D4.7: FUTURE GOVERNANCE OF THE MODELS

LEAD AUTHOR: Torbjörn Jansson

OTHER AUTHORS: Gustaf Eskhult, Hyungsik Choi, Pavel Ciaian, Petra Salamon, Martin Banse, Petr Havlik, Hans van Meijl, Jan Peter Lesschen.

DATE: (3 – July – 2020)

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.4.7 Future governance of the models
WP NAME/WP NUMBER	Project coordination and management / WP4
TASK	4.5
VERSION	1.0
DISSEMINATION LEVEL	Public
DATE	14/07/2020 (Date of this version) – 29/06/2020 (Due date)
LEAD BENEFICIARY	SLU
RESPONSIBLE AUTHOR	Torbjörn Jansson
AUTHOR(S)	Torbjörn Jansson (SLU), Gustaf Eskhult (SLU), Hyungsik Choi (SLU), Pavel Ciaian (JRC), Petra Salamon (TI), Martin Banse (TI), Petr Havlik (IIASA), Hans van Meijl (WUR), Jan Peter Lesschen (WUR).
INTERNAL REVIEWER	Approval by WP leader

DOCUMENT HISTORY

Version	Initials/NAME	DATE	COMMENTS-DESCRIPTION OF ACTIONS
1			First draft
2	TJ	2020-06-26	Extended with review of other models, publications and deepened analyses
3 (final)	TJ	2020-07-03	Revised after review and feedback from the participating modelling teams

Table of Contents

TABLES

FIGURES

Executive summary

The purpose of this deliverable is to analyse the structure and governance of the six models involved in SUPREMA, with the aim to derive recommendations on how to move towards a more sustainable business model in case they are lagging behind. The analysis is based on self-assessment by the teams, augmented with a publication metric and other publicly available information.

Sustainable simulation models in agriculture are discussed based on a review of the survival rate over 20 years of sixteen models reviewed in 1999. It is found that six of those are still actively developed, while ten are dormant or no longer existing. Without any further data on how those models were managed in detail, it is difficult to find explanations to why certain models were discontinued. A hypothesis is that changes in key human resources (staff moving) is an important factor, but some models seem to have survived despite changes in the core modelling teams.

The models in SUPREMA are categorized along the two dimension of how open the networks are in terms of code access and contributions of code. Ideally, this information would be supported by metrics on e.g. commits or pull requests in versioning software. Albeit all models use some kind of software versioning system, those system look differently and are generally not subject to open access. Therefore, the categorization is ultimately based on subjective judgement, supported by descriptions of the model networks developed in collaboration with the modelling teams and data on the number of institutions and persons involved in each model network. The classification is used to generalize the discussion of challenges faced by model development teams.

Based on the descriptions and analytical model, we attempt to identify the key challenges for each modelling network and some broad recommendations that the different networks might want to consider for their future work. Since the author is a member of the CAPRI development network, there is a certain bias in that direction in terms of detail. In summary, the SUPREMA models face different challenges, but all face the partly conflicting needs to, on the one hand, open the development networks to find a wider base of expertise and contributions, and on the other hand to coordinate developments and share overhead costs among partners. CAPRI and AGMEMOD have more open and wider developer networks compared to MAGNET and GLOBIOM. The key challenge for those models is to find ways to coordinate overhead processes and share costs. For CAPRI, there is also a need for more coordination in the software development process itself, and one possibility to explore is the establishment of a new legal body that institutionalizes the governance of the model. For MAGNET and GLOBIOM, the challenges are rather to open or extend the networks for development contributions. IFM-CAP is the youngest model in the family and might benefit from finding a wider user base. This also holds true for MITERRA, which is comparatively specialized and currently relying on the expertise of a small group of developers.

Changes with respect to the DoA

No changes with respect to the DoA

Dissemination and uptake

This report is intended for the modelling teams involved in the SUPREMA family of models. It might also be relevant for other stakeholders in these models such as the connected research institutes and the European Commission.

Short Summary of results (<250 words)

The SUPREMA models face different challenges, but all face the partly conflicting needs to, on the one hand, open the development networks to find a wider base of expertise and contributions, and on the other hand to coordinate developments and share overhead costs among partners. CAPRI and AGMEMOD have more open and wider developer networks compared to MAGNET and GLOBIOM. The key challenge for those models is to find ways to coordinate overhead processes and share costs. For CAPRI, there is also a need for more coordination in the software development process itself, and one possibility to explore is the establishment of a new legal body that institutionalizes the governance of the model. For MAGNET and GLOBIOM, the challenges are rather to open or extend the networks for development contributions. IFM-CAP is the youngest model in the family and might benefit from finding a wider user base. This also holds true for MITERRA, which is comparatively specialized and currently relying on the expertise of a small group of developers.

Evidence of accomplishment

This report in itself is the evidence of accomplishment.

GLOSSARY / ACRONYMS

AGMEMOD	AGRICULTURAL MEMBER STATE MODELLING
CA	CONSORTIUM AGREEMENT
CAPRI	COMMON AGRICUTLURAL POLICY REGIONALISED IMPACT MODEL
CAPRI-RD	COMMON AGRICULTURAL POLICY REGIONALISED IMPACT - THE RURAL DEVELOPMENT DIMENSION
CCAT	CROSS-COMPLIANCE ASSESSMENT TOOL
COMEXT	EUROSTAT INTRA- AND EXTRA-EUROPEAN TRADE DATABASE
DG CLIMA	THE COMMISSION'S DIRECTORATE-GENERAL FOR CLIMATE ACTION
EUROCARE	EUROPEAN CENTRE FOR AGRICULTURAL, REGIONAL AND ENVIRONMENTAL POLICY RESEARCH
FADN	FARM ACCOUNTANCY DATA NETWORK
FLOSS	FREE/LIBRE OPEN SOURCE SOFTWARE
GLOBIOM	GLOBAL BIOSPHERE MANAGEMENT MODEL
GTAP	GLOBAL TRADE ANALYSIS PROJECT
IFM-CAP	INDIVIDUAL FARM MODEL FOR COMMON AGRICULTURAL POLICY ANALYSIS
IIASA	INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
INPE	INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS
JRC-IPTS	JOINT RESEARCH CENTRE AND THE INSTITUTE FOR PROSPECTIVE TECHNOLOGICAL STUDIES OF THE EUROPEAN COMMISSION
MAGNET	MODULAR APPLIED GENERAL EQUILIBRIUM TOOL
MOU	MEMORANDUM OF UNDERSTANDING
OA	OPEN ACCESS
SLU	SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES
TERM	EXPLANATION / MEANING
TI	THÜNEN-INSTITUT
WECR	WAGENINGEN ECONOMIC RESEARCH
VSN	SOFTWARE VERSIONING SYSTEM
WUR	WAGENINGEN UNIVERSITY & RESEARCH

1 Introduction

Developing a new large-scale simulation model is a very costly exercise, typically involving several person-years of work to reach the production phase. At the same time, most models are developed with a particular type of analysis in mind. As the political and economic surroundings change, so do the requirements on the models. Furthermore, models require continuous updates in terms of data in order to stay relevant. Therefore, models require continuous development and maintenance work after the initial investment in order to continue to be relevant to empirical questions.

If we look at modelling in the European Union and the needs of the policy makers, which has been a driver for much modelling work, we see several shifts in focus from 1990 until today. The common agricultural policy (CAP) has gradually shifted from market price support, via coupled direct support, to de-coupled farm payments. Environmental measures, both in terms of subsidies and regulations have gained increased attention. And with the last and the present reforms, the focus has returned to the national markets via voluntary coupled payments, and national strategic plans. Furthermore, member states have shown growing concern for national competitiveness and food security issues. Obviously, a single agricultural economic model cannot cope with all of these shifts without adaptation and development

Some models cope well with changes and are continuously being applied and developed over longer time periods. Other go dormant and eventually disappear. Why is that so, and how should modelling work be organized in order to produce a “sustainable business model”? Clearly, there are many ways to skin a cat, and each model is in some way unique. Nevertheless, this deliverable attempts to analyse what sustainability in modelling is and to derive recommendations for the models of the SUPREMA family. There are six models, of which four are more widely used and with a broader scope of application (CAPRI, MAGNET, AGMEMOD, GLOBIOM), whereas the two remaining models are more narrowly focussed and less used (IFM-CAP, MITERRA).

