HAL open science

The European agro-economic model AROPAj

Pierre-Alain Jayet, Athanasios Petsakos, Raja Chakir, Anna Lungarska, Stéphane De Cara, Elvire Petel, Pierre Humblot, Caroline Godard, David Leclère, Pierre Cantelaube, et al.

To cite this version:

Pierre-Alain Jayet, Athanasios Petsakos, Raja Chakir, Anna Lungarska, Stéphane De Cara, et al.. The European agro-economic model AROPAj. INRAE-PSAE. 2023. hal-04109872v1

HAL Id: hal-04109872

https://hal.inrae.fr/hal-04109872v1

Submitted on 8 Jun 2023 (v1), last revised 28 Aug 2023 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

The European agro-economic model *AROPAj*

Pierre-Alain Jayet Athanasios Petsakos Raja Chakir
Anna Lungarska Stéphane De Cara Elvire Petel
Pierre Humblot Caroline Godard David Leclère
Pierre Cantelaube Cyril Bourgeois Mélissa Clodic
Laure Bamière Nosra Ben Fradj Parisa Aghajanzadeh-Darzi
Gaspard Dumollard Ancuţa Isbăşoiu Juliette Adrian
Gregory Pilchak Myriam Bounaffaa Delphine Barberis
Coline Assaiante Maxime Ollier Loïc Henry
Alessandro Florio Ines Chiadmi Eva Gossiaux
Erica Ramirez Maxence Gérard Julie Reineix
Olga Zuravel Lisa Baldi Mingzhu Weng

PSAE - INRAE-AgroParisTech - Université Paris-Saclay
22 place de l'agronomie, F - 91220 Palaiseau

access: https://www6.versailles-grignon.inra.fr/economie_publique/Publications

DOI: https://doi.org/10.17180/nxw3-3537
<2023/06/08 16:23:28>

email pierre-alain.jayet@inrae.fr

Contents

Abstract

This publication presents the European agro-economic *AROPAj* model. Throughout recent years the model has been adapted to suit *FADN* data and the expansion of the European Union (*EU*). The revised model, which is presented here, is based on the same principles, modularity, and generic conceptions of the previous renditions. The strength of the model lies in the evolution of computing capacities and facilities, especially when new scientific problems arise.

The starting premise of the model resulted from the combination of different factors, namely new policy tools announced to be implemented by the changing Common Agricultural Policy (*CAP*), the need to account for a wide diversity of farming systems, and the changing geographical area following the *EU* enlargement. In complement of econometric approaches relying on past information, the model introduces new mathematical programming (*MP*) approaches using the most relevant data.

The Farm Accountancy Data Network (*FADN*) is considered as a permanent, robust and representative source of information, useful for *MP* models delineated at the farm level. The information in the *FADN* database consists of surveys of individual farm samples. The database provides individual information mainly accounting of accountancy and financial or economic variables. The *AROPAj* model is also capable of using data on physical variables such as livestock, production, and land. Nevertheless, a lot of *AROPAj* parameters are not directly sourced by raw *FADN* data. For example, the model makes use of other data sources, including expert reports and books, such as livestock feeding. Other sources utilized by some *AROPAj* modules are *inter alia* data concerning land use and land cover (Land Use/Cover Area frame Statistical Survey (*LUCAS*), CORINE Land Cover (*CLC*)), the European Soil Database provided by the European Joint Research Centre (*JRC*), the Digital Elevation Model (*DEM*), and the Climate information provided by the European *MARS JRC* Unit. These variables are grouped by means of a clustering method to form a group of similar farms that present similar characteristics and thus can form a unit of modeling for the *AROPAj* model: the farm type (also denominated by farm group). The overall model is based on farm-groups, designed on a regional scale, for all of the *EU*.

In large part, the *AROPAj* model originates from modeling tools aiming at *CAP* impact assessments. Mathematical programming models are effective in this way, especially when the implementation of new policy tools in economic models has to rely on non-time dependent data, unlike time-inferred econometric models. Due to the wide and complex diversity of *CAP* tools renewed throughout decades, the *CAP* block size quickly increased. The generic modeling approach embraces the wide variety of European farming systems which are subsumed by the continually-growing *EU*. It involves a multi-scale analysis from the farm type level to the *EU* level, accounting for realistic discontinuities at the farm level.

The model is also designed to assess global change impacts and analyze policy while accounting for the complexity of interactions between agriculture and the environment. In this regard, it is a tool used for the analysis of a large range of economic and environmental domains, such as greenhouse gas (*GHG*) emissions and abatement, nitrogen sourced pollutions, biomass provision from perennial crops, ozone impacts and adaptation to climate change. Other collateral outputs may exist when providing farm-based estimates and data related to physical parameters.

The interest of linking the *AROPAj* economic model with biophysical models relies on the possibility of accounting for environmental (physical impacts such as soil and climate) and technical variable impacts on the production functions, as well as on the agricultural system

output relevant to environmental concerns. The general goal is to bring to light the relationship between crop production inputs (e.g. fertilizer applications) and crop production outputs (e.g. yields). The potential relation between the economic and biophysical models is to be found in the management practices of the farmer, such as irrigation, application of pesticides, and fertilizer.

Some refinements are required when the farm group's optimal solution is made of non-marketed crop output (or of non-marketed nitrogen input). In other words, this occurs when, at the optimum, the crop production is entirely re-used on-farm (e.g. for livestock feeding), or when no N-input comes from the market (e.g. manure sourced). In this case, market prices have to be replaced by relevant shadow prices, which are part of the optimal solution. In practice, the problem is solved by iterated *AROPAj* runs with shadow prices replacing the prices used at the previous iteration for the yield function derivative. This process of convergence is fast in most *AROPAj LP*'s, except for a few cases where solutions have periodic forms over iterations, with slight amplitudes.

The model is running using the *GNU* software and applications, except for the marketed optimization solving tool (*GAMS*). Year after year, the choice has been made to promote the use of *GNU* software for *AROPAj* improvement, not only regarding the financial cost, but mainly to facilitate the maintenance and the evolution of our programs and applications. The model also produces some interesting spatially disaggregated outputs in line with multi-scale environmental problems. The computing and algorithm facilities embedded in the modeling process allow new improvements and new model exploitation exercises.

