

D4.8: DATA MANAGEMENT REPORT

LEAD AUTHOR: Floor Brouwer (Wageningen Research)

OTHER AUTHORS: Arnaldo Caivano (JRC), Mariia Bogonos (JRC), Pavel Ciaian (JRC), Javier Castro-Malet (JRC), Maria Blanco (UPM), Stefan Frank (IIASA), Alexander Gocht (THUENEN), Ana Gonzalez-Martinez (WR), Petr Havlik (IIASA), Roel Jongeneel (WR), Myrna van Leeuwen (WR), Jan Peter Lesschen (WR), Hans van Meijl (WR), Petra Salamon (THUENEN), Andrzej Tabeau (WR), Peter Witzke (EUROCARE)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 773499 SUPREMA

PROJECT	Support for Policy Relevant Modelling of Agriculture (SUPREMA)
PROJECT NUMBER	773499
TYPE OF FUNDING	Coordination and Support Action
DELIVERABLE	D.4.8 Data Management Report
WP NAME/WP NUMBER	Project coordination and management / WP4
TASK	Task 4.4 (Data management)
VERSION	Final
DISSEMINATION LEVEL	Public
DATE	31/07/2020 (Date of submission)
LEAD BENEFICIARY	WR
RESPONSIBLE AUTHOR	Floor Brouwer
AUTHOR(S)	Arnaldo Caivano (JRC), Mariia Bogonos (JRC), Pavel Ciaian (JRC), Javier Castro-Malet (JRC), Maria Blanco (UPM), Stefan Frank (IIASA), Alexander Gocht (THUENEN), Ana Gonzalez-Martinez (WR), Petr Havlik (IIASA), Roel Jongeneel (WR), Myrna van Leeuwen (WR), Jan Peter Lesschen (WR), Hans van Meijl (WR), Petra Salamon (THUENEN), Andrzej Tabeau (WR), Peter Witzke (EUROCARE)
INTERNAL REVIEWER	Approval by WP leader

DOCUMENT HISTORY

Version	Initials/NAME	DATE	COMMENTS-DESCRIPTION OF ACTIONS
1	Version 1	27/7/2020	First draft, to update Deliverable D4.3 (Data management plan)
2	Version 2	28/7/2020	Draft for internal review
3	Final	29/7/2020	Final version

Table of Contents

Executive summary.....	4
Glossary / Acronyms	5
1 Introduction.....	7
1.1 Structure of the document	7
1.2 Why is a Data Management Plan needed?	7
2 DataM.....	9
2.1 The DataM Information System.....	9
Life-cycle of scientific data and DataM.....	9
The DataM data management tool	14
The DataM BI/analytics platform: Qlik-Sense	16
The DataM portal	16
DataM governance	17
Architecture of the data base.....	17
Evolution of DataM during the SUPREMA project (2018-2020)	17
Data privacy.....	17
3 Open access release of scenarios	18
3.1 Uploading data to DataM	18
3.2 Uploading data to DataM	18
Format of the data	18
Rules to provide data.....	19
Dimensions to provide model data.....	19
Unit of measurement to provide model data.....	19
3.3 Comparison of the 2030 and 2050 baselines.....	20
3.4 Comparing a 2030 CAP scenario	20
3.5 Comparing a 2050 climate scenario.....	21
3.6 Intellectual Property Rights (IPR)	21
4 Open access release of data	23
5 Conclusions.....	26
6 References.....	27

Executive summary

This deliverable presents the Data Management Report on the open access release of data from model runs in SUPREMA. Here, Open access (OA) refers to the practice of providing on-line access to scientific information that is free of charge to the end-user and reusable. 'Scientific' refers to all academic disciplines. In the context of research & innovation, 'scientific information' can mean: (i) peer-reviewed scientific research articles (published in scholarly journal) or (ii) research data (data underlying publications, curated and raw data) (see also: European Commission, 2017). For SUPREMA, the Data Management Plan (DMP) is defined as 'the development, execution and supervision of plans, policies, programmes and practices that control, protect, deliver and enhance the value of data and information assets' obtained. The Data Management Report is updating the DMP. This report describes the release of research data collected and generated during the project, and after it is completed. This also includes data to be generated, methodologies and standards, data privacy/openness, and preservation measures.

Changes with respect to the DoA

No changes with respect to the DoA

Dissemination and uptake

The deliverable is publicly available. The report explains which of the research data generated and/or collected are made open.

Short Summary of results

DataM is the European Commission's Data portal of agro-economics Modelling. It contains the outcomes of research activities, and is operated by the Joint Research Centre (JRC) of the European Commission. DataM, including the web portal but also the Information System, will be used to release model runs that are considered for open access release. A baseline comparison and harmonization action includes the models represented in SUPREMA. A medium-term (until 2030) assessment of European agricultural policy alternatives will cover CAPRI, IFM-CAP and AGMEMOD-MITERRA Europe. Finally, SUPREMA will also use different modelling tools for the long-term (until 2050) assessment of climate change goals, using GLOBIOM and MAGNET as leading models.

Evidence of accomplishment

The deliverable itself can act as the evidence of accomplishment.