Before looking at the SUPREMA models, we have a look at models that were surveyed twenty years ago. Some of these models are still active, while others are not, and this in itself is a workable definition of “sustainable management”. Then, we turn to the SUPREMA models, and analyse the models from three different perspectives: how widely are they used, what is their scope and what are the characteristics of the development networks? The final section summarizes the conclusions for each model in concise recommendations for changes, if any changes seem necessary.

Before proceeding, it should be noted that all conclusions and recommendations of this deliverable are the responsibility of the main author. The author list of this deliverable contains the names of at least one person per modelling team, which all have contributed to varying degrees to the production or served as main point of contact, without attempting to exhaust the list of persons providing input in some form or the other. This does not necessarily imply that the people in the list endorse all conclusions and recommendations reached.

2 Methods

The purpose of this deliverable is to

1. analyse the governance of the models of the SUPREMA family from the perspective of whether they have a “sustainable business model”, and to
2. derive recommendations for models “lagging behind in this respect”.

David Hume (1711-1776) noted that what *ought* to be cannot be derived solely from statements about what *is*. This is true in this deliverable too, where the objective is to derive recommendations (what ought to be) from an analysis of what is (descriptions of the modelling networks in SUPREMA). The step from “what is” to ”what ought to be” involves an element of preference, and depends on traits of the author involved. Therefore it is important to point out that this analysis was carried out by Torbjörn Jansson, with 20 years of experience from the CAPRI developers’ network as a member of various research institutions, but little experience of other software development.

There are related research fields termed “Software Engineering Governance”, “Software Development Governance” and “IT Governance”. Studies of governance are frequently investigating the structure of Free/Libre Open Source Software (FLOSS), because there the exchange of information among developers is easily available and measurable. Via messages in chat forums, bug reports, code comments and so-called pull requests (requests to include particular code into a code repository) and commits (the actual inclusion of code into a code repository) metrics can be computed and network charts drawn. For commercial or proprietary software, that kind of information is more difficult to obtain. The models of SUPREMA float in the lands between FLOSS and proprietary software, with a tendency towards the proprietary side. None of the models has made their versioning software repository publicly readable, albeit some publish versions of the source code.

Within FLOSS (Free/Libre Open Source Software) that there are two sorts of studies regarding how FLOSS cooperatives organises themselves. One focuses solely on one or a few open source projects to judge how they are built up. The other one focuses more broadly and uses existing open-source repositories for data mining in order to draw inference. This has been used to draw inference on what factors contributes to open-source development.

Few studies looks at a specific open-source project. Zhou and Mockus (2015) takes the approach of focusing solely on two open-source projects, Mozilla and Gnome, by studying the Issue Tracking System (ITS) data. To understand what increases the odds of modelling participants to become long term contributors, logistic regression has been used. Main conclusions is that the responsiveness to the contributor from the environment she is working in affects the odds of becoming a long-term contributor.

Crowston and Howison (2005) have taken a wider approach and examined the typical social structure of open-source development in debugging processes through Sourceforge, and concluded that there are no clear patterns pointing towards highly centralised or de centralised network. However they have identified a pattern of decentralisation in projects with large number of contributors, and concluded that very few projects in the open-source environment contains high level of participants and a high degree of centralisation. Beecher *et al* (2009) studies whether there are different process and product characteristics depending on the repository used by the project. Their conclusions are to some extent that the success of an open-source project could very well depend on the repository they are linked to. Medappa and Srivastava (2019) uses Github data to show whether superposition could influence the successfulness of an open-source project. Their findings were that sequentially layered and individual task work, known as superposed organization of work, can enhance the popularity of a project in contrast to the common software development. A main finding is that the repository matters to the success of the project. Communication is important to engage people into long-term contribution and how the development is organised matters. Open-source development tends to be clustered and decentralised.

Our approach in this deliverable is rather exploratory, largely based on self-assessment by the model teams involved and complemented with public documentation of the models and other information. In an initial stage (2018), semi-structured interviews were carried out with representatives of the four “core models” (CAPRI, AGMEMOD, GLOBIOM, MAGNET). This resulted in an initial appreciation of how the networks were organized and which teams expressed concerns with organizational issues. Based on the interviews and publicly available information, we reached some initial conclusions that we discussed bilaterally with representatives of the modelling teams. At this stage, we also extended the scope of the analysis to include MITERRA and IFM-CAP. We also added a bibliographic metric to the analysis, and the models were classified along two dimensions (scope of the model, spread of the network) in order to be able to generalize the discussion. The analysis and recommendations were presented to the teams in a project meeting on June 25, and the analysis was somewhat revised and updated after feedback from the teams.

3 Sustainably managed models

A review by van Tongeren, van Meijl and Surry (2001) on global models applied to agriculture investigated 16 different models that were actively developed and used in 1999. Six of those, listed in Table 1, seem to be actively developed and applied to this day, whereas ten are dormant or have disappeared. One way of defining sustainability *ex-post* is to use model survival as a criterion. Viewed that way, models in Table 1 were sustainably managed and the ones in Table 2 were not. Two of the models, AgLink and the FAO World Model, were essentially merged into one, AgLink-COSIMO, which is now jointly developed by OECD and FAO. FAPRI, GTAP and INFORUM are similar to the extent that they are maintained by research institutes tied to American universities (Missouri, Purdue and Maryland). Among those, the GTAP arguably stands out as the most successful model, backed by a consortium of strong public partners such as OECD, FAO, EC (JRC), World bank, IFPRI, Wageningen Economic Research, Thünen Institute, USDA, but also consultancy companies like McKinsey and KPMG. G-Cubed is still developed and applied by one of the original developers, and is now maintained by a dedicated commercial software/consultancy firm (McKibbin Software Group) in cooperation with the University of Maryland.

Table 1: Models still active

Model	Developer
AGLINK	OECD
FAO World Model (now: COSIMO)	FAO (OECD)
FAPRI (FAPRI-MU)	FAPRI, Uni Missouri
G-Cubed	McKibbin and Wilcoxen (McKibbin Software Group)
GTAP	GTAP Center and Consortium
INFORUM	The Interindustry Forcasting Project at the University of Maryland

It may be interesting to have a closer look at The Global Trade Analyses Project (GTAP). GTAP is a global consortium in quantitative economic analysis of pressing global concern in the areas of Trade and Development and Global Environmental Issues. It consist of 33 members including OECD, FAO, EC (JRC), World bank, IFPRI, Wageningen Economic Research, TI, USDA, but also consultancy companies like McKinsey and KPMG. The core of its success is an institutional innovation in economic modelling through international collaboration to increase quality of data and analysis. The idea is to cover fixed costs together, to create a public good with lower entry barriers in and to better serve the policy analysts and decision makers. GTAP is truly a global network with users in almost every country in the world. The centrepiece of the Global Trade Analysis Project is a global database describing bilateral trade patterns, production, consumption and intermediate use of commodities and services. The current GTAP Data Base may be purchased by anyone who is interested in using it. The proceeds from the model help to offset the cost of producing the next release. This permits users to share development costs and it prevents needless duplication of effort in creating this public good. Older versions are available free of charge (except for the preceding version). The GTAP Data Base is fully documented and produced by the GTAP Centre and the quality is enhanced by all consortium members and data contributors.

The Swedish Agricultural Sector Model (SASM) may serve as an example of a small system that was successfully maintained by a single modeller for more almost 30 years. Initially developed at SLU (Apland and Jonasson, 1992), it has subsequently been applied on a consultancy basis by the main developer (Lars Jonasson) for clients within the public administration as well as the private sector. However, there is no recruitment of new modellers and essentially no methodological extensions.