The European agro-economic *AROPAj* model is built of a set of linear programming models, aiming at a good representation of European farming systems in term of geography and farm type diversity as well. As a result, the model aims at being a tool and a reference, which benefits a significant numbers of PhDs and MSc students, and contributes to many published peer-reviewed papers (to which this document refers).

1 Introduction

1.1 Objectives

Key-points:

Policy analysis (agricultural and environmental policies)

Scope (crops, geographical cover)

Agricultural land allocation

Linkage between agriculture and the environment

Mitigation and adaptation of the EU agriculture (to global changes)

New challenges agriculture could face over next years, when no past information leads usual econometric approaches to fail in feeding analysis

The basic premise of the *AROPAj* model is derived in order to suit a combination of different factors: the newly introduced policy tools of the ever-changing Common Agricultural Policy (*CAP*); the need to account for the heterogeneity of farming systems, and the growing geographical area of the European Union (*EU*). The *AROPAj* approach of the *AROPAj* model uses a combination of mathematical programming (*MP*) and statistical techniques. This choice is motivated by the availability of micro-economic and technical data of sufficient quality. The Farm Accountancy Data Network (*FADN*) -"an instrument for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy" [1]- is considered to be a permanent, robust, and representative source of farm level information for *MP* models.

The *FADN* data, combined with technical information, allows one to adequately describe the "production set" of the representative economic agents. Production sets are viewed as sets of technical constraints. Furthermore, the numerous policy instruments, which are currently implemented in the *CAP* framework, as well as those that will be implemented in the future, are easily managed by *MP* modeling. Following conventional approaches, the basic structure of production sets is linear. The modeling of the farming systems by means of technical relationships between agricultural activities has to be informed by behavioural elements, which are to be selected in accordance with micro-economic standards. In our case, farming system decisions result from gross margin optimization. Stated otherwise, the rationale for the *AROPAj* modeling approach is predicated on linear programming (*LP*) techniques. The *AROPAj* model is composed of a set of *LP* models in order to correctly represent the geographical and typical diversity of European farming systems.

The *AROPAj* model is a supply-side model. The entire price system is included in the set of parameters. Of course, the "small country" assumption does not hold when the agriculture of the European Union is considered as a whole. Nevertheless, there exists already a number of well-performing general or partial equilibrium models which can supply us with the price vectors when necessary.

The *AROPAj* model is a mono-periodic short run model. Results are delivered at an annually level. Since the model is static, the estimated results can be interpreted as the long run impacts of economic policy (*CAP*) or physical (climate) changes under the *ceteris paribus* hypothesis:

[1] http://ec.europa.eu/agriculture/rica/

no change in farms structure or in land endowment except through exogenous signals via the parameters. Included in the model, land allocation is joined to (bovine) livestock adjustment.

Because of the linear structure of the model and the necessity of implementing many *CAP* tools, there is a need for binary variables. Particular attention must be paid to the restrictions imposed by the *CAP* measures. Set-aside requirements as well as milk and sugar beet quotas fall into this category. Mandatory and voluntary set-asides are accounted for, each type of set-aside being treated as a productive activity associated with the corresponding payments. The different types of sugar beet quotas (A, B, and C) are also included. Many of the *CAP* policy instruments included in the model involve the use of binary or integer variables whenever producers have to face mutually exclusive "discrete" choices. At this stage, *AROPAj* leaves the general *LP* framework and calls upon Mixed Integer Programming solvers (*MIP*).

After the seminal period devoted to the *CAP*, strong opportunities for *AROPAj* improvement were provided by environmental issues. The first occurred in the middle of the nineties with the emerging problem of climatic change (*CC*) impacts. A greenhouse gas (*GHG*) block has been added, following the methodology developped by the International Panel for Climate Change (*IPCC*), in order to compute agricultural *GHG* emissions (**?**). It makes it possible to assess the impacts of *CAP* reforms on GHG emissions as well as to assess the impact of *GHG* emission mitigation or *CC* adaptation on *CAP* sourced agricultural support.

The interest in agricultural pollution other than that of *GHG* emissions, and the improvement of the agronomic block of the *AROPAj* model lead to a reexamination of the linkage between nitrogen (N) sources and crop yields. The seminal work builds on a stylized plot analysis, based on crop modeling (i.e. the *STICS* crop model use), which has expanded to cover a serie of crops throughout the various European farming systems. This expansion of the model is made possible by the myriad of physical data at the European scale, particularly concerning soil and climate, as well was the digital elevation information (**?**). Mathematically, the implementation of these functions makes the *AROPAj* model non-linear, even when *LP* solvers are used as part of a two-step calculation process (Aghanjanzadeh-Darzi et al., submitted paper).

The introduction of the nitrogen block in the model creates a gap in the applications originating from the *AROPAj* model. Subtle analyses of environmental regulation are provided when dose-response functions, initially devoted to nitrogen-to-yield functions, were expanded to N-pollutant functions (**??**). Another field of application refers to autonomous adaptation of agricultural systems to climate change, when adaptation applies at the plot level via the N-yield functions which are transformed by climate change (**?**).

Another gap arises when probabilistic contribution of *AROPAj* representative farming systems are geographically distributed within a region. This spatialization block is based on spatial econometric methods (developed by **?**). The methodology and applied calculation process are explained by **?**. By using tropospheric ozone (O_3) concentration yield functions as well as spatialization blocks, we are allowed to improve the assessment ofO_3 impact on farming systems of various scales (**?**). The spatialization block also allows us to couple models based on different spatial grids. An example of such a coupling of models is the concomitant uses of the *AROPAj* and hydroecological *MODCOU* models, the former dedicated to agricultural surfact activities and related pollution factors and the latter to water and pollutant transfers to aquifers (Bourgeois et al., submitted).

The last on-going step is to incorporate water as a major production factor, as was done with

nitrogen. Academic papers are on progress and PhD students are currently working on this front.