Glossary / Acronyms

AGMEMOD	AGRICULTURE MEMBERSTATES MODELLING
AGMIP	AGRICULTURAL MODEL INTERCOMPARISON AND IMPROVEMENT PROJECT
BI	BUSINESS INTELLIGENCE
CA	CONSORTIUM AGREEMENT
CAPRI	COMMON AGRICULTURAL POLICY REGIONALISED IMPACT MODELLING SYSTEM
CSV	COMMA SEPARATED VALUES
DATAM	DATA PORTAL OF AGRO-ECONOMICS MODELLING
DBA	DATABASE ADMINISTRATOR
DG	DIRECTORATE GENERAL
DG AGRI	DIRECTORATE GENERAL FOR AGRICULTURE AND RURAL DEVELOPMENT
DG COMM	DIRECTORATE GENERAL FOR COMMUNICATION
DMP	DATA MANAGEMENT PLAN
EUROCARE	EUROPEAN CENTRE FOR AGRICULTURAL, ENVIRONMENTAL AND REGIONAL RESEARCH
GA	GRANT AGREEMENT
GDP	GROSS DOMESTIC PRODUCT
GLOBIOM	GLOBAL BIOSPHERE MANAGEMENT MODEL
IFM-CAP	INDIVIDUAL FARM MODEL FOR COMMON AGRICULTURAL POLICY
IIASA	INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
IPR	INTELLECTUAL PROPERTY RIGHTS
JRC	JOINT RESEARCH CENTRE
LCA	LIFE CYCLE ANALYSIS
LULUCF	LAND USE, LAND-USE CHANGE, AND FORESTRY
MAGNET	MODULAR APPLIED GENERAL EQUILIBRIUM TOOL
NUTS	NOMENCLATURE OF TERRITORIAL UNITS FOR STATISTICS
OA	OPEN ACCESS
SUPREMA	SUPPORT FOR POLICY RELEVANT MODELLING OF AGRICULTURE
UAA	UTILIZED AGRICULTURAL AREA
WR	WAGENINGEN RESEARCH

1 Introduction

1.1 Structure of the document

Section 1.2 will outline the need for a Data Management Plan in SUPREMA, and proposes the use of DataM for the release of public datasets that is traceable through the publication of related meta-data in the JRC Data catalogue and in major European open data portals. Section 2 will present the DataM Information System, including the data management tool, and the software used. Moreover, the DataM portal is presented, its governance and architecture, as well as data privacy considerations. Section 3 will summarise the steps for open release of scenarios from CAPRI, GLOBIOM, MAGNET, AGMEMOD, MITERRA-EUROPE and IFM-CAP, including IPR. The release in DataM is presented in Section 4 and some conclusions are presented in Section 5 of the report.

1.2 Why is a Data Management Plan needed?

SUPREMA participates in the Open Access and the Open Research Data Pilot of the European Research Council (ERC). From 2017 all H2020 projects will participate in a pilot project to make the underlying data related to project outputs openly available for use by other researchers, innovative industries and citizens (<https://www.openaire.eu/what-is-the-open-research-data-pilot>). According to the Open Research Data Pilot, 'Open data is data that is free to access, reuse, repurpose, and redistribute. The Open Research Data Pilot aims to make the research data generated by Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access.' (<https://www.openaire.eu/what-is-the-open-research-data-pilot>). Data will be released in open formats, with proper documentation to support their use in other research. After the project completion, and if there is no objection by any of the project partners and use anonymization is preserved, the data are foreseen to be published in an Open Data portal (for example in <http://open-data.europa.eu>) for future research.

The DMP specifies the implementation of the pilot for: data generated and collected, standards in use, workflow to make data accessible for use, reuse and verification by the community, and definition of a strategy of curation and preservation of the data. Therefore, we refer to the SUPREMA Grant Agreement (GA), Article 29.3 on "Open Access to research data":

Regarding the digital research data generated in the action ('**data**'), the beneficiaries must:

- a. deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:
 - the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
 - other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan'.
- b. provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex 1, would be jeopardised

by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.

The data management policy described in this document reflects the current state of consortium agreement on data management. Data is stored in a database developed by JRC (DataM). Project participants will have secured web access to the databases, which will have been automatically checked for consistency, homogeneity and completeness. After project completion, provided that there are no objection by project partners, DataM dataset are open to the public for future research always consistent with exploitation and Intellectual Property Rights (IPR) requirements. DataM public datasets are made traceable through the publication of related meta-data in the JRC Data catalogue (<http://data.jrc.ec.europa.eu>) and in major European open data portals (EU open data portal, <http://data.europa.eu> and European Data Portal, <https://www.europeandataportal.eu>).

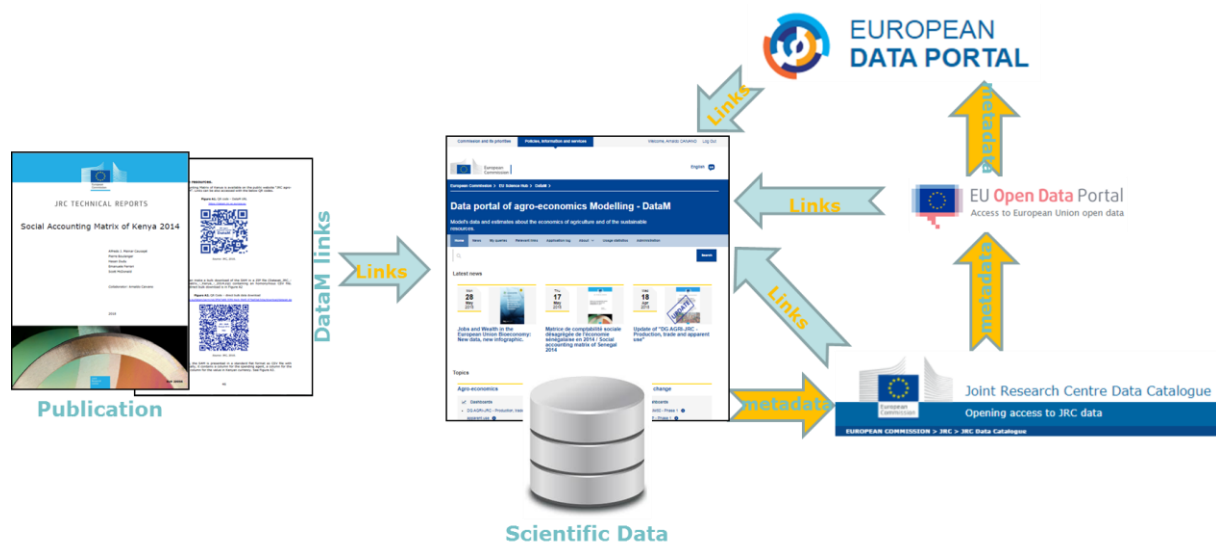


Figure 1 - DataM and open data

2 DataM

DataM is publicly known as a web site (<https://datam.jrc.ec.europa.eu>): it is the European Commission's Data portal of agro-economics Modelling. DataM contains model's data and estimates about the economics of agriculture and of the sustainable resources. By definition, DataM does not deal with official statistical data. DataM contents are the outcomes of research activities. Indeed, DataM is operated by JRC, the Joint Research Centre of the European Commission. Data is presented both in terms of raw CSV datasets, ready to download, and in the form of advanced interactive dashboards or interactive infographics that allow the self-analysis of data.