Table 2 lists the ten models that appear to be no longer actively developed and maintained. Two of those models (RUNS and GREEN) were developed at the OECD, which currently is a key partner in developing AgLink-COSIMO. Thus, lack of a sufficiently strong institutional backing is not a shared property for those models going out of business. In fact, it is difficult to find any general pattern in the institutional surroundings that determines a model's long-term survival, in particular since the models and networks are no longer there. However, one can at least hypothesize that the personal arrangements are key. For instance, the RUNS model at OEDC was developed by Dominique van der Mensbrugghe. When he had left the organization, the model was no longer being developed and applied[1]. Similar explanations may be found behind other models in Table 2. If a model depends on a single key person, the model is obviously very exposed to the professional or career path of that particular person, regardless of the size of the backing institution. Sustainable management of models seems to involve keeping key knowledge and skills in the long run, either by keeping key staff or by de-personalizing the model in a wider network, while ensuring that sufficient resources are being devoted to its maintenance. That relying on a small number of core people can present a risk to sustainability is noted also by others. For instance, Crowston, K., & Howison, J. (2005) find, in relation to an open source software project, that "overly heavy reliance on a small number of (possibly corporate funded) developers was a major threat to the sustainability of the project and thus to the suitability of the project"

Table 2:Models that are no longer actively developed

Model	Developer
ESIM	USDA/ERS + Göttingen
GAPsi	FAL
MISS	INRA
SWOPSIM	USDA/ERS
WATSIM	Uni-Bonn/FAL
GREEN	OECD
Michigan BDS	University of Michigan
RUNS (Rural-Urban-North-South)	OECD
WTO Housemodel	WTO
MEGABARE/GTEM	Australia Department of Agriculture, Fisheries, and Forestry (ABARES)

What can we learn from this? Why do certain models go out of business whereas others do not? We can formulate a few hypothesis:

- Because they are no longer relevant?
  - Technically outdated? Many of the surviving models depend on older methods that are basic but robust. GTAP follows essentially the same methodology since the 1990's.
  - Strategic decisions or developments are lacking? This is almost equivalent to stating that the model is no longer developed and used, and begs for additional questions.
- Because they cannot be maintained and updated? This explanation can be further disaggregated:

[1] Source: Personal communication with Yves Surry, professor emeritus in agricultural economics at SLU and one of the co-authors of the report of 2001.

- There are insufficient resources to fund model maintenance. This is clearly relevant in relation to the value of the model and the costs of its maintenance, but not alone. Financially strong organization such as the OECD and FAO were home to three of the models in Table 2.
- The costs of model maintenance are too high. This is difficult to quantify, but arguably plays a larger role, the larger and more complex a model becomes.
- The Human Resources are not available (the core modellers leaves). This is very likely a critical explanation in many cases when models are being left dormant or disappear. Nevertheless, we can observe that e.g. the CAPRI network lost its main developer Wolfgang Britz about 2014, but still exists six years on.

4 Characteristics of the modelling networks in the SUPREMA family

4.1 CAPRI

Within SUPREMA, the CAPRI network undertook an investigation of problems and possibilities for changes in the governance structure of CAPRI. The topic has been regularly revisited in CAPRI developers' meetings and consortium meetings during the SUPREMA project life.

4.1.1 Present structure

The network around the CAPRI model is not governed by any specific legal document or contract, but is rather an informal collaboration of researchers sharing common interests. The explanation to this lack of formalized governance is found in the background of the model. From the initial CAPRI project, starting in 1995, and up to 2014, most developments were involving the main developer and coordinator Wolfgang Britz in Bonn. The work was organized around a sequence of multi-partner EU funded projects that were coordinated by Bonn University (CAPRI, CAP-STRAT, CAPRI-DYNASPAT, CAPRI-RD). This central source of funding has been increasingly supplemented and replaced by smaller and medium sized projects that were not explicitly targeting CAPRI developments as such but specific research questions where some development work may be needed or useful. The consequence is that developments are not "automatically" aligned by the co-ordinator of the "main" CAPRI project, but require specific efforts on the part of the network.

During the first decade, there was no software versioning system in place, but access was managed via a network drive at Bonn University. The only way for "satellite" developments to find their way into the code base or database was via Wolfgang Britz. When the software was moved to the software versioning system (Subversion, or "SVN"), decentralized development by other network members became much simplified, but the need for coordination rose. Nevertheless, developments very frequently involved Wolfgang Britz for as long as he was involved in the network.

The SVN system has become the main way for the network to coordinate their efforts, because it requires the cooperative resolution of conflicting developments in a shared repository called the "trunk". All developers in the network can contribute code to the trunk. Many developments for specific projects or relating to specific model features take place in "branches". A branch is a copy of the trunk (or another branch) for which the SVN-software is able to keep track of differences. All members of the network can create model branches at their own discretion as a way of isolating "their" particular model version from changes in the trunk.

The CAPRI versioning system incorporates, the software that builds the various CAPRI databases from raw data, as well as the raw data itself. In many cases the raw data are associated with metadata that follows the data through the system. However, the CAPRI databases themselves are not versioned except for the (public) Stable Release versions. The reason is practical: binary data files cannot be versioned in the same way as the source code text files, but each change, however small, implies that a new copy of the data has to be stored. The programs that process raw data into CAPRI databases are also subject to the Stable Release process, and therefore tested for functionality and stability issues before being versioned and published.

Coordination in the network takes place in targeted e-mail traffic and weekly virtual "modellers' meetings". Strategic issues, such as the organization of the annual training session, are discussed in an annual "consortium meeting", generally taking place in connection with the annual training session.

However, lacking a formal contract or governance structure, the decisions taken at the consortium meeting have the character of recommendations. No document defines which types of decisions can be taken at the consortium meeting or how voting rights should be allocated.

The CAPRI Trunk is in constant flux. In order to provide a stable point of reference, the network publishes a “Stable Release”. The entire code base and all the compiled databases for all Stable Release versions of CAPRI are available to anonymous users via the CAPRI model homepage. This should be advantageous from a dissemination perspective, since at least in principle, the model is accessible to blind peer review. In practice, it means that the network is not aware of who is using the model and for what purpose, and there is no mechanism to collect fees or to limit the distribution and application of the model.

4.1.2 Areas requiring coordination and overhead

4.1.2.1 Database and baseline updates

The CAPRI system needs a database and a baseline. As applies to the development work we may also note here a shift in responsibilities and efforts from the large research programs of the past (CAPRI, CAP-STRAT, CAPRI-DYNASPAT, CAPRI-RD) to smaller projects that were critically dependent on the database and baseline updates. The need for updates depends on the kind of applications of the system. For “pure” research questions or empirical applications to long run horizons (2050 or longer) it may be perfectly ok to work with a moderately outdated database (e.g. base year 2004 or 2008) or baseline (AgLink projections from 2014 or 2017). In contrast, need for updates is usually urgent for policy relevant applications or for applications for private sector clients.

The key databases for the current build of the model are the consolidated FAO data, the “national database (COCO), the regional database (CAPREG), and the global database (GLOBAL). Albeit there is an aspiration to rely on homogeneous public data sources such as Eurostat, the “Eurocentric” focus of CAPRI led to an extensive search for supplementary, special data sets to complete or supplement the main data. Such complementary use of data sources is typically caused by the discovery of data problems, such as the dataflow involves a complex overlay of possible multiple sources.

The updates of COCO and CAPREG were mostly handled by EuroCARE and TI in the past, with TI clearly focusing on the extensive database for different farm types. The absence of division of labour in the database work is also triggered by the specialized experience accumulating within the consortium, insufficient written documentation and some original files not even subjected to SVN versioning. If the chance to share COCO/CAPREG database work with other agencies is small, it does not pay off for EuroCARE or TI to invest more in documenting their work.

The specialization is similar for baseline work. Most baseline work has been performed in the last years by EuroCARE and JRC-IPTS (and sometimes also UPM) as these agencies were requiring up to date baselines for their work. For EuroCARE this was mostly triggered by the demand of DG CLIMA projects (policy oriented work) and for JRC-IPTS by the annual DG AGRI outlook. Again similar is the situation for the CAPRI disaggregation module CAPDIS which so far was exclusively managed (both in term of updating necessary input data sets and development of the code) by JRC-ISPRA, and changes made without bothering to ‘warn’ potential other users.

4.1.2.2 Stable release cycle

In 2016 JRC-IPTS launched the first stable release project (star1) after prior discussion of project ideas in the network which was followed by a second one in 2017. The implementation of the stable release work is so far largely centralized at SLU with supplementary tasks for EuroCARE and TI. SLU (Torbjörn Jansson) effectively acts both as a “general” as well as a “release manager” for the stable release process with JRC-IPTS influential in the selection of supported features and supplementary improvement work. Technical commits are limited to the small circle of immediate project partners.

The concept seems to work well with tested release versions for general use emerging and gradual improvements achieved in the areas of supply model calibration, feed and fertilizer calibration. Also some knowledge spread on the testing procedures may be recognized. Potential problems could be the capacity problems of the SLU team working on the model, compared with the numerous cases of non-reproducible results and other issues to be tackled under the stable release heading.

