeaWiFS ocean color data, JGOFS, ClimPP, VGPM, TRANSCOM, Takahashi DB (SOCAT), Modis, Other satellite data, Litterbag project, CARBOOCEAN database, EUROFLUX, AMERIFLUX, EMDI, Norby et al., IGBP, Zinke et al., Olson et al., GEOLAND Project, Sabine et al., Carr et al., Kettle et al.	26 out of 69
3	CHEM	UARS, MOPPIT CO, GAW database, Aircraft measurements, MIPAS, SCIAMACHY, NIWA, TOMS	8 out of 69

Name	Model Type
C-ECM	CARBON CYCLE
	CLOUD
	PRECIPITATION
	RADIATION
CNRM	CARBON CYCLE
	CHEM
	CLOUD
	PRECIPITATION
	RADIATION
CCSM3	AEROSOL
	CARBON CYCLE
ECHAM	AEROSOL
HadGEM	AEROSOL
	CARBON CYCLE
IPSL	CARBON CYCLE

	Dataset	Used to evaluate:	Used by following models:	Usage %
2	AEROCE	Aerosol Composition (AE Composition) Aerosol Budget (AE Budgets) Aerosol Deposition (AE Budgets)	COSMOS (AEROSOL)	6.67 %
3	AERONET	Component Aerosol optical depth (AE Optical Properties) Single scattering albedo (AE Optical Properties) Aerosol Optical Properties (AE Optical Properties) Aerosol Budgets (AE Budgets)	HadGEM (AEROSOL) COSMOS (AEROSOL) ECHAM (AEROSOL)	20.00 %

Datasets not listed above are: SHADOZ, LOGAN, TES, AIRS, HALOE, TRMM, MOPITT, ODIN,.....

Lastly, from the homepage we can access a full list of available models inserted (sorted either by model name or by model type) and some simple statistics on a dataset basis. In detail, we can see the total number of available datasets, which ones are used per model type, the parameters they are used to evaluate and the usage percentage based on the total number of available models. At the bottom of the page, we can see a listing of datasets that are not used by any model.



This is the scheme of the database used. Main table is DAT_PARAM which holds the Model/Parameter/Dataset/Space-time/Comment combination information. The building blocks are the tables model_type (Aerosol, Carbon Cycle, etc), param_type (Budget, Optical Properties etc), space (Global, Grid, etc), time (seasonal, mean, etc) datasets_available (dataset name, link to home page and short description), tools (IDL, GrADS, etc). In a level between those two, come the model_info, parameters and analysis_tools, which hold the necessary associations between models used, parameters evaluated and analysis tools used.

3. Evaluation Results

In total 40 questionnaires were distributed (five ESMs times eight model components). From those questionnaires 26 were returned answered to the AA. For this reason the following results are preliminary and they are expected to be altered when the total of 40 questionnaires are returned answered. The main findings are following.

	CMCC	CNRM	COSMOS	Hadley	IPSL
Aerosols	X		X	X	
Carbon Cycle	X	X	X	X	X
Chemistry	X	X	X		
Clouds	X	X		X	X
Ocean					
Precipitation	X	X		X	X
Radiation	X	X		X	X
Surface Fluxes	X	X	X		

Table 1: The questionnaires that have been returned answered per ESM and per model component

- **Aerosols.** 3 ESMs have answered which use in total 11 different databases. AERONET is the most popular used by all ESMs while EMEP, IMPROVE and MODIS & MISR are used by 2 out of 3.

Issues and problems are reported by ESM COSMOS. Gridded MODIS products in HDF format cannot be used with Ferret, the software tool that they use usually. Also an ESM group which is not listed, referred to as ECHAM, reports problems concerning the consistency between different satellite datasets. Also they report lack of consistency between different satellite datasets.

- **Carbon Cycle.** All 5 ESMs have answered and they use 27 databases. GLODAP and NOAA/ESRL/GMD are the most popular data bases used by 60% of the ESMs.

Issues and problems are reported by 2 ESMs.

ESM COSMOS in the evaluation procedure caused not by the C-cycle data, but by the missing knowledge on land biosphere heterogeneity and process knowledge on C-cycle dynamics.

ESM Hadley emphasizes the difficulties of using inversion estimates for 1980s/90s because this can only be accomplished by after running the model from pre-industrial years. Ditto for CO₂ trend. Also some processes are still unclear and this results in added difficulty, e.g. it is unclear how much of present day uptake is forest regrowth and how much is deposition?

Finally they mention the problem of needing global gridded products which unfortunately are not as reliable as the site-level ones. For example global gridded NPP data are half-data and half-model which is not ideal. On the other hand site level NPP data are more reliable but their representativity is questioned.

- **Chemistry.** 3 ESMs have answered but one of them (CMCC) does not contain a chemistry-climate coupling scheme. 22 databases are used for evaluation by the other two ESMs and 5 of them (TOMS, MIPAS, SCHIMACHY, MOPITT and aircraft measurements) are used by both ESMs.

ESM COSMOS mentions the issue of the variety of data formats which forces the modelers to use several different tools in order to convert the data to the same format.

- **Clouds.** 2 ESMs have answered but one of them (CMCC) does not perform any evaluation of cloud properties and processes. CNRM uses 3 databases for the

evaluation: ERBE, ISCCP and NOAA.

- **Precipitation.** 3 ESMs have answered and they use 5 databases for the evaluation of their precipitation schemes. CMAP is the most popular as it is used by both ESMs.
- **Radiation.** Two ESMs have answered and they use 6 databases. ERBE is the most popular, used by both ESMs.
- **Surface Fluxes.** 3 ESMs answered and they use 16 databases. The most popular are CMAP and ISCCP used by 2 ESMs.

Problems and issues were reported by two ESMs.

CNRM mentions the poor quality/reliability of the datasets which is crucial for the estimation of the uncertainties. For example the available datasets often lead to unbalanced global heat budget.

Hadley ESM is concerned about the format the satellite data. Modern datasets tend to converge towards NetCDF, although the satellite community standard is HDF, which still poses some problems and requires extra work. Also, a large amount of data generated in past decades come in many different formats, which also requires extra pre-processing to convert it to a common format.