2.1 The DataM Information System

Internally in JRC, and in the context of SUPREMA, the term DataM does not refer only to the web portal but to the Information System in broader terms. The DataM Information System includes also a "data management tool" and a "Business Intelligence tool".

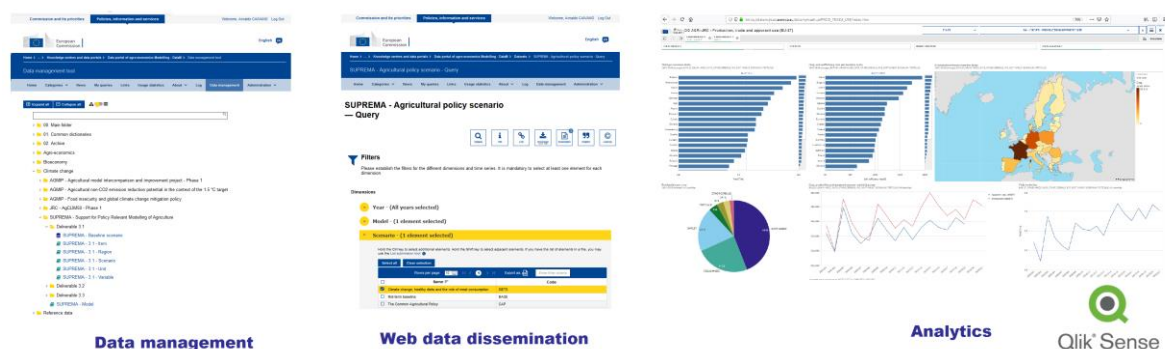


Figure 2 - DataM Information System

Life-cycle of scientific data and DataM

We can consider the life cycle of scientific data as composed by three main phases: construction, analysis and dissemination. The DataM Information System deals principally with the last two.

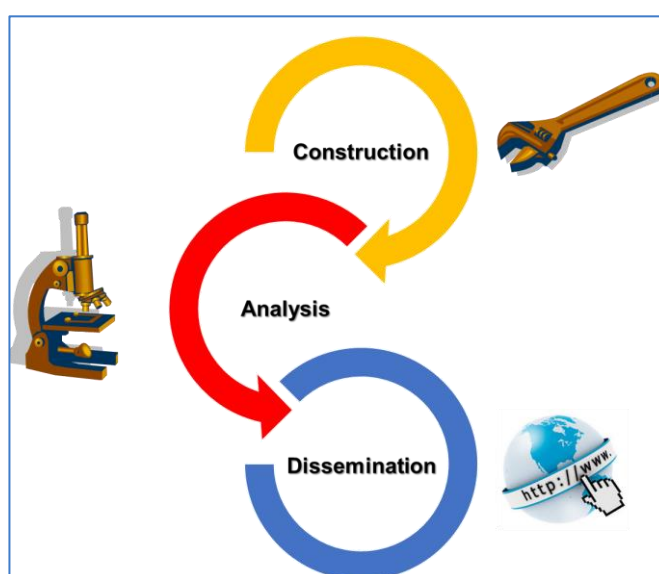


Figure 3 - Life-cycle of data

Actually the processes of strict construction of the data is arbitrary and depends on the individual scientific activities. Typical DataM outcomes are the result of modelling activities (i.e. GAMS software processing). However DataM contains also results of data processing, following a scientific methodology, over data coming from other sources. The following table (Table 1) lists the origin of DataM contents at time of writing.

#	CONTENT	PUBLICATION	RAW DATASETS	INTERACTIVE DASHBOARD	INTERACTIVE INFOGRAPHICS	ORIGIN
1	ASGTS - KENYA	2018	NO	NO	YES	MODELLING (CGE)
2	AGCLIM50 - PHASE 1	2017	YES	YES	NO	MODELLING (CAPRI, IMAGE, GLOBIOM, MAGNET, MAGPIE)
3	AGMIP - AGRICULTURAL NON-CO2 EMISSION REDUCTION POTENTIAL IN THE CONTEXT OF THE 1.5°C TARGET	2018	YES	NO	NO	MODELLING (CAPRI, IMAGE, GLOBIOM, MAGNET)
4	AGMIP - FOOD INSECURITY AND GLOBAL CLIMATE CHANGE MITIGATION POLICY	2018	YES	NO	NO	MODELLING (AIM, CAPRI, EPPA, ENVISAGE, FARM, GLOBIOM, GCAM, GTEM, IMPACT, MAGPIE, MAGNET)
5	AGMIP - PHASE 1	2017	YES	YES	NO	MODELLING (CAPRI, IMAGE, GLOBIOM, MAGNET, MAGPIE)
6	BIO-BASED INDUSTRY AND BIOREFINERIES	2020	NO	YES	NO	DATABASE COMPILED BY JRC
7	BIOSAMS EU MEMBER STATES - 2010	2018	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC-ECONOMICS OF AGRICULTURE DEPARTMENT