4.1.2.3 Training

Training activities have been carried out in the past in the context of larger research projects where such activities have been explicitly written into the terms of reference (CAPRI, CAP-STRAT, CAPRI-DYNASPAT, CAPRI-RD, TRUSTEE). Two training sessions after the TRUSTEE project have been organized based on own funding and capacities by JRC-Ispra (2017) and SLU (2018). In both cases (and any previous training) numerous working days have been spent on top of the local organizing agency, by the network members acting as "trainers" to prepare the test simulations and training material.

The planning is initiated in the last years at previous year's consortium meeting with the selection of the local organizing agency. In the course of the year this agency then seeks in its own interest to obtain help from network members working in some subject area to achieve a broad coverage of topics and the filling of time slots foreseen to be used during the session.

4.1.2.4 Documentation

The comprehensive documentation by Britz and Witzke 2014 (https://www.capri-model.org/docs/CAPRI_documentation.pdf) has been updated in selected parts on various occasions such as before the 2019 training session in Seville. From 2020, a move of the entire documentation to an online "Wiki" format was initiated by TI on their own resources. As of June 2020, the work was well under way but not finalized. Despite the recent efforts to move the existing documentation into a more dynamic Wiki-format, the documentation is lagging behind in several areas.

4.1.2.5 Website, mailing list, SVN user rights

The website (www.capri-model.org) and mailing list (capritalks@listserv.dfn.de) are currently mainly managed by TI even though all developers have access to improve the website. Access rights to the SVN system, i.e. the Trunk, are currently managed in a collaborative manner by Wolfgang Broehl (U Bonn) and Alexander Gocht (TI), based on informal agreements. There is no dedicated content managing person for the mailing list and website in the network.

4.1.2.6 Code development and maintenance

Branching off for new code developments protects the trunk against disturbances, but also creates the need for subsequent reintegration into the trunk, if this is useful. Currently, there is no overview of what developments are available and how they are mutually compatible. In fact, there is no formal mechanism in place to decide on which such parallel developments are to be maintained and taken into the trunk. In the past few years, such decisions have been taken by consensus in the consortium meeting at the annual training session.

By the construction of a Trunk containing "everything", but also by the age and size of the model, the code contains many functions that have become obsolete. Obsolete code should eventually be removed, otherwise the code will become more difficult to understand for humans, more difficult to maintain, and the risk of mistakes will increase.

Under CAPRI-RD a red book of coding standard has been developed and some efforts have been made to initiate an update. However, the coding practices develop with increasing experience, and what was considered good coding twenty years ago is no longer necessarily up to standard. The continuous revision and maintenance is a continuous overhead process of the network.

4.2 MAGNET

The MAGNET model was developed at Wageningen Economic Research (WECR) as a successor to the LEITAP model. With agricultural issues increasingly connected to other fields in matters concerning, for instance, bio-energy, sustainability and climate change, LEITAP became increasingly complex. The increasing complexity of LEITAP led to a corresponding increase in the costs associated with using the model. First, it increased the costs of changing model aggregations due to the use of external data, which needed to be manually adjusted when the aggregation was changed. The LEITAP aggregation also had a high level of detail (mainly at the EU level) which increased runtime even when the detail was not needed to answer a specific research question. A final issue was the high cost of training new staff members, who were discouraged by the increasing complexity of the model. To reduce the costs of using the model, LEI made a considerable investment in recoding the model, an effort that has resulted in MAGNET. Development has been driven by the following key principles:

A modular setup around a GTAP core: the modular setup has been designed such that model extensions can be switched on through choices in a single parameter file, sometimes in combination with changes to the closure file. This allows new users to start with GTAP and then add extensions as needed. For experienced users, it facilitates the tailoring of the model to the research question at hand and eases debugging when developing a model. The GTAP model was chosen as a basis for MAGNET not only because it is the premier CGE model, but also because the GTAP network provides a common background, which enables comparison across a wide variety of other CGE models developed from GTAP.

Data are kept and processed at the lowest level of detail; all databases are kept in their original format and processed at the lowest level of detail to increase aggregation flexibility. All data changes and adjustments are coded in GEMPACK to enhance tractability and quality control. This approach also facilitates the updating of datasets, since the same code can be applied to the updates.

MAGNET is in several aspects a modular extension to the standard GTAP model. As such, it benefits from the GTAP database, from the GTAP courses for users, and from part of the theoretical structure. This is in contrast with other models in the SUPREMA family, which have their own specific databases and theoretical structures. The MAGNET database embeds next to the GTAP data, many supplementary datasets that enables it to disaggregate sectors beyond the GTAP level and include modules on land, land use and nutrition for example. Currently the MAGNET database is transformed to a database system called Datawarehouse that enables better version control, lower costs for updating the database, and is connected with Power-BI in order to have much more opportunities for visualisation. MAGNET is run by a consortium of three institutions: Wageningen Economic Research, the Joint Research Centre and Thünen Institute. WEcR hosts the MAGNET model and database and bilateral consortium agreements between WEcR and partners conceal the consortium. The consortium agreement specifies obligations of partners and the consortium is open for new partners. The strategy is discussed among partners in consortium meetings where partners indicate their priorities and where synergy effects are identified. Model developments remain however, largely project based as core funding is very limited. Dedicated MAGNET training courses are organised within the consortium. The consortium has a website, with information on model documentation and a list of all magnet related publications by all consortium members (https://www.magnet-model.org/).

Funding for the model comes from both projects and own resources of the partner institutes involved. There is a user fee however, which is required for the GTAP database and the GEMPACK software.

4.3 AGMEMOD

4.3.1 Current set-up

AGMEMOD is formally organized as a partnership, governed by a Memorandum of Understanding (MOU) under European law. The partnership has three main bodies:

- The Executive Committee;
- The Core Modelling Group;
- The National Experts

All parties to the partnership (all the groups listed) are invited to meet at least annually in a general assembly, where certain decisions can be taken. There are also more frequent (bi-annually at least) ordinary meetings of the Executive Committee and the Core Modelling Group.

The **Executive Committee** is responsible for the overall coordination and strategic planning. It is the legal entity acting on behalf of the AGMEMOD Partnership. Most importantly, it provides the combined model versions on a central server, and coordinates baseline work and results review. It consists of researchers from Thünen (Martin Banse) and WEcR (Myrna van Leeuwen). Currently, both organisations prepare and compile most of data updates and model revisions made available on the central server (based at the Thünen Institute), partly with the support of the EU Commission and especially with the support of other sources (Thünen gets national public funding for maintaining and extending the model, a German baseline and scenarios, whereas WecR gets public/private support to provide an annual national baseline for Dutch agriculture). The cross-funding is crucial for financing the core team's effort, but it creates risks in case when this financial support fails or substantially declines that neither maintenance nor updates could be conducted. The groups involved in the Executive Committee meet at least once a year to discuss important issues of the Partnership with respect to model development, databases and related tools.

The **Core Modelling Group** is responsible for work relating to scientific contents and output. From the MoU 6.2.1: "The Members of the Core Modelling Group […] are responsible for the technical co-ordination between the different parties undertaking the modelling activities and for monitoring the different achievements. They shall also undertake provisions concerning model-based commodity market analysis and policy expertise. Their task comprises analysis of the baseline results across countries for different groups of products". Members are approved by the Executive Committee and the Core Modelling Group. National experts contributing significantly to the AGMEMOD project can become members of the Core Modelling Group.

All parties (signatories of the MoU) have access rights to the combined model and the baseline. If a party does not comply with the MoU, it "defaults", and will lose their access rights. All parties fund their own work, i.e. there are no joint financial resources in development of the Central Version apart from joint projects. There may be financial inputs for specific non-core partners unknown to the Partnership. Parties have different options to sign the MoU, either as institution or scientist involved. Those who have signed the MoU download the latest version via a link.

The group of **national experts** are the broadest group of model users. In principle, they are invited to provide data and validate baseline results for their respective countries, and may provide or improve their respective country models. The group of national experts is an informal group and members are not part of the MoU.

Next to national teams the AGMEMOD partnership has access to a European Network of National Market Experts, reviewing AGMEMOD outcomes for AGMEMOD Outlook workshops. They are informed by an Excel tool which presents model outcomes on a country-base either as tables or graphs. This tool has been developed by the teams at Thünen and WEcR and is regularly extended and improved. The Market Experts provide their reviews and feedback either on an AGMEMOD workshop or in written form by email.