1.2 State of the art and early future

> Key-points:
>
> What runs
>
> What is in progress

As the *AROPAj* model continuously evolves, some discrepancies may arise among itself various key elements: namely the model itself, its construction, extensions, and applications. These elements are clustered as follows:

1. the core of the model, including different internal blocks added over time

(see also Section **??**),

1. the three steps of the model building, namely:
   (a) typology and sample farm clustering (see also Section **??**),
   (b) parameter estimation (see also Section **??**),
   (c) farm cluster model calibration (see also Section **??**),
2. the dose-response functions dedicated to nitrogen fertilizing and crop yields (see also Sections **??** and **??**),
3. the spatialization processing tools (see also Section **??**),
4. the shell and its user-friendly interface (see also Section **??**).

An overall overview of data and processes involved in and close around the *AROPAj* model is presented in Figure **??**. At the present stage, the "spatialisation" and "crop yield function" blocks are dissociated from the *AROPAj* kernel in the way that each block building is made separately. The computing limitations now being overcome, each block building is scheduled to be refined through interacting with others. More accurately, spatializing *AROPAj* farm groups will be used for improving the yield function selection (through soil characteristics), and yield functions will be included in the *AROPAj* calibrating process.

At the core of the model, we put the design of farming systems and the way crops, pastures and animals interact, as well as their relation to demand for inputs such as mineral fertilizer and feedstuff and to the export of marketed products and non marketed externalities. The basic *LP* structure of *AROPAj* has remained stable and robust for many yearswhile constantly integrating new blocks, which improve the representational quality of the production set and the assessment of agricultural externalities. The integration of new blocks is rather simple and their impact on relative computer run time is minimal. In a word, there is no difficulty in adding parameters and primal or dual variables in the model as long as relevant data exists for parameter estimates for the model.

The computing tools that enable the model to evolve are themself constantly evolving. The goal of this is to make programmers and users as less dependent as possible on the operating system and

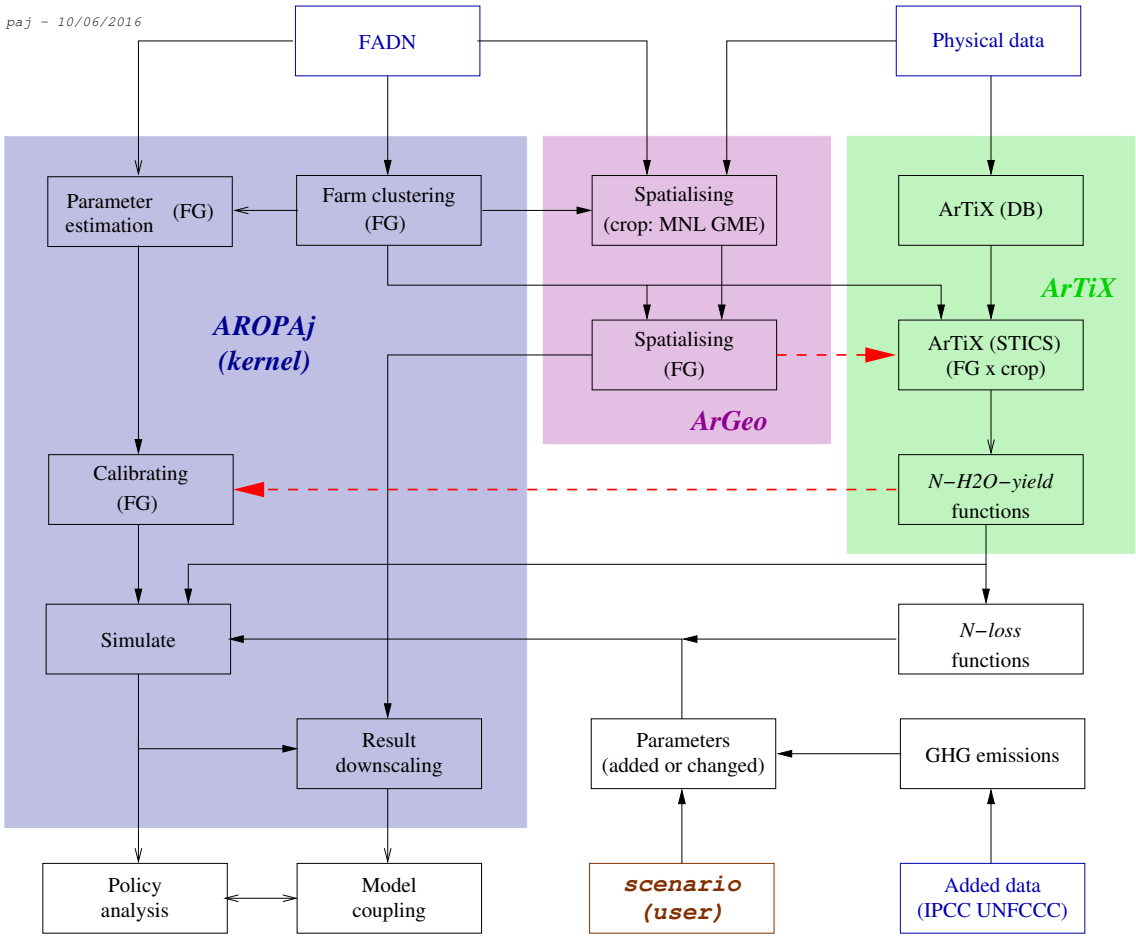


Figure 1. General overview of data and processes associated to the *AROPAj* kernel and its surrounding applications. Red arrows refer to improving the interface between the core model and other processes.

computers that hosts the model. Nevertheless tools operate in *Unix* or *Linux* environment. When it comes to software, things get a little complicated. At the beginning, linear programming tools were based on commercial software, namely *MGG* dedicated to matrix generating and *SCICONIC* as a *LP* solver. Currently we use the *MGG* structure as if it is composed of *FORTRAN* instructions making that now free of license costs. The *AROPAj* code under *MGG-FORTRAN* is automatically converted in *GAMS* ; we use *MP* solvers proposed in the *GAMS* library. Other software contributed to the building of the *AROPAj* model and to external blocks such as *SAS* and *ArcGIS*. These blocks are now replaces by *R* modules (*PostGIS* and *QGIS* are also used for a series of specific tasks and works). The core of the model is designed in such a way that all past applications developped for any version of the model are still operational.