8	BIOMASS ESTIMATES	2017	YES	NO	NO	APPLICATION OF CALCULATIONS OVER PUBLICLY AVAILABLE DATA (MAINLY FAO PRODSTAT) AS FROM JRC/NOVA METHODOLOGY
9	BIOMASS USES AND FLOWS	2017	YES	YES	NO	APPLICATION OF CALCULATIONS OVER PUBLICLY AVAILABLE DATA (MAINLY FAO PRODSTAT)
10	COUNTRY DASHBOARDS	2019	NO	NO	YES	DATA COLLECTED FROM PUBLIC SOURCES (DEVCO, FAO, EIU, ITU, JRC, MIT, UNDP, UNESCO, UN, WHO, WB)
11	COVID-19 PANDEMIC DATA	2020	YES (ON GITHUB SITE)	YES	NO	DATA COLLECTED FROM NATIONAL BODIES
12	DG AGRI-JRC - PRODUCTION, TRADE AND APPARENT USE	2018	YES	YES	NO	APPLICATION OF DG AGRI EXPERTS' COEFFICIENTS AND DG AGRI/JRC METHODOLOGY TO COMBINE COMEXT DATA WITH DG AGRI DATA (SHORT TERM OUTLOOK)
13	FOODSECURE - FOOD AND NUTRITION SECURITY IN LONG TERM PERSPECTIVE	2017	NO	NO	YES	MODELLING (IMAGE, GLOBIOM, MAGNET)

14	FTA - FREE TRADE AGREEMENTS STUDY	2017	YES	NO	YES	MODELLING (AGLINK-COSIMO, MAGNET)
15	FOOD PRICE CROWDSOURCING AFRICA	2019	NO	YES	NO	DATA ON NIGERIAN FOOD PRICES COLLECTED ON THE MARKETS IN CROWDSOURCING VIA MOBILE/SMS
16	JOBS AND WEALTH IN THE EU BIOECONOMY / JRC - BIOECONOMICS	2017	YES	NO	YES	APPLICATION OF CALCULATIONS OVER PUBLICLY AVAILABLE DATA (MAINLY EUROSTAT COMEXT) AS FROM JRC METHODOLOGY BASED ON COEFFICIENTS PROVIDED BY NOVA INSTITUTE
17	JOBS CALCULATOR	2019	NO	NO	YES	ELABORATION OF THE SAM BY CGE TEAM AT JRC-ECONOMICS OF AGRICULTURE DEPARTMENT
18	MATRICE DE COMPTABILITÉ SOCIALE - KENYA - 2014	2018	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC-ECONOMICS OF AGRICULTURE DEPARTMENT
19	MATRICE DE COMPTABILITÉ SOCIALE - SÉNÉGAL - 2014	2018	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC-ECONOMICS OF AGRICULTURE DEPARTMENT

20	PLAN SÉNÉGAL EMERGENT 2019-2023	2019	YES	NO	YES	MODELLING (CGE)
21	RJOC - ETHIOPIA	2019	NO	NO	YES	MODELLING (CGE, DEMETER)
22	SAM - ETHIOPIA - 2015/16	2020	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC- ECONOMICS OF AGRICULTURE DEPARTMENT
23	SAM - KENYA - 2014 (EN/FR)	2018	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC- ECONOMICS OF AGRICULTURE DEPARTMENT
24	SAM - KENYA - 2017	2020	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC- ECONOMICS OF AGRICULTURE DEPARTMENT
25	SAM - SÉNÉGAL - 2014 (FR/EN)	2018	YES	YES	NO	ELABORATION OF THE SAM BY CGE TEAM AT JRC- ECONOMICS OF AGRICULTURE DEPARTMENT
26	SUPREMA - SUPPORT FOR POLICY RELEVANT MODELLING OF AGRICULTURE	2020	YES	NO	NO	MODELLING (AGMEMOD, CAPRI, GLOBIOM, IFM- CAP, MAGNET, MITERRA)
27	SCENAR 2030	2017	YES	YES	YES	MODELLING (CAPRI, IFM- CAP, MAGNET)
28	SURVEY: COVID EFFECTS ON AGRICULTURAL- FOOD SUPPLY CHAIN	2020	TBC	TBC	YES	DATA BEING COLLECTED WITH A SURVEY

Table 1 - DataM contents (July 2020)

A real-time future-updated situation of the open-data contents of DataM is accessible at this link of the EC data catalogue: <https://data.jrc.ec.europa.eu/collection/datam>

A real-time future-updated overview of DataM contents' usage is accessible at this link: https://datam.jrc.ec.europa.eu/datam/mashup/USAGE_STATISTICS

For those contents whose source is not modelling, data construction is based on ad-hoc techniques. In these cases, the typical technologies in use are: (i) the python language (data extraction from file or web sources, and data crunching); and (ii) database tools such as Oracle and SQL-lite (data crunching in SQL or PL/SQL).

DataM Information System is used for the final integration of data into consolidated datasets, and for the post-construction part. DataM can have also an impact in the construction phase since DataM datasets can work as input for further elaborations.

The DataM data management tool

With "data management tool" we mean a software layer above the data base management system that factorizes common needs for the management of (scientific) datasets.

The DataM data management tool allows:

- loading data from external sources.
- storing data in standard format (time-series oriented star-diagram with one unique measure and arbitrary dimensions; time and "indicator" are mandatory dimensions)
- avoiding "data manipulation" operations (Creating, modifying, dropping tables and indexes) / no need of DBA support
- managing dictionaries of common reference data
- allowing harmonization of data by:
 - mapping different nomenclatures
 - converting units of measurement
 - aligning different granularities
 - aligning different taxonomies
- Managing data versioning
- Managing standard meta data for all contents
 - Description
 - Contact point
 - Geographical coverage¹
 - Time coverage
 - Copyright
 - Update frequency
 - Domain
 - Keyword (tags for search operations in open data portals)
 - Contributors:
 - Name, Surname
 - Email
 - ORCID

¹ In DataM the time coverage and geographical coverage are automatically calculated on the basis of the data every night.