Up to now the AGMEMOD teams provide regular training on the development of databases, model use, scenario development and model revisions which are often free off-charge for participants to enable knowledge transfer and to broaden human resources in model use. These sessions are often funded by involved institutions, sometimes by projects or third parties. Upon request additional trainings sessions can be scheduled. Training on AGMEMOD is also part of the Doctoral Certificate Program in Agricultural Economics of German universities. Additionally, there are regularly organised training courses for African economists who aspire to work with the model.

The Partnership has a website (https://agmemod.eu/) where public and partners can get information about the model, latest developments, see selected results from projects applying AGMEMOD, and upcoming trainings.

4.3.2 Areas for improvement

The AGMEMOD database was originally based on EUROSTAT production and price statistics and EUROSTAT supply and use tables. For countries where such data is not available, national statistics has been used. After the end of regular publication of supply and use tables by EUROSTAT the use of national data became more important, but at the same time proved more difficult to compile as national institutions partly gave up certain data collection. The regular publications of Short-term Outlook by the EU Commission serve as a basis for most recent data on area use, production, and number of livestock. But this data is not harmonized with data provided by national agencies at Member State level which creates sometimes conflicts. Trade data are taken either from COMEXT database or FAOSTAT (fish), in some cases also from national statistics. The AGMEMOD database is updated twice a year by groups organised in the Executive Committee, supported by the EU Commission and some national teams. Based on these updates new baselines are established by the groups organised in the Executive Committee and are then made available to all members of the partnership in terms of a new Central Version of AGMEMOD.

The use of data and the flow of work make it complicated to request regular support from national teams with respect to data and model updates:

- The use of national data may conflict with data from the EU Commission and Eurostat.
- National teams have different timeframes and schedules than the centralized procedures.
- Provision of the Central Versions creates a free-rider problem and does not allow creating opportunities to distribute overheads.
- Useful developments and extensions of national teams are often not reported back to the Central Version, but if they are reported resources are missing regularly to integrate those developments in the Central Version.
- Most of the cost related to software update and -maintenance is financed by the Thünen Institute and WEcR. Only a smaller part of the cost is covered by projects. There is a need for a long-lasting financial support for the software development of AGMEMOD.

With respect to governance it can be noted, that although national teams are invited to meetings and to provide feedback to the Outlook, the integration in model development is quite fragmented for those members of the Partnership which are not part of the Core Modelling Group. A joint acquisition of research funds and combined efforts to develop a common research strategy is crucial, while at the moment the development is ruled by erratic appearance of calls for tendering and research proposals. To achieve this objective, more exchange meetings (for example bi-annual meetings) might help. To limit the cost and resources a series of physical meetings combined with online tools might be appropriate to better integrate more partners. Also sharing the financial burden by opening the Executive Committee and the Core Group might be an additional option.

Over the last years the establishing of an informal Network of National Market Experts became an important approach for developing a sound and broadly accepted outlook by both Member States and the executive committee. A detailed analysis of medium-term development of agri-food markets not only at the aggregated level of the entire EU, but also at individual Member States is an important contribution of AGMEMOD which requires a long-term strategy. Here, an intensive exchange of members of the Partnership with short-term expertise provided by market experts also provides valuable input for further model developments. To strengthen this network a special access to output related parts of the website may prove helpful as it can help to catch additional feedback. Further, the Outlook Workshops with detailed analyses at Member State level can be extended by online versions. But the most important topic is to place the Partnership within a growing network of active partners.

4.4 GLOBIOM

GLOBIOM (https://iiasa.github.io/GLOBIOM/) has been developed at IIASA since 2007, and has been accessible to external collaborators in the context of specific joint projects or for strictly academic purposes by PhD students, often in the context of the IIASA Young Summer Student Program (YSSP). The most successful example of this type of collaboration is GLOBIOM-Brazil, a refined version of GLOBIOM for Brazil, which has been developed by the INPE team with support of IIASA since 2012. For external collaborators GLOBIOM code includes data that are available via GITHUB once authorization have been provided. This is a pre-release of a fully open source version which is under preparation. For the IIASA GLOBIOM team, SVN serves as a platform for exchanging code developments. So far, the external users can modify their versions of GLOBIOM, but do not have the possibility to change the central GLOBIOM code at IIASA, and thus cannot directly contribute to the model development. Thus, GLOBIOM is functionally an institutionally owned model of IIASA. Overhead costs such as data updates, documentation and software architecture, has so far been minimal, and mostly covered by external projects. Training for students or project partners is given on a per-project-basis and typically involves a two-week stay at IIASA.

4.5 IFM-CAP

IFM-CAP was developed by JRC starting from 2013 at the request from the Directorate General for Agriculture and Rural Development (DG AGRI) of the European Commission with the purpose to improve the quality of the ex-ante CAP policy assessments upon the existing aggregate (regional, farm-group) models. The main driver for the IFM-CAP model development was the gradual introduction of farm-specific measures within the CAP (e.g. farm specific decoupled payments, CAP greening measures) which demanded the application of a modelling system able to capture policy representation and impacts disaggregated at the micro level. A simplified IFM-CAP prototype version was finalised in 2015 (Louhichi et al. 2015), whereas an improved version was completed in 2018 (Louhichi et al., 2018a).

All independent users of IFM-CAP are at the JRC, where the development team officially consists of 3 ½ people. One of them is permanently employed, while the others are on limited term contracts. Currently, one IFM-CAP team member has been assigned to research activities on other topics not related to IFM-CAP. Each time the staff composition changes, some model competence has to be re-built. However, the staff leaving might also continue to contribute to the model from their new affiliation. Currently, the team assesses that the model is in a good position, with a successful and extensive revision of the code to make it more transparent, modular, computationally faster and user friendlier.

The use of IFM-CAP is guided by a Graphical User Interface. The reporting and the visualization of the results is done through the software Qlik Sense. The model version is controlled through the SVN-software.

As regard for the future of the model, it is continuously applied for policy analyses for the Commission and will continue to be developed in the near future. The main applications of IFM-CAP for policy analyses in the past include the analysis of the future pathways for the European agriculture sector (M'Barek et al., 2017), the evaluation of the impact of CAP greening (Louhichi et al. 2017, 2018b) and the contribution to the impact assessment of the European Commission proposal for the CAP post-2020 (European Commission 2018).

IFM-CAP faces some unique challenges in the SUPREMA family: From a staff and development perspective, the JRC does not have students in the way the other research institutes have, and hence parts of the development work is outsourced to external consultancies. This limits the way that developments can be done, as the external resource is typically tied to a limited domain and not available as a resource in parallel developments. More importantly, the IFM-CAP is based on micro data from FADN, which is governed by strict access and dissemination rules. This makes an expansion of the developers' network to external institutions cumbersome. Otherwise, extending the network to broaden the group of users and developers would be interesting, if it could be done based on external project funding.

The data access challenges might be overcome by a separation of program code and data. The model code in itself could be shared with external partners. If external users could apply the model to data that they have the rights to use, this might be an option for extending the network. However, this would require that each external partner who gets access rights to the FADN data, estimates the required model parameters and calibrates the model to the base-year by herself, based on their local FADN sample. An alternative option is that the IFM-CAP team in JRC shares the calibrated parameters with partners, but only for their local FADN sample, as long as the external partner provides proof of access rights to all variables in their local FADN sample that are needed to calibrate the model.

4.6 MITERRA-EUROPE

The MITERRA model was initially developed at WUR/ALTERRA in 2006-2007, and is only used internally at WUR. It could therefore be considered an in-house model at this point, as the modellers involved are based at WUR, besides a few external guest researchers. In the CCAT (Cross-Compliance Assessment Tool) a two-way linkage was developed with CAPRI to assess impact of environmental measures. Later most of the calculation rules for nitrogen emissions have been included directly in the CAPRI model. Within the framework of SUPREMA a linkage has now been made with AGMEMOD, which is foreseen to be expanded in the future. Such links, however, do not include the full model, but rather a selection of key indicators based on simplified data. The data that drives the full MITERRA model is spatially disaggregated on the sub-national level, whereas e.g. AGMEMOD runs for entire countries, and this limits the degree of integration that can be achieved.

The funding from MITERRA comes from projects. During the last few years, there were fewer projects involving MITERRA, and consequently the staff has worked with other research. However, in the last year new projects started (e.g. Nutri2Cycle and internal WUR projects) and an update of the model is ongoing. Besides, the model is being revised and updated with the ambition to reach what WUR calls "Status A", a quality label for models that can be used in policy advice. This requires external assessment and a consolidated documentation.