The nitrogen block refers to parameterized yields by default. When *N*-yield functions are available, mineral nitrogen intakes and crop yields become *AROPAj* endogenous variables. These functions result from the *STICS* - *AROPAj* linkage, which is based on a tremendous undertaking involving biophysical databases and database management systems. An earlier cluster of computer networks at the Laboratoire des Sciences du Climat et de l'Environnement (*CEA* - *CNRS* Saclay) was used for this purpose. This project was realized by the efforts of D. Leclère for the *AROPAj-V2* version on the European scale. Our current work (2014-2016), managed at our workstation by P. Humblot, entails extending the model, facilitating database management, undertaking *STICS* calculation runs, and processing results.

In the early 2000's during the EC-FP5 *GENEDEC* period, P. Cantelaube and his colleagues at the Join Research Center in Ispra, Italy, created a means of incorporating calculations related to Geographical Information System (*GIS*) and the crossing of spatial databases.

A. Lungarska and P. Humblot are managing the renewed spatialization block on our workstation with the support of open source softwares (2017-2021). Over the years, colleagues from *INRA-EA* teams (Orléans, Laon, Avignon) contributed to the improvement of the yield function blocks in *AROPAj* in relation with the development of the *STICS* crop model. All these factors strengthen the model's capacity to build the link between the farming system's representation and the geo-referenced soil characteristics, giving us opportunities to enlarge our analysis field in term of agro-environmental problematics.

Table 1. List of *AROPAj versions*.

version	EU cover	*FADN* year	yield functions (MS/number in case)	spatialization (FADN-region number)	*OS* & stat. software
V1 (obsolete)	EU-12	1997			*Unix* & *SAS*
V2	EU-15	2002	(15)	(12)	*Unix* & *SAS*
V3	EU-24*	2004	(EU except Malta)	(specific way)	*Unix* & *SAS*
V4 (for test)	EU-27	2009			*Unix/Linux* & *SAS*
V5	EU-27**	2007		(109/136)	*Linux* & *R*
V5	EU-27	2008	(EU27 except Malta, Cyrus)	(110/135)	*Linux* & *R*
V5	EU-27	2009	(EU27 except Malta, Cyrus)	(111/135)	*Linux* & *R*
V5	EU-27	2010	(EU27 except Malta, Cyrus)	(110/135)	*Linux* & *R*
V5	EU-27	2011	(EU27 except Malta, Cyrus)	(111/135)	*Linux* & *R*
V5	EU-27***	2012	(EU27 except Malta, Cyrus)	(104/134)	*Linux* & *R*
V5	EU-28	2013	(EU28 except Malta, Cyrus)	(130/135)	*Linux* & *R*
V5	EU-28	2014	(EU28 except Malta, Cyrus)	(129/134)	*Linux* & *R*
V5	EU-28	2015	(EU28 except Malta, Cyrus)	(130/135)	*Linux* & *R*
V5	EU-28	2016	(EU28 except Malta, Cyrus)	(130/135)	*Linux* & *R*
V5	EU-28	2017	(EU28 except Malta, Cyrus)	(130/135)	*Linux* & *R*
V5	EU-28	2018	(EU28 except Malta, Cyrus)	(129/135)	*Linux* & *R*

*EU-25, except Malta

**Some regions subdivided compared to next years in Portugal

***The *FADN* covers 3 French overseas Regions, and some Hungarian regions grouped

One *version* of *AROPAj* is animated by the annual *FADN* sample on which farm clustering, parameter estimating, and *LP* calibrating are based. Table **??** lists the various versions which have been or still are in use. The *V1* version is not maintained for operational uses. All versions refer to one or more European research programs, as cited in Section **??**.

In 2016, a "water block" is included in the model under the $V5$ -version. It allows to integrate irrigation water as well as nitrogen fertilizing through 2 -input yield functions. This was first operating for the $V5/2009$ -version of the model applied to France, for which such yield functions were estimated in 2016 (see the seminal paper by **?**). The extension is completed and operational in 2020 for the *FADN* years 2008-2012.

All improvements inside and around the model, including the kernel updating, the spatial downscaling of results at a fine level of resolution (within *FADN* region), and the implementation of crop model based yield functions, are designed and built in a generic way to run the *V5* class of *AROPAj* for all *EU* countries (see the map covering *FADN* regions of the *EU-28*, Figure **??**). The yield functions in "future climate" were first built in line with the *FADN*-years 2009 and 2012 (*IPCC* RCP4.5 scenario, CC-years 2010-2100) in a situation of weak adaptation of production systems (sowing dates).