- Distributions:
 - Link for bulk download of raw data
 - Link for interactive data download
 - Link to interactive dashboard or infographic
- Related publications / methodology documents
 - Title
 - Authors
 - DOI

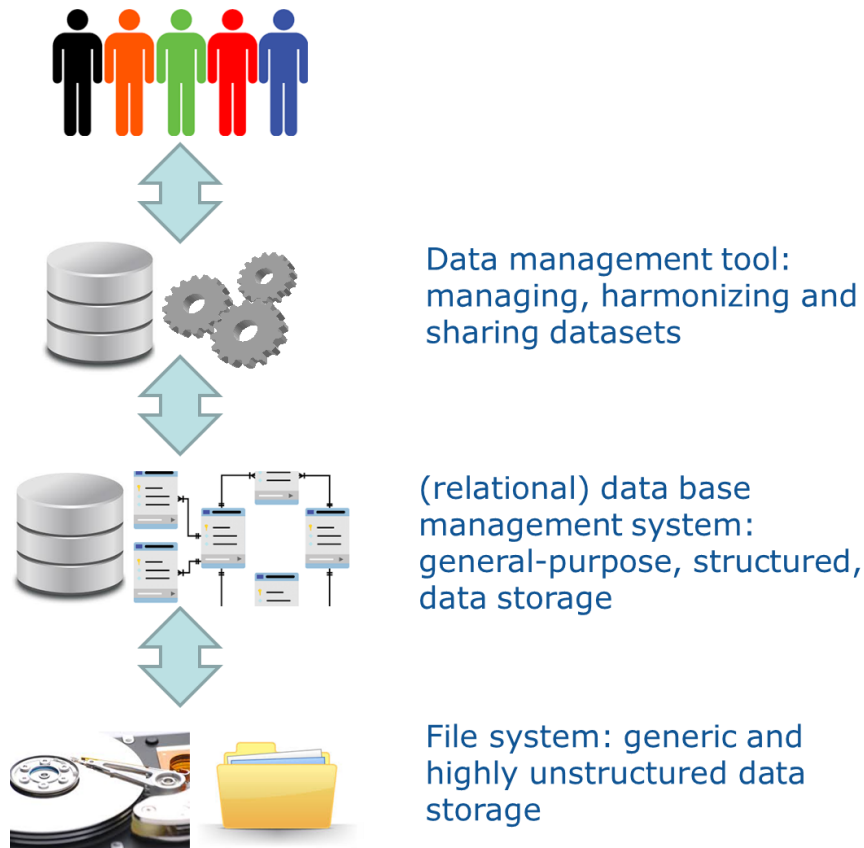


Figure 4 - Data management tool

The data management tool of DataM is a software entirely developed at JRC oriented to reference data dictionaries and star-diagram data structures (principles that are common to most data warehouse technologies), while integrating the advanced management of standard meta-data.

The current data management tool of DataM, at time of writing, does not include yet features for:

- Integrating in the system the Python routines used for extracting data from arbitrary sources and transforming it in a structure suitable for the load in the database.
- Workflow management for the collaborative construction of datasets by distributed working groups.
- Management of big data
- Management of geo-spatial data
- Management of multilingualism

The DataM BI/analytics platform: Qlik-Sense

Qlik-Sense is a modern software for business intelligence (BI). It is considered one of the top leaders on the BI market.² In use in many Directorate General, is a standard de-facto in the European Commission. JRC was a pioneer in the European Commission for the adoption of this software through the experience of the Economics of Agriculture department and DataM. Whereas at beginning of SUPREMA (2018) Qlik was a product only in use by the EoA Unit, at time of writing (2020) this has become a standard service at JRC.

Business intelligence (BI) comprises the strategies and technologies used by enterprises for the data analysis of business information. BI technologies provide historical, current and predictive views of business operations. Common functions of business intelligence technologies include reporting, online analytical processing, analytics, data mining, process mining, complex event processing, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics.

JRC is adopting the BI for scientific purposes, in particular, in DataM BI is used for:

- Analysing the constructed datasets:
 - Comparison with other sources
 - Outliers detection
 - Check of the respect of "business rules"
- Exchanging results in proper way within the scientific circuit or with other stake-holders (i.e.: policy DG's)
- Disseminating results on the web allowing user to make self-analysis of data:
 - interactive dashboards = dashboards with interrelated charts/tables/maps with scarce or absent narrative
 - Interactive infographics = web pages with much narrative and embedded Qlik contents.

The DataM portal

The DataM portal (<https://datam.jrc.ec.europa.eu>) is an official part of the European Commission web presence³, and official "knowledge service" (https://ec.europa.eu/info/research-and-innovation/knowledge-publications-tools-and-data/knowledge-centres-and-data-portals_en); it has been approved in 2018 by DG COMM when was refactored following EC web standards and is inspired to the principles of the knowledge-for-policy platform⁴ being developed by JRC. It offers the following functionalities:

- Search contents by specifying keywords
- Download of raw data:
 - Bulk download, zip file with CSV with data, citation text, legal notice and all metadata
 - Interactive data download by filtering the parts of interest of the dataset
 - Obtaining automatically the citation text, for correctly citing the dataset in scientific publications
 - Visualizing the meta-data

² Source: Gartner – BI Magic quadrant

³ http://ec.europa.eu/info/about-commissions-new-web-presence_en

⁴ <https://ec.europa.eu/knowledge4policy>

- Visualizing the copyright notice
- Accessing the related publications / methodology documents
- Accessing the Qlik interactive dashboards and infographics
- API for automatic synchronization of the meta-data with the JRC data catalogue (and subsequent dissemination to EU open data portal and European Data Portal).
- API for transmitting data on-demand to other computer systems

DataM governance

At time of writing, DataM is developed and managed by a team of four IT professionals / data scientists working under the JRC department for Economics of Agriculture within the directorate for Sustainable Resources. It is powered by JRC IT, in particular by the staff (and on machines) of the JRC Seville site for the database (Oracle) and web server, and of JRC Ispra (Italy) for the business analytics platform Qlik-Sense. In long term, the complete migration to Ispra is a rather possible scenario.

Architecture of the data base

The data base of DataM underlying the data management tool is implemented in Oracle. Each dataset is basically implemented in a fact table and a number of specific tables for the reference.

The strength of this model is the plasticity: it suits to host almost all typical datasets of our domain with a simple, common, structure.