Currently, the development team is small and with a limited capacity. The network could be extended internally within WUR and externally to other research teams in Europe. Both developments would require lower entry barriers to the system, and such work is indeed under way in the process of reaching “Status A”. The model code is versioned on a GIT server and might be made available as open-source in the future. For the external extensions, the time horizon is longer term.

5 Comparison of SUPREMA models

An objective indicator of the extent to which a model is being applied and developed is the number of publications that are produced. Figure 1 shows publications per model over ten years (including 2020). The statistics were extracted from Uppsala University Library Search engine, and complemented with a comparison of publications reported on the model home pages, if any. It consists of the sum of (i) peer-reviewed articles, (ii) working papers, (iii) conference papers and (iv) a few book chapters (mainly AGMEMOD). Not included are publications where said models are only mentioned or used as a source reference. In the publications that are included in the statistics, the model was either used as the main method of analysis or as a part of a wider analysis consisting of two or more models. There is a lag in reporting in the library database, as evident from the sharp drop in publications in 2019 and 2020, likely leading to an underestimation of the true number of publications. This should not bias the results.

Models have different purposes. In-house specialized models such as the IFM-CAP are likely to produce less scientific output, whereas open networks working with models with broader scope are more likely to aim for journal publications. By including working papers and technical reports, we hope to level out some of those differences in aims. Based on the numbers of Figure 1, the models can be broadly grouped in two categories: the four “core” models (CAPRI, MAGNET, AGMEMOD and GLOBIOM) together account for 377 of the 399 publications in the figure, against 22 for IFM-CAP and MITERRA together. In the former group, CAPRI and MAGNET account for 134 and 137 publications respectively, with CAPRI being more prolific in the beginning of the period and MAGNET more so towards the end.

The number of publications is not only indicating how much the models are being used, but should also bear some positive correlation with the probability of attracting project funding. Based on the figure, but without any economic data, we conjecture that MAGNET and GLOBIOM have experienced a period of more extensive project work.

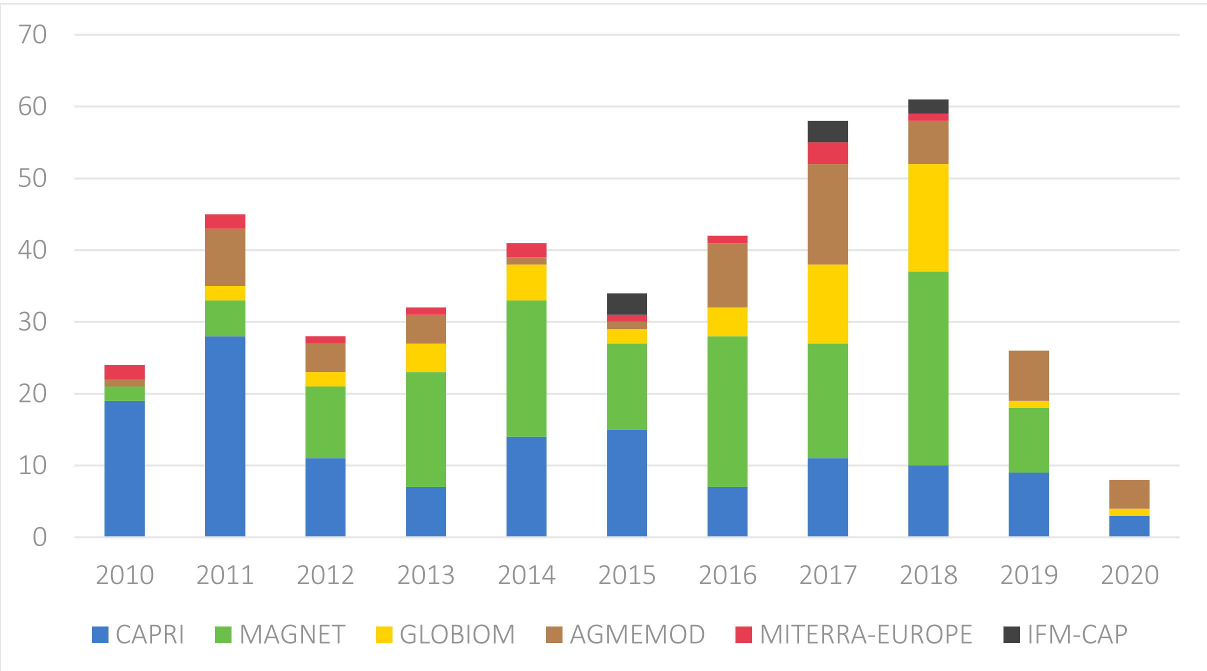

Figure 1: Publications per model over time, including journal papers, conference proceedings, working papers and technical reports. Source: own research.

In order to characterize the developer networks, the research teams were asked to estimate three numbers: (1) the number of individuals involved in the development and applications of the models, (2) the number of full-time equivalents engaged, and (3) the number of institutions involved. Clearly, any given individual can be involved at different levels, ranging from a user that does not contribute to the development in any way to a core developer of the code. Here, we are interested in the developer networks and therefore attempt to exclude users that do not contribute in any way to the development. Since the models involved are generally not openly available, all users are to some extent also developers. The following formulations were used in the enquiry:

- How many "full time equivalents" (FTE) are working with the model?
- How many different people are involved, even if sporadically or part time?
- How many different institutes/institutions/organizations are involved in the network?

The responses are summarized in Table 3. The FTE estimate is an indicator for how much human resources are commanded by the project network. In that respect, GLOBIOM and MAGNET dominate the list, with estimated 18 and 15.8 FTE respectively, followed by CAPRI and AGMEMOD with 9 and 4.5 FTE respectively. For AGMEMOD, mentioned FTE only cover the human resources employed by Thünen and WEcR, human resources working with and developing AGMEMOD in other institutions are not included. Numbers mentioned for AGMEMOD are, therefore, not comparable with the other models. IFM-CAP reports a number of FTE that is only slightly less than that of AGMEMOD (three to four compared with 4.5 for AGMEMOD). The human resources commanded by MITERRA in terms of FTE (1.5) is presently the lowest of any model in the family.

We can also see that for instance the number of institutions linked to the AGMEMOD partnership (34) is large compared to the other networks, which is explained by the nature of the model. The number of institutions involved in CAPRI and GLOBIOM are similar (10) even though the number of FTE differ. The developer network of MAGNET appears from this perspective to be the most highly concentrated, with four institutions accounting for all of the staff. However, such comparisons across models hide fact such that the AGMEMOD network only counted FTE at two core institutions and not those engaged at the partner institutions. Nevertheless, the institutional networks of CAPRI, AGMEMOD, IFM-CAP and MITERRA appear to be on average “thinner” than those of GLOBIOM and MAGNET in terms of FTE per institution.

If we divide the number of full time equivalents by the number of individuals, we obtain an indication of the degree to which work is shared among the staff involved, albeit this likely conceals the fact that a few people work close to full time whereas others are only weakly linked with the network. Nevertheless, for MAGNET, the staff on average works close to full time with the model, and for IFM-CAP, the ratio is also above 50%. For AGMEMOD, MITERRA and to some extent also CAPRI, the network seems to contain many individuals that devote only a minor share of their time to the model.

Table 3. Sizes of the modelling networks reported by the respective teams

Model	A: Full time equivalents	B: Number of individuals	C: Number of institutions	A/B****
GLOBIOM	18	40	10	0.45
MAGNET*	15.8	22	4	0.72
CAPRI**	9	30	10	0.30
AGMEMOD***	4.5	21	34	0.21
IFM-CAP	3-4	6	4	0.58
MITERRA	1.5	5-10	2	0.20

* MAGNET is connected to the GTAP network. That has not been accounted for in the number of institutions.

** CAPRI can be anonymously downloaded and used. Such users are excluded from the accounting, and also excluded are participants of the annual training courses that do not turn into actively contributing model users.
*** The number of full time equivalents and persons only account for the staff at Thuenen and WEcR, whereas the resource use at the members of the Partnership is unknown.
**** If we use the mid-points of the ranges, i.e. 3.5 FTE for IFM-CAP and 7.5 persons for MITERRA

The models and their associated developer networks can be categorized along several dimensions. From a governance or business model perspective, it seems useful to classify them according to

- the broadness of the scope of the model, and
- the character of the developer community

One way to assess the scope would be to count the number of variables and indicators produced, in combination with the number of types of outputs computed. At the one extreme, model can be narrowly focussed on a particular task, even if that task is done very comprehensively. Towards this end of the spectrum we find MITERRA and IFM-CAP. At the other end of the spectrum we find "do-it-all-models" that model a wider range of phenomena, from the economic behaviour of different agents, trade flows, biophysical balances and environmental impacts. Towards this "broad scope" end of the spectrum, we find the four core models. For reference, we put the standard GTAP model into the chart as a more narrowly focused model with a large and open network of contributors.