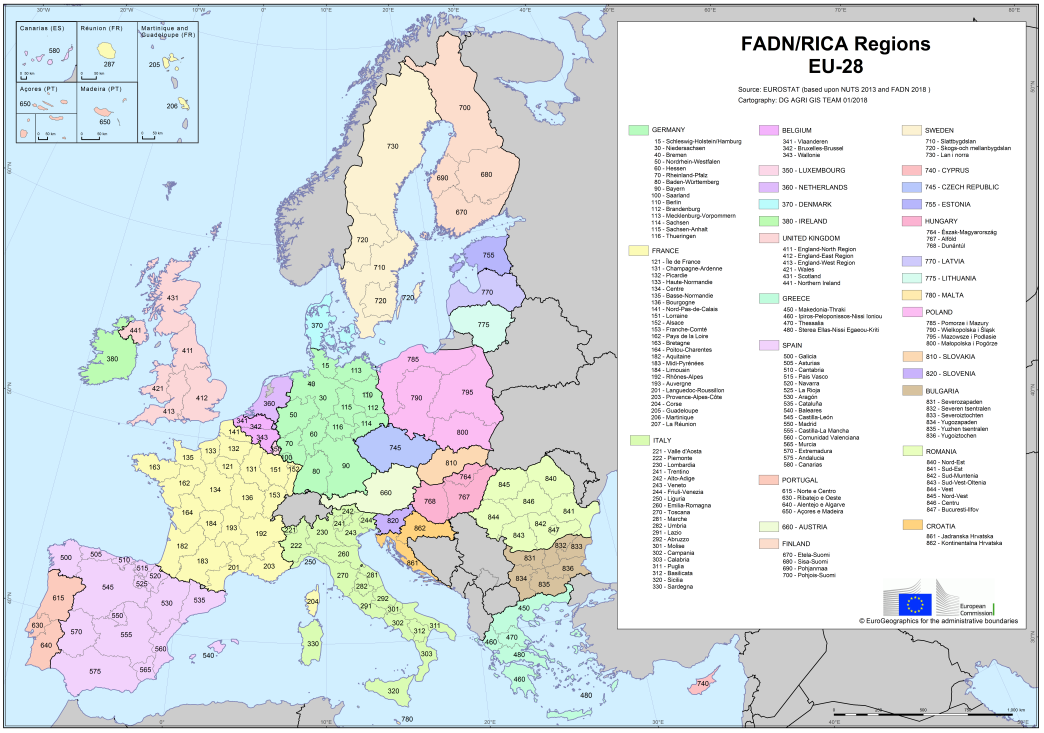


Figure 2. Map of FADN regions - 28 countries of the EU (2018).

Regional delineation and map on line

https://ec.europa.eu/agriculture/rica/regioncodes$_f r.cfm?CodeCountry = EUR$

In 2018, the downscaling step is implemented for the 6 *FADN* years 2007-2012. The coupling of *AROPAj* with the crop model is completed and operational for different countries and for 5 of the 6 *FADN* years, 2008-2012.

In 2020-2022, the *AROPAj* calibration is realized for the 6 *FADN* years 2013-2018. Downscaling

is realized for *FADN* years 2013-2018, including Croatia. Re-calibration of yield functions is carried out in 2022-23 for *FADN* years 2012-2018. The yield functions related to "future climate" (*IPCC* RCP4.5 scenario) have to be renewed, after adjustments in physical parameters and after reshaping of yield functions in "present time". This function building is realized for 3 *GCM*-years (2016, 2043 and 2073 as considered of statiscally representative of three-decade periods, respectively 2011-2040, 2140-2070, 2071-2100) in a context of climate change threw a RCP4.5 model scenario (based on the *GCM* IPSL2014-CM5A-WRF [2]) with strong adaptation of crops (including changes in crop cultivars and technical routes).

New projects involving *AROPAj* in a climate change context will refer to greater warming under RCP8.5 scenarios. This means building new yield-functionslp for "future climate" years.

In terms of technical modules, the AROPAj model will soon benefit from the inclusion of the transformation of organic matter of plant and animal origin into biogas.

[2] http://www.drias-climat.fr/document/Doc_DRIAS_database_IPSL2014-IPSL-CM4_WRF.pdf

2 *AROPAj* among other agricultural economic models

2.1 MP models vs macro-econometric models

> Key-point:
>
> Completeness over competition

Compared to the econometric methods that are used for policy analyses in the agricultural sector, *MP* models offer a number of advantages. The most important difference is arguably that there are significantly fewer data requirements, allowing *MP* models to be built with just one observation (data from a single year). Econometric models, on the contrary, require large amounts of time series or panel data in order to estimate production relationships; they relate varied input to the observed output on the basis of an assumed model. Since their estimates depend on the data set that was used for estimation, they are only valid for the reference period corresponding to this data. This means that they cannot be used for economic analyses in cases of drastic policy changes, unlike *MP* models, which have the ability to examine different production methods and policy contexts as well as to analyze the interrelation between the different resources and production activities.

Econometric models are considered "descriptive" (positive) since they seek to answer the "why" and "how" questions, that is, to what extent, and how certain observable economic factors can affect a specific economic phenomenon (e.g. production). On the other hand, *MP* models are normative (determenistic) in the sense that they try to identify the optimal solution to a specific economic problem, under the assumption that economic agents behave "rationally", i.e., they optimize their utility or welfare, under a set of constraints related to resource endowments. This framework of mathematical formulation for economic analysis presents the additional advantage that it allows the estimation of shadow prices for all scarce resources, which is important when there is no market or when their social value is not reflected in market prices.

2.2 One side of the market in agricultural models

> Key-points:
>
> Fine *LP*s aggregation regarding one side of the market
>
> Wide opportunities for demand - supply coupling

Two inter-connected concerns are often raised by the use of *LP* one-market side models. Linearity in itself and linear programing by extension may be negatively perceived when one desires smooth functions that link prices and quantities (i.e. when there is a preference for many order differential functions). Clearly, usual *LP* corner solutions and stepwise result functions do not meet this expectation. At the same time, reality is rarely, if ever, amenable to representation by smooth functions. At the farm level, for instance, there is no doubt that discontinuous land sharing changes when prices or *CAP* tools change. This is why linearity should not be rejected *per se*. The *LP* design of farm level models is discussed hereafter (see the **??**).

Using the *LP* option at the farm level does not preclude results with smooth curves at the sectoral level. Clearly, the aggregation of *LP* models will rarely yield a "linear" result. But, there remains the question of aggregation bias, which was raised early on by **?**, and subject to relaxed condition discussions that (**?**) dealt with among others. This is a typical problem of production sector representation when the sector is composed of numerous economic agents. Since computing tools allow us to extend the limits of modeling capabilities,the real limit is data availability and the diversity of surveyed farm systems (see Sections **??** related to *FADN* and Section **??** related to parameter estimation). The diversity of farm systems covered by the *AROPAj* model is addressed in Section **??**.