Evolution of DataM during the SUPREMA project (2018-2020)

Activities on DataM in 2018 have been focusing on:

- Implementation of a Qlik-based system for the data quality check of model outcomes. An important application was produced for the CAPRI outcomes
- Implementation of a data connector for the direct integration of GAMS and Qlik (G2Q), that was completed and released as open source on last April <https://datam.jrc.ec.europa.eu/datam/perm/news/789>
- Refactoring of the data management tool (project completed in 2019). A further phase for the integration of python scripts is envisaged in 2021.

Data privacy

DataM is mainly thought for public dissemination of open data, which means that most of functionalities do not require a login. However, by logging in, authorised users can access restricted contents that require specific access rights. Restrictions normally apply to contents still under study, and this has been the case for the SUPREMA project: contents under preparation were restricted within the SUPREMA community by accessing the restricted area of DataM, before the official publication.

The username and password are obtained through the European Commission's user authentication service (EU Login).

For other users, an EU Login account can be obtained through a simple registration procedure:

- Click on "Log In" on the top-right corner of the screen
- Click on "Create an account" and follow the instructions.

3 Open access release of scenarios

3.1 Uploading data to DataM

Data files are prepared by the modelling teams, coordinated by the task leaders and send to JRC for upload. Every single data point is a numerical value, typically identified by a number to represent a dimension:

- The model that produced it
- The scenario (or simulation)
- The spatial unit (or region)
- The time unit, typically a year to which the projection refers to
- The indicator (name of some variable, normally endogenous)
- The item to which the indicator refers)
- Other dimensions could be added, if relevant to the model. For example, a farm-oriented model could add dimensions like farm size or farm type.

A baseline comparison and harmonization action is addressed for all models represented in SUPREMA (Task 3.1 – Inter-model baseline comparison and harmonization). In addition, a medium-term (until 2030) assessment of European agricultural policy alternatives will cover CAPRI, IFM-CAP, and AGMEMOD-MITERRA Europe (Task 3.2 – Using SUPREMA for a medium-term assessment of European agricultural policy alternatives). Finally SUPREMA will also use the different modelling tools for the long-term (until 2050) assessment of climate change goals (Task 3.3), using GLOBIOM and MAGNET as leading models. All scenarios are released open access soon afterwards.

3.2 Uploading data to DataM

Format of the data

A format is provided with guidelines to present data, distinguishing between.

Model data must be provided as a table (=file), in which every line is one data point (see below).

	A	B	C	D	E	F	G	H
1	MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE	
2	CAPRI	SCEN1	IT01	2030	PRODUCTION	WEATH	30	
3	MAGNET	SCEN2	ES02	2040	IMPORT	RICE	40	
4								
5								

In case you use a CSV, you can also see in a text editor the physical structure (see below).

```

2020_06_25_ModelDataExample - Notepad
File Edit Format View Help
MODEL,SCENARIO,REGION,YEAR,INDICATOR,ITEM,VALUE
CAPRI,SCEN1,IT01,2030,PRODUCTION,WEATH,30
MAGNET,SCEN2,ES02,2040,IMPORT,RICE,40
  
```

Rules to provide data

F4								
	A	B	C	D	E	F	G	H
1	MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE	
2	CAPRI	SCEN1	IT01	2030	PRODUCTION	WEATH	30	
3	MAGNET	SCEN2	ES02	2040	IMPORT	RICE	40	
4								
5								

A header is mandatory at the first line (see below).

	A	B	C	D	E	F	G
1	MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE

It is recommended to use codes (i.e. IT01 and not "Abruzzi"). It is forbidden to have two lines that refer to the same combination of dimensions (see below).

	A	B	C	D	E	F	G	H
1	MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE	
2	CAPRI	SCEN1	IT01	2030	PRODUCTION	WEATH	30	
3	MAGNET	SCEN2	ES02	2040	IMPORT	RICE	40	
4	MAGNET	SCEN2	ES02	2040	IMPORT	RICE	50.3	??

Dimensions to provide model data

The facts table is the most important one, and a single dimension is specified for each column (see below).

	A	B	C	D	E	F	G
1	MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE

For the benefit of the reader, additional information may be provided in separate tables (see below):

	A	B
1	REGION	NAME
2	IT01	Abruzzi

DIMENSION TABLE OF REGIONS

Unit of measurement to provide model data

The unit of measurement could be added to the fact table. In case the unit of measurement is not included in the fact table, you need to specify it in a dimension table of the indicators. The latter is possible only if one indicator is uniquely linked to a single unit of measurement (see below).

	A	B
1	INDICATOR	UOM
2	PRODUCTION	TONS
3	IMPORT	M EUR
4		

In case a single indicator could be depicted in different units, you can either indicate in the table (see below).

MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE	UOM
CAPRI	SCEN1	IT01	2030	PRODUCTION	WEATH	30	TONS
MAGNET	SCEN2	ES02	2040	IMPORT	RICE	40	M EUR
MAGNET	SCEN2	NL04	2040	IMPORT	RICE	50.3	TONS

Or split the indicator accordingly (then you have again the option to indicate the unit of measurement in fact or dimension table) (see 2 tables below).

MODEL	SCENARIO	REGION	YEAR	INDICATOR	ITEM	VALUE	UOM
CAPRI	SCEN1	IT01	2030	PRODUCTION	WEATH	30	TONS
MAGNET	SCEN2	ES02	2040	IMPORT (VALUE)	RICE	40	M EUR
MAGNET	SCEN2	NL04	2040	IMPORT (QUANTITY)	RICE	50.3	TONS

INDICATOR	UOM
PRODUCTION	TONS
IMPORT (VALUE)	M EUR
IMPORT (QUANTITY)	TONS

3.3 Comparison of the 2030 and 2050 baselines

A selection of commodities is made to compare the simulation results of the 2030 baselines, including soft wheat, corn, rapeseed, rapeseed oil, beef, pork and raw milk from cow. These commodities are among the most straightforward in their definition. The basis for comparison of the baseline results of AGMEMOD and IFM-CAP is the baseline results of CAPRI. The reason for choosing CAPRI over the other two models is that IFM-CAP uses growth rates of prices and yields projected by CAPRI to generate its baseline. Thus, comparing the two models provides with additional insights on the outcome of alignment of these two models.