As regards the developer community, this can be open in the sense that many developers from different institutions contribute to the development of the underlying software and databases, or closed in the sense that development is centralized to one or few institutions. Furthermore, the model can be openly accessible or access can be restricted in various ways, either by formal arrangements or by technical barriers. If again the SUPREMA models are categorized along these two dimensions, we would find CAPRI and AGMEMOD in the more open range whereas the other models are more restricted in terms of access. For AGMEMOD, the distributed nature of the model calls for a wider community of national experts, which is formalized in the MOU. In the CAPRI case, the open nature has been partly a consequence of the history of the model, partly a conscious strategy to facilitate peer review and acceptance.

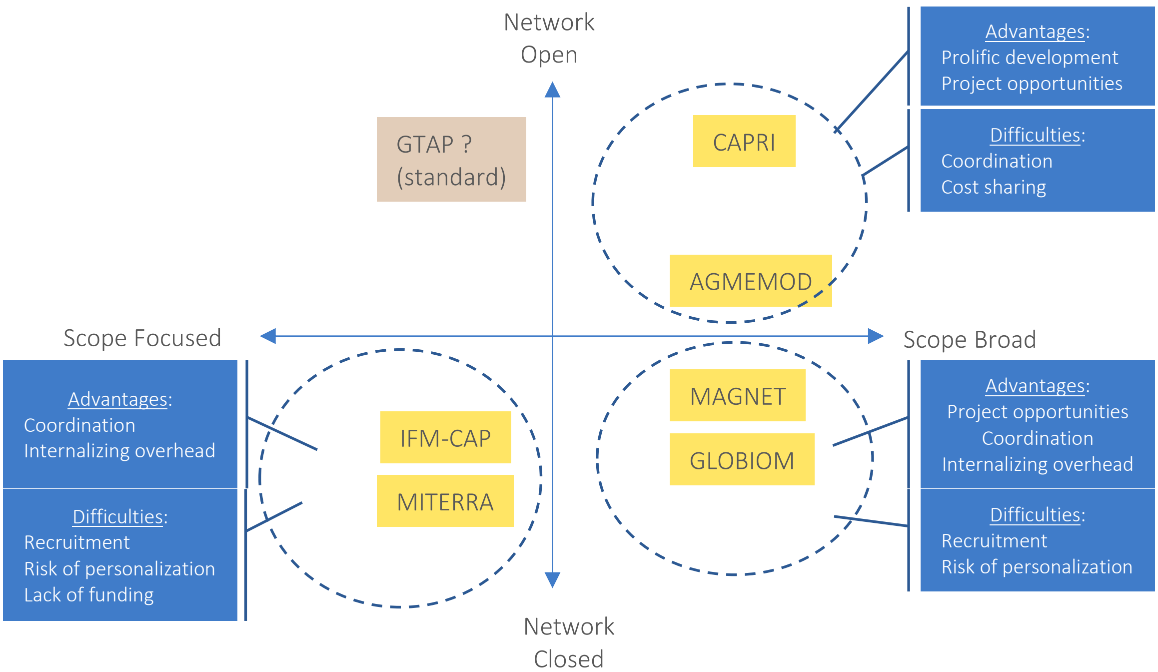

Figure 2: Two dimensions of models and networks with strengths and weaknesses (opportunities and threats)

If we start in the lower left quadrant, we find IFM-CAP and MITERRA. Being in this quadrant is associated with two main advantages: Coordination of development and decision-making is simplified because of the smaller/closer network. Internalization of overhead costs such as documentation and code maintenance should be straightforward. The main difficulties are likely to be recruitment of new modellers and users, a high risk of personalization and loss of competence in connection with staff turnover, and possibly also a lack of funding if the development is depending on external projects. The latter seems to be the case for the MITERRA model, where the team reports that lack of funds has forced some staff to work on other issues in the past years and documentation to lag behind. For IFM-CAP, the situation from the funding respect looks better with presently three persons working full time and another person half time with the model.

In the lower right quadrant, we find MAGNET and GLOBIOM, with broader scope but still a relatively closed network structure. Due to the smaller size of the networks, these models share the advantages of lower coordination efforts and the internalization of overheads. In addition, they are clearly benefitting from a much larger array of project funding opportunities. This is also reflected in the very good recent publication record of those models, evident in Figure 1. On the side of difficulties, we find the same risks on the human resources side. Recruitment of new competence can be difficult, and there is a certain risk of personalization of the skills and knowledge built into the core modelling team. This might be a potential area of development for these models.

If we finally look at the upper right quadrant, we find the models CAPRI and AGMEMOD, with wide scope and larger/more open developer networks. The key advantages of this quadrant come from the prolific development of new modules or technical solutions that can arise from a wider user/developer community. However, this advantage comes at a high cost in terms of high costs for coordination of model developments and of decision making, whereas the later has not occurred in the case of AGMEMOD so far, and also in free-rider problems when it comes to overhead costs. In particular, CAPRI has yet to find a way to handle the coordination issue, which AGMEMOD to some extent has solved using an MOU. In the long run, it seems critical to find solutions to ensure that overheads are funded.

Table 4 summarizes some of the characteristics of the four core SUPREMA models, based on interviews carried out by Hyungsik Choi in 2018. Regarding Ownership, no institution claims ownership of CAPRI, whereas the other models have either public institutions behind them (IIASA, WEcR) or are managed by a formalized organization (the AGMEMOD partnership). As regards networking, the AGMEMOD network is the largest, followed by CAPRI, whereas the development of GLOBIOM is comparatively centralized to IIASA, and MAGNET developments are shared by three institutions. The funding, in all cases, comes from projects, but all networks report that some work is carried out using "own resources". There is no fee-based funding of overheads except for the GTAP database fee extracted for MAGNET. Only the CAPRI model is publicly available for download, but also there, development is limited to a club of network members. The models all have some documentation, but it might be partly obsolete (CAPRI), difficult to obtain (AGMEMOD) or in development (GLOBIOM). Asked about whether they think that their particular modus operandi gives rise to governance issues, only the CAPRI and AGMEMOD teams reported that this might be the case.

Table 4 Characteristics of modelling networks

	CAPRI	GLOBIOM	MAGNET	AGMEMOD
Ownership	No ownership	IIASA	WEcR	AGEMEMOD Partnership

Network	EuroCARE, Thünen, SLU, JRC, UPM (Madrid), Bonn University, NIBIO (Norway), UC Dublin, and a few more in varying constellations.	INPE (Brazil), CARR (China), RANEPA (Russia), RTI (USA), KU Leuven (Belgium), Fundacion Bariloche (Argentina) and others	WEcR, TI, JRC	WECR, TI and other 30 institutions
Funding	Public/project	Public/project	Public/project, license fee	Public/project
Availability in model/database	Freely available	Closed, but progress towards public GitHub	Closed, License fee for GTAP database	Sign MOU, no fees involved
Documentation	Exists	In progress	Exists	Exists*
Problematic issues	Low administration costs, but high technical barrier for using CAPRI, not strong leadership for model governance	None reported	None reported	Different interests in a large consortium

Source: Based on interviews by Hyungsik Choi, 2018.
* The documentation is integrated in a help-system, which is part of the model and was not independently available from a website. It is integrated on the website as well. Also a book has been published describing AGMEMOD (Chantreuil et al., 2011, https://link.springer.com/book/10.1007%2F978-94-007-2291-0)

6 Recommendations

In this section, we provide a very brief summary and a few bullet-point recommendations per model. The recommendations are limited to aspects relating to the networks of the developers, based on self-assessment and the analyses presented above. It would certainly be valuable to further investigate into many other aspects, such as the technical procedures used in code development and testing, on the structure of software itself, on funding structures and of the details of the contractual frameworks governing the developers. Such extended analyses would require the systematic collection of a much range of data from the different models, and together with some standards for good practice in the field, it could be the basis of a richer set of recommendations. However, such data are difficult to obtain. The software and versioning platforms are partly different and typically not open to external access, and financial data of sufficient detail may be unavailable even to the organizations themselves. Therefore, such extended analyses are beyond the scope of this deliverable. Nevertheless, the preceding analyses and comparisons together with these concluding remarks should be useful for the modelling teams.

In the final part of this section, we conclude with a brief discussion of to what extent there are cross-model possibilities for mutual benefits of the network. That kind of conclusions are somewhat beyond the scope of this deliverable, but nevertheless the author considered it useful to document ideas and suggestions in that direction that arose during the project life, in particular the final project meeting.

6.1 CAPRI

CAPRI would likely benefit from a more formalized governance structure, striking a balance between on the one hand open access and a prolific methodological development and on the other hand ascertaining adequate funding of overheads and maintenance. The following three points could make a start:

- A legal entity should be created to govern the model maintenance and development, and in particular the financial aspects that are presently informally shared among several parties. The legal body could have both institutional and individual members.
- License fees for the database can help fund overhead while keeping the model open.
- Better separation of database and model code would facilitate peer review and licensing.

It would be the task of the new legal body to work out the details of shared responsibilities and technical and administrative processes for ensuring the long term survival of the model.

6.2 MAGNET

MAGNET has established a suitable governance structure with CA governing 3 partners
MAGNET is not lagging behind, but might nevertheless benefit from making the model more available for external contributors both in order to vitalize developments and for peer review.

- No change is needed, but the team might want to consider publishing a limited version of the model.
- MAGNET may serve as inspiration for GLOBIOM and CAPRI.

6.3 AGMEMOD

AGMEMOD has a large network of national teams, a smaller network of core developers, and a very small executive committee. In addition, there is a Network of National Market Experts, which provides feedback to the outlook outcomes.

Key challenges: In total, this structure seems to serve its purpose, but also presents some caveats

- It is difficult to share overhead cost for database updates, baseline updates and tool developments (free-rider problem).
- There are difficulties to feed developments by national teams into the Central Version due to different degrees of (perceived) involvement, which may lead to differences in felt responsibilities to care for the whole.
- The structure hinders streamlining strategic developments.

Recommendations (derived in dialogue with core modellers): Improvements may be achieved by

- Broadening the Executive Committee and Core Group to share overhead cost
- More frequent developer and strategy meetings with participants of the core group to get broader understanding of challenges and sharing of responsibility
- Try to reap economies of scale and scope by combining and harmonize scenario studies for different government and non-governmental clients by different Partners at EU, at Member Stare or at groups of Member States (e.g. COVID19 assessments for different countries)

6.4 GLOBIOM

So far, an ”informal club” model worked well. GLOBIOM has taken steps towards open access (OA)

- OA could bring benefits in terms of distributed knowledge and prolific development
- OA should be preceded by a governance structure
  - to share overhead
  - for decision making
  - for coordination of tasks
- It may be wise to look at the business model of MAGNET

6.5 IFM-CAP

IFM-CAP development depends on a team of 3-4 people in the same institution (JRC in Seville) which devote a large share of their time to the model development. This should be a good basis for a model that so far has a fairly focused area of application.

Challenges:

- The use of the model outside of the core institution seems limited. That may hamper the inflow of innovations to the model and the benefits in terms of scrutiny that come with an extended user base.
- It is difficult to disseminate the model due to the confidential nature of the FADN data.

Recommendation:

- Think of ways to extend the user and developer network beyond the JRC, for instance to institutions working with national level analyses. A key challenge would be to find ways to parameterize the models with FADN data.

6.6 MITERRA

The development team is small and with limited capacity. Documentation, training etc. is lagging behind.

- The network could be extended
    - internally within WUR and
    - externally to other research teams in Europe
- Lower entry barriers: process of reaching “Status A”.
- External: ”co-branding” as indicator module for broader models such as CAPRI, AGMEMOD and perhaps others too?

6.7 Cross-cutting discussion points

- There is potential to learn from each other. For instance, the governance strategy of MAGNET with a formalized consortium agreement might be useful for GLOBIOM and CAPRI
- Data bases are expensive. MAGNET relies on GTAP to a large extent. Can we find other ways to share data among ourselves, such as e.g. the processed FAOSTAT data of CAPRI? Such data exchange would require strictly formalizing the data definitions and make update schedules highly predictable to enable planning of modelling work that depends on the data. The coordination efforts required would be considerable, but so also the potential benefits.
- Can the models complement each other better so that each becomes more focused, such as including MITERRA in a component fashion in the broad-scope models? This was to some extent successfully tested by CAPRI and AGMEMOD already.
- Testing procedures from CAPRI stable releases might be applicable to other systems as well. Some initial steps were taken to implement the GUI of CAPRI for the GLOBIOM model. If that is completed, many of the GAMS routines for systematic stability checks might be transferable too.
- The modelling networks probably bring mutual benefits already via the exchange of staff and knowledge. Staff/developers leaving one network has some likelihood to find new employment in another, since the market for modellers is thin. Certain generic skills are transferable.
- Finally, we note that the “soft linking” in terms of harmonization of baseline assumptions and ensamble modelling studies has already resulted in certain mutual benefits across the modelling teams.

7 References

Apland, J. & Jonasson, L. (1992). *The conceptual background and structure of SASM : a Swedish Agricultural Sector Model* . Rapport, Sveriges lantbruksuniversitet, Uppsala.

Beecher, K., Capiluppi, A. & Boldyreff, C. 2009, "Identifying exogenous drivers and evolutionary stages in FLOSS projects", *The Journal of Systems & Software*, vol. 82, no. 5, pp. 739-750.

Chantreuil, F., Hanrahan, K.F. van Leeuwen, M., 2011, The Future of EU Agricultural Markets by AGMEMOD. doi: 10.1007/978-94-007-2291-0.

Crowston, K., & Howison, J. (2005). The social structure of free and open source software development. *First Monday*, 10(2). https://doi.org/10.5210/fm.v10i2.1207

European Commission. 2018. "Proposal for a Regulation of the European Parliament and of the Council for 'CAP post 2020'." European Commission, Available online at: https://eur-lex.europa.eu/resource.html?uri=cellar:aa85fa9a-65a0-11e8-ab9c-01aa75ed71a1.0003.02/DOC_1&format=PDF

Louhichi, K., Ciaian, P., Espinosa, M., Colen, L., Perni, A. and Gomez y Paloma, S. 2017. "Does the crop diversification measure impact EU farmers' decisions? An assessment using an individual farm model for CAP analysis (IFM-CAP)." Land Use Policy 66: 250–264.

Louhichi, K., M. Espinosa, P. Ciaian, A. Perni, B. Vosough Ahmadi, L. Colen and S. Gomez y Paloma 2018a. "The EU-Wide Individual Farm Model for Common Agricultural Policy Analysis (IFM-CAP v.1): Economic Impacts of CAP Greening," European Commission, Joint Research Centre, EUR 28829 EN, doi:10.2760/218047

Louhichi, K., P. Ciaian, M. Espinosa, A. Perni, and S. Gomez y Paloma 2018b "Economic impacts of CAP greening: application of an EU-wide individual farm model for CAP analysis (IFM-CAP)." European Review of Agricultural Economics 45(2): 205–238.

M'barek, R., J. Barreiro-Hurle, P. Boulanger, A. Caivano, P. Ciaian, H. Dudu, M. Espinosa Goded, T. Fellmann, E. Ferrari, S. Gomez y Paloma, C. Gorrin Gonzalez, M. Himics, K. Louhichi, A. Perni Llorente, G. Philippidis, G. Salputra, P. Witzke and G. Genovese (2017): Scenar 2030 - Pathways for the European agriculture and food sector beyond 2020. No JRC108449, JRC Working Papers, Joint Research Centre (Seville).

Medappa, P.K. & Srivastava, S.C. 2019, "Does Superposition Influence the Success of FLOSS Projects? An Examination of Open-Source Software Development by Organizations and Individuals", *Information systems research*, vol. 30, no. 3, pp. 764-786.

van Tongeren, F., van Meijl, H. & Surry, Y. (2001). Global models applied to agricultural and trade policies: a review and assessment. *Agricultural Economics*, vol. 26 (2), pp. 149–172 John Wiley & Sons, Ltd. DOI: https://doi.org/10.1111/j.1574-0862.2001.tb00060.x

Zhou, M. & Mockus, A. 2015, "Who Will Stay in the FLOSS Community? Modeling Participant's Initial Behavior", *IEEE Transactions on Software Engineering*, vol. 41, no. 1, pp. 82-99.