A thorough representation of one side of the market by a model leads modelers and model users to seek information on price feedbacks. Undoubtedly, the "small country" assumption has to be rejected when the model aims at covering tje agriculture of the entire *EU*. Consequently, we can expect that any change in European agricultural supply would impact world and domestic prices. Scientific challenges encountered here are of different categories. First of all, adding a demand side model in the modeling framework or including the supply side model in an existing partial or general equilibrium model would be questionable in terms of modeling cost-benefit analysis. The demand system may be of high complexity comparable to the supply side, and using a simple demand elasticities matrix might not meet basic requirement of well balanced sides of the markets. More importantly, and paradoxically, possibly incorrectly calculated price feedback impacts may hide what is also of interest, namely direct impacts of environmental or economic change on the agricultural sector *ceteris paribus*. Some attempts to account for clearing market price impacts were undertaken with the linking of *AROPAj* to partial equilibrium models (e.g. *PEATSim* as illustrated in Chapter 3 of (**?**), and *ESIM*, as illustrated by (**?**)).

Another application of market clearing calculus is based the *MODANI* framework. The *MODANI* model incorporates the agricultural supply side model and a feed industry model developed by Olivier Lapierre (AgroParisTech) in the nineties. Raw matters supplied by *AROPAj* farmers are used by the feed industry, meanwhile concentrates produced by the latter are used by *AROPAj* breeders. More than 40 markets (raw matters and concentrates) are simultaneously taken into account (**?**,**?**). An adaptive process was developed for market clearing computation, as illustrated by Figure **??**.

The need for coupling supply and demand remains challenging, yet useful in many cases. This is not only a question of prices, but above all, a question of quantity limits. *AROPAj* has yet experienced partial equilibrium approach with spatial and industrial capacity limits. The development of a three stage modeling chain was an attempt to assess European farmers' capability to meet the target set by the *EC* Directive on the use of bioenergy (see Section**??**).

2.3 LP vs NLP models

Key-points:

Basic linear structure

Linear basis does not prevent from non-linearity inclusion

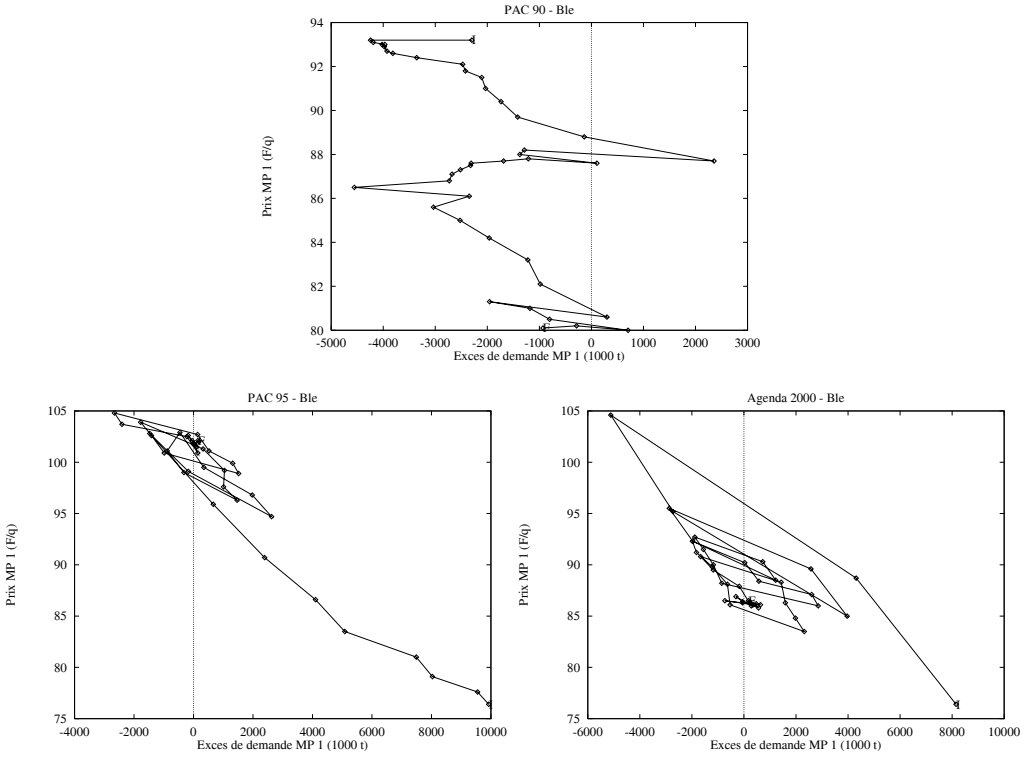


Figure 3. Looking for French market clearing approximated by the *MODANI* model: the case of one raw market ("mp4" refers to soft wheat) up to 40 iterations.

Since *LP* models have been traditionally used for farm planning, moving to aggregate geographical levels pushes the limits and renders questionable the ability to adequately simulate production technologies. The Leontieff technology implied by *LP* is regarded as an oversimplification that can only be justified by a lack of data necessitated by more appropriate profit functions (**?**). Clearly, developing models that do not correctly represent the technology used, means that model outcomes do not necessarily reflect reality and cannot be considered as valid analytical tools (**?**). The validation of the model is thus one of the most important steps in the process of model building. *MP* models are commonly validated with the production test, which entails examining how well the model is able to replicate the observed production levels (**?**). Given that most constraints that are used for farm models cannot be formulated at the regional level, the problem of overspecialization (or corner solutions) is of great concern for *LP* models.

In order to overcome the limitations of the *LP* framework, the literature suggests the use of nonlinear objective functions with which interior solutions can be achieved. It should be noted that nonlinearities in optimization models were not sought initially for calibration reasonsbut rather for their ability to answer specific questions of economic research, such as price formation in agricultural markets and farmers' risk considerations (such models will be presented in the next paragraphs), for which *LP* models are inadequate. Nevertheless, their ability to provide optimal solutions that are not limited to the edges of the constraint set (corner solutions) is supposed to increase the validity of the model.

Probably the most widely used type of nonlinear programming (*NLP*) models built for policy analysis are those which are based on the principles of Positive Mathematical Programming (*PMP*). *PMP* was formalized by **?** and relies on the assumption that there is some marginal implicit information that needs to be captured in order to calibrate the model. The traditional *PMP* method uses a three-step algorithm, beginning with a model that includes calibration constraints, whose *shadow prices* are supposed to contain all information not observed by the analyst. In the following steps, the *PMP* algorithm introduces (or re-estimates) nonlinear terms in the objective function using the previously estimated implicit information, so that the model's first order conditions are satisfied exactly at the observed production level. The main advantage of *PMP* is its ability to exactly calibrate any type of optimization model –typically *LP* ones– by using minimal data, such as a total land constraint. Despite criticisms (e.g. **?**), *PMP* has continued to be used in numerous applications, including major European research projects such as *CAPRI* and *EUROTOOLS* with the aim to analyze the effects of the continuously changing European *CAP* regime.

NLP models can be built using simulated partial equilibrium conditions by finding the optimal price-quantity correspondance for clearing the market of one or more activities and/or factors of production (e.g. **?**). Since such models (also called "price-endogenous" models) consider both the supply and the demand side of inputs and outputs, they relax the basic assumption of traditional *LP* models that farmers are price takers. In this sense, such a model no longer maximizes aggregate profits but instead returns the total producer and/or consumer surplus at the point of market equilibrium. Price-endogenous models are intrinsically nonlinear since the price of one or more activities is defined as a function of the aggregate quantity produced. The main difficulty with such models is that they require a large amount of information that is rather difficult to find, namely pre-estimated demand and supply elasticity values. Although they can offer important insights on how the global agri-food system works, partial equilibrium models are usually highly

aggregated representations of agricultural markets and cannot easily capture the heterogeneity of agricultural production within a specific region. Examples of such models are the *IMPACT* model from *IFPRI* and the *GLOBIOM* model from *IIASA*.

Mean-Variance (E-V) models are another type of nonlinear MP models used for policy analysis. E-V models, building on the pioneering work of **?**, are able to capture the effect of uncertainty and risk on farmers' decision making process. This is achieved with a nonlinear objective function which maximizes the expected utility (EU) of farm profit (or more generally, wealth). The most commonly used specification of an E-V model is the linear one, which was first presented by **?** and which is expressed as: $\mathrm{EU} = \mu - \frac{1}{2}\varphi\sigma^2$, where μ is the mean farm profit, φ is the risk aversion coefficient, and σ^2 the variance of farm profit. A serious shortcoming of E-V models is that they are difficult to build when a large number of activities is considered because the variance term is typically based on time series of prices and/or yields, which are rarely available at the farm or regional level.[3] This may lead to a biased representation of uncertainty faced by the farmers. Another problem with linear E-V models is that the risk aversion coefficient is not easy to estimate. Nevertheless, modelers usually solve their model for different values of φ and select the one which yields an optimal solution that best approximates the observed production plan. As a result of this practice, a level of calibration is also achieved.

[3] In sector models that include representations of both the output supply and demand side, the initial source of uncertainty is the variation of yields, since the final market price is endogenously calculated by the model.

3 Input data

When working with a model, it is essential to carefully choose data. For the implementation of *AROPAj*, the researching and update of data required a significant effort. Indeed, the data used are regularly changed and every year it appears new. Therefore we must constantly search for updates and ensure that our data sets remain up to date and relevant. Moreover, the model *AROPAj* model necessitates the processing of collected data into a specific format (see **??**).

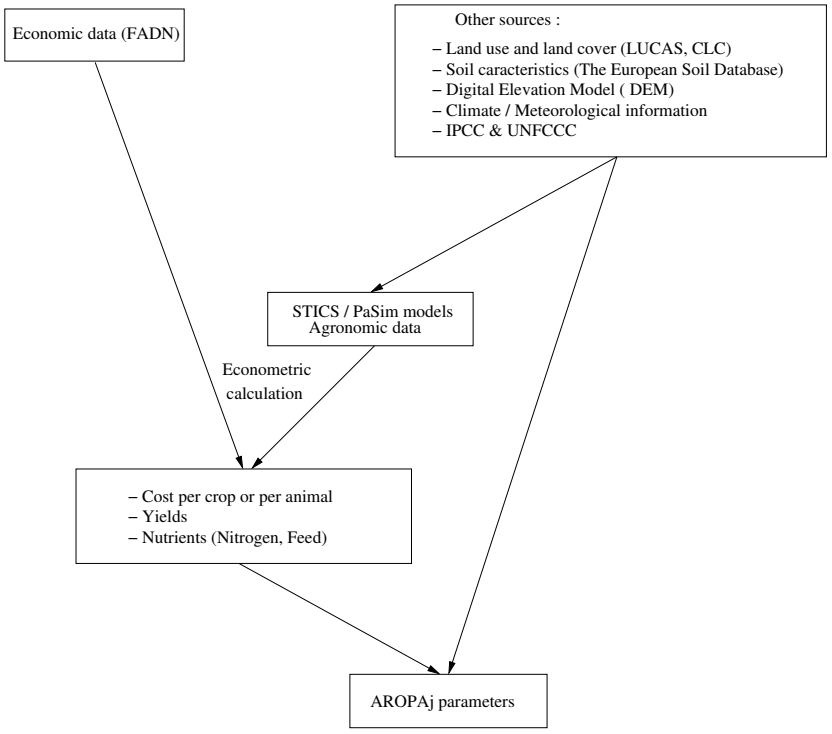


Figure 4. General illustrative scheme of data management.

3.1 The *FADN* at the core of the farming system representation

Key-points:

Data as major limiting factor for economic analysis
A chance with the *FADN* to relax this constraint
Associating micro-economic and bio-physical data

The *FADN* is the main data source of the *AROPAj* model, since it statistically represents farming system heterogeneity on the regional and covers the entire European Union. The various versions of the *AROPAj* model rely on different annual *FADN* samples with respect to restricted conditions, e.g. no *FADN* sourced estimate should be based on less than 15 sample farms. Description and characteristics of the *FADN* are of free access on-line: `http://ec.europa.eu/agriculture/rica/`. Researchers must sign an individual form each time they use raw data from

the European Commission. Data access is permitted by the European framework programmes for research.

As a partner or as the cordinator (*GENEDEC*), the *INRA-Economie Publique* unit was granted access to the *FADN* data for the following research programmes:

- FP4-FAIR *EUROTOOLS* (*EU*-12): *FADN* - 1997: "Tools for evaluating EU agricultual policy at different decision levels"
  `http://cordis.europa.eu/project/rcn/48045_en.html`
- FP6 *GENEDEC* (*EU*-15 and *EU*-25): *FADN* - 2002, 2004: "A quantitative and qualitative assessment of the socio-economic and environmental impacts of decoupling of direct payments on agricultural production, markets and land use in the *EU*"
  `http://cordis.europa.eu/result/rcn/49674_de.html`
- FP7 *AnimalChange* (*EU*-27): *FADN* - 2007-2009: "An Integration of Mitigation and Adaptation Options for Sustainable Livestock Production under Climate Change" `http://www.animalchange.eu/`
- FP7 *FoodSecure* (*EU*-27): *FADN* - 2007-2011: "An interdisciplinary research project to explore the future of global food and nutrition security" (in progress)
  `http://www.foodsecure.eu/`

Other European programmes referred to *FADN* and *AROPAj*, i.e. *INSEA* (FP6, "Integrated Sink Enhancement Assessment" `http://esdac.jrc.ec.europa.eu/projects/INSEA/`), *CCTAME* (FP7, "Climate Change – Terrestrial Adaptation & Mitigation in Europe" `http://www.cctame.eu/`).

The *FADN* database provides farm-level information mainly in terms of accountancy and financial or economic variables. The *FADN* database also offers information regarding physical content such as areas, livestock, and products. Nevertheless, much of the *AROPAj* parameters are not directly sourced by raw *FADN* data. For instance, the model's need for costs per crop or per animal necessitates the use of econometric calculation. In addition, the techno-economic character of the *AROPAj* model demands that information be gathered from other data sources, including expert reports and books (among them attention is paid to livestock feeding).

The need for external information is even more critical for the technical parts of the *AROPAj* modeling chain, namely the "dose-response functions" (see the section devoted to *N*-yield function building), and the geographical probability distribution of *AROPAj* representative farming systems (see the section devoted to spatial econometrics and farm group spatialization). In any case we need *FADN* data, either for micro-economic estimates or for economic and physical data coupling.

3.2 Other sources required by some *AROPAj* modules

> Key-points:
>
> Associating micro-economic and bio-physical data
> Physical data related to *CLC*, *LUCAS*, *ESB*, climate, *DEM*
> Using the *IPCC* guidelines for comparison with *UNFCCC* data

Numerous sources have been used for the spatialization procedure.

3.2.1 Data concerning land use and land cover

1. Land Use/Cover Area frame Statistical Survey (*LUCAS*)

   Two *LUCAS* data sets (2001/2003 and 2012) are used for the spatialization procedures of *V2* and *V5* of *AROPAj*. *LUCAS* survey provides information concerning the observation of different crops at different points in space. This information serves as a regressant in a multinomial logit model. The database accounts also for irrigating practices when observed. We aggregate this information at the scale of a 10 km grid in order to integrate it into the subsequent steps of the spatialization.

   Source:

   `http://epp.eurostat.ec.europa.eu/portal/page/portal/lucas/introduction`

2. CORINE Land Cover (*CLC*)

   Information on land cover at a fine scale (100 m) is provided in the *CLC* database. Land cover is presented in 44 classes. In the first version of the spatialization procedure, we use numerous classes which are often used in the agriculture covered by the *FADN*. In the latest version of the spatialization, the procedure only covers activities involved in *AROPAj*.

   Sources:

   `http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-100-m-version-1`

   `http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-3`

3.2.2 The European Soil Database (*ESDB* v2.0)

The European Soil Database and the Pedotransfer Rules Database are provided by the European Joint Research Centre (*JRC*).

Source:

`http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_Data_Distribution/ESDB_data.html` (Registration required)

3.2.3 Digital Elevation Model (*DEM*)

Geographical information concerning altitude is obtained from the United States Geographical Survey's model *GTOPO30* at the scale of appoximately 1km (30 arc seconds). Slope is then evaluated.

Source:

`https://lta.cr.usgs.gov/GTOPO30`

3.2.4 Climate

Climate information is provided by the European Union Joint Research Centre *MARS* Unit. The *AGRI4CAST* Interpolated Meteorological Data is available at the scale of a 25 km grid and covers weather variables since 1975.

Source:

`http://mars.jrc.ec.europa.eu/mars/About-us/AGRI4CAST/Data-distribution/AGRI4CAST-Interpolate` (Registration required)

3.2.5 *IPCC* guidelines

The *IPCC* parameters for the *EU* Member States can be found in their National Reports of *GHG* Inventories, more specifically in their Tables 4 "Sectoral Report for Agriculture". These are submitted to the United Nations Framework Convention on Climate Change (*UNFCCC*), whose objective is to stabilize greenhouse gas (*GHG*) concentrations in the atmosphere at a level that would prevent and reduce dangerous anthropogenic interference with the climate system. Achieving this objective is dependent on an accurate knowledge of *GHG* emissions trends and on our collective ability to alter these trends.

The use of *IPCC* guidelines is detailed in **??**.

3.3 Data related to other models

> Key-point:
>
> Added external data required for model coupling and applications

Due to the wide variety of models employed with *AROPAj*, a wide variety of data is mobilized in the modeling chain. Consistency of data involved in the various parts has to be checked. At the same time, each of the different models is developed independently. In some cases, climate, soil and farm management data are not from the same sources nor of the same scale, time period, and horizon. The reader has to keep this in mind, not only for discussing model assumptions but also for discussing scenarios which govern *AROPAj* based simulations.

See Section **??** for more information.

3.3.1 *STICS*

In order to run at the farm-type level, the *STICS* modelhad to make the most of data from existing databases (see Section **??** above for details): European Soil Database (*ESDB*), Digital elevation model (*DEM*), Climate database from *MARS JRC* Unit, and *FADN*. To infer *