GLOBIOM, MAGNET and CAPRI are included in the long-term 2050 baseline, because they are capable of producing long-term projections. Indicators focus on greenhouse gas emissions and agricultural land use. Crop production and market balances are included as well. The commodities considered include rice, wheat, oilseeds, non-ruminant and ruminant meat products, as well as aggregates, such as agriculture, crops and livestock. The regional aggregation considered is EU before 2019.

3.4 Comparing a 2030 CAP scenario

A selection of commodities is made to compare CAPRI results with AGMEMOD-MITERRA. Variables include:

- Production at EU level (% deviation from baseline in 2030) – beef, dairy and sugar.
- Net trade at EU level (% deviation from baseline in 2030) – beef, dairy and sugar.
- Consumption per capita EU28 (% deviation from baseline in 2030) – beef, dairy and sugar.
- Average price EU28 (% deviation from baseline in 2030) – beef, dairy and sugar.

- Additional indicators EU28 (% deviation from baseline in 2030) – including utilized agricultural area, cereals, oilseeds, pulses, potatoes, sugar beet and vegetables and permanent crops).
- Emissions EU28 (% deviation from baseline in 2030) – CH₄ emissions, N₂O emissions, greenhouse gas emissions, NH₃ emissions, N leaching.

3.5 Comparing a 2050 climate scenario

This scenario does assess the development of long-term global climate change mitigation scenario, and include the greenhouse mitigation potential of the EU's agricultural sector and related environmental, economic, socio-economic impacts conditional on different levels of GHG mitigation efforts outside the EU, using GLOBIOM, MAGNET and CAPRI. The following variables are compared:

- Change in agricultural GHG emissions in MtCO₂eq/yr in the 00% buy in scenario in 2050 compared to the reference scenario.
- Relative change in ruminant production and consumption/domestic use in the 00% buy in scenario in 2050 compared to the reference scenario across world regions.
- Relative change in EU agricultural production and consumption/domestic use in the 00% buy in scenario in 2050 compared to the reference scenario.
- Impact of different levels of rest-of-the-world mitigation ambition on EU agricultural sector in 2050 across models.
- Relative change in EU agricultural production in the 50% buy in scenario and full-buy-in scenario in 2050 compared to the reference scenario.
- Global GHG mitigation potential from agriculture in MtCO₂eq/yr in 2050 across models.
- Relative change in livestock production in the 10% and 50% buy in scenario in 2050 compared to the reference scenario across world regions.
- Relative change in livestock calorie intake in the full buy-in scenario in 2050 compared to the reference scenario across world regions.

3.6 Intellectual Property Rights (IPR)

Intellectual Property Rights (IPR) will receive special attention from the beginning. All rules regarding management of knowledge and IPR will be governed by the Consortium Agreement (CA). SUPREMA was based on DESCA (Consortium Agreement Model) H2020 model for the Consortium Agreement (CA). SUPREMA will adhere to the rules laid down in Annex II of the Grant Agreement. The CA will address background and foreground knowledge, ownership, protected third party components of the products, and protection, use and dissemination of results and access rights.

The following principles will be applied:

- Pre-existing know how: Each Contractor is and remains the sole owner of its IPR over its pre-existing know-how. The Contractors will identify and list the pre-existing know-how over which they may grant access rights for the project. The Contractors agree that the access rights to the pre-existing know-how needed for carrying out their own work under the project shall be granted on a royalty-free basis.
- Ownership and protection of knowledge: The ownership of the knowledge developed within the project will be governed by an open source license.
- Open data: Data and results obtained during the project that are based on open public-sector data will be made available free of charge.

The procedures for the dissemination, protection and exploitation of intellectual property rights (IPR) are clearly covered in the Consortium Agreement (in Section 6: Governance Structure, Sub-section 6.2.4: Veto rights). The intention has been to balance the requirements necessary to protect such

intellectual property and the foreseen dissemination objectives. IPR will be applied according to the rules of the employer under the applicable European and national laws and regulations.

4 Open access release of data

DataM is the European Commission's data portal of agro-economics modelling (<https://datam.jrc.ec.europa.eu>), containing the outcomes of research activities, and is operated by the Joint Research Centre (JRC) of the European Commission. A news item is posted on 22 July 2020 <https://datam.jrc.ec.europa.eu/datam/perm/news/849> announcing that data files are accessible in DataM and catalogued as open-data at this permanent URL: <https://data.jrc.ec.europa.eu/dataset/d6ef74c6-ba91-4e37-827e-d0854fbe85dd>

Related resources

Data access



SUPREMA - Agricultural policy scenario

Dataset - Interactive download - CSV format



SUPREMA - Agricultural policy scenario

Dataset - bulk download - zip file with CSV inside



SUPREMA - Baseline scenario

Dataset - Interactive download - CSV format



SUPREMA - Baseline scenario

Dataset - bulk download - zip file with CSV inside



SUPREMA - Climate change mitigation scenario

Dataset - Interactive download - CSV format



SUPREMA - Climate change mitigation scenario

Dataset - bulk download - zip file with CSV inside

Publications



SUPREMA - Deliverable 3.1 - Inter-model baseline harmonization and comparison



SUPREMA - Deliverable 3.2 - Agricultural policy scenario description and divergence analysis



SUPREMA - Deliverable 3.3 - Analysis of climate change mitigation scenarios

Additional information

Last modified	2020-07-15
Issue date	2020-07-15
Landing page	https://datam.jrc.ec.europa.eu
Geographic area	World
Temporal coverage	From: 2000-01-01 – To: 2050-12-31
Update frequency	unknown
Language	English
Data theme(s)	Agriculture, fisheries, forestry and food; Environment
EuroVoc domain(s)	52 ENVIRONMENT; 60 AGRI-FOODSTUFFS
Identifier	http://data.europa.eu/89h/d6ef74c6-ba91-4e37-827e-d0854fbe85dd

The open dataset

The [dataset](#) includes 3 components.

1 - Baseline scenario

Links: [bulk download](#), [query](#), underlying [report](#).

Models: AGMEMOD, CAPRI, GLOBIOM, IFM-CAP, MITERRA

These are the harmonized baselines for 2030 and 2050.

Please note that the baseline projections do not take into account the 2020 and possible future effects of the SARS-CoV-2 pandemic

2 - Agricultural policy scenario

Links: [bulk download](#), [query](#), underlying [report](#).

Models: AGMEMOD, CAPRI, MITERRA

Medium-term horizon scenarios aiming comparing different models and/or model combinations, that have a large degree of 'similarity' such as joined indicator variables, i.e.: AGMEMOD-MITERRA (combined) modelling tool and the CAPRI model. The main focus was comparing model results in both agronomic and biophysical domains. Two variants of the agricultural policy scenario have been simulated and compared: (i) a CAP greening scenario; and (ii) a sustainable diet scenario. Both scenarios are hypothetical but have been chosen in such a way that they can provide insights in future policy issues as: (i) a further greening of the CAP fits in the policy implementation space as it is included in the ongoing policy reform of the CAP after 2020; and (ii) as increasing consumer awareness about healthy diets and their relation to meat consumption, as well as the footprint/climate consequences are highly relevant with respect to the Green Deal roadmap (December 2019) and the Farm to Fork Strategy (May 2020) documents that have been recently published.

3 - Climate change mitigation scenario

Links: [bulk download](#), [query](#), underlying [report](#).

Models: CAPRI, GLOBIOM, MAGNET

This scenario quantifies the GHG mitigation potential of the EU's agricultural sector and domestic and global impacts of the EU policy, conditional on different levels of GHG mitigation efforts in the rest of the world. These are obtained through the SUPREMA models CAPRI, GLOBIOM and MAGNET and include scenarios where the EU only takes ambitious unilateral climate action up to scenario where the 1.5 C target is pursued globally

The model runs presented in WP3 are accessible through:

<https://data.jrc.ec.europa.eu/dataset/d6ef74c6-ba91-4e37-827e-d0854fbe85dd>



Organisation: European Commission, Joint Research Centre

Point of contact: ✉ jrc-datam@ec.europa.eu

Title: SUPREMA - Support for Policy RElevant Modelling of Agriculture

D ATAM

Description

Impact assessments for agriculture are partly based on projections delivered by models. Sectoral policies are becoming more and more interrelated. Hence, there is a need to improve the capacity of current models, connect them or redesign them to deliver on an increasing variety of policy objectives, and to explore future directions for agricultural modelling in Europe.

SUPREMA (Support for Policy RElevant Modelling of Agriculture) is a project that has received funding from the European Union's Horizon 2020 research and innovation programme (under grant agreement No 773499 SUPREMA) and that came to address this challenge by proposing a meta-platform that supports modelling groups linked already through various other platforms and networks.

SUPREMA should help close the gaps between expectations of policy makers and the actual capacity of models to deliver relevant policy analysis. The SUPREMA model family includes a set of 'core models' that are already used in support of key ... [Show more](#)

Related resources

Data access



SUPREMA - Agricultural policy scenario

Dataset - Interactive download - CSV format



SUPREMA - Agricultural policy scenario

Dataset - bulk download - zip file with CSV inside



SUPREMA - Baseline scenario

Dataset - Interactive download - CSV format



SUPREMA - Baseline scenario

Dataset - bulk download - zip file with CSV inside



SUPREMA - Climate change mitigation scenario

Dataset - Interactive download - CSV format



SUPREMA - Climate change mitigation scenario

Dataset - bulk download - zip file with CSV inside

Publications



SUPREMA - Deliverable 3.1 - Inter-model baseline harmonization and comparison



SUPREMA - Deliverable 3.2 - Agricultural policy scenario description and divergence analysis



SUPREMA - Deliverable 3.3 - Analysis of climate change mitigation scenarios

SUPREMA Horizon 2020 agro-environmental project. Model outcomes are online.

The [EC Data portal of agro-economic Modelling \(DataM\)](#) publishes today [baselines and model outcomes](#) of the Horizon 2020 project "SUPREMA".

The project

Impact assessments for agriculture are partly based on projections delivered by models. Sectoral policies are becoming more and more interrelated. Hence, there is a need to improve the capacity of current models, connect them or redesign them to deliver on an increasing variety of policy objectives, and to explore future directions for agricultural modelling in Europe.

SUPREMA (SUpport for Policy RElevant Modelling of Agriculture) is a project that has received funding from the European Union's Horizon 2020 research and innovation programme (under grant agreement No 773499 SUPREMA) and that came to address this challenge by proposing a meta-platform that supports modelling groups linked already through various other platforms and networks.

SUPREMA should help close the gaps between expectations of policy makers and the actual capacity of models to deliver relevant policy analysis.

The SUPREMA model family includes a set of 'core models' that are already used in support of key European impact assessments in agriculture, trade, climate and bioenergy policies.

SUPREMA has been coordinated by [Wageningen Research](#) with the participation of [EuroCARE](#), [Thünen Institute](#), [Swedish University of Agricultural Sciences \(SLU\)](#), European Commission [Joint Research Centre \(JRC\)](#) and [Research Executive Agency \(REA\)](#), [International Institute for Applied Systems Analysis \(IIASA\)](#) and [Universidad Politécnica de Madrid \(UPM\)](#).

5 Conclusions

DataM is the European Commission's data portal of agro-economic modelling. It contains the outcomes of research activities, and is operated by the Joint Research Centre (JRC) of the European Commission. DataM, including the web portal but also the Information System, will be used to release model runs that are considered for open access release. A baseline comparison and harmonization action is addressed for all models represented in SUPREMA. A medium-term (until 2030) assessment of European agricultural policy alternatives covers CAPRI, IFM-CAP and AGMEMOD-MITERRA Europe. Finally, SUPREMA does also use different modelling tools for the long-term (until 2050) assessment of climate change goals, using GLOBIOM and MAGNET as leading models.

6 References

European Commission (2017). Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in H2020. European Commission, Directorate-General for Research & Innovation.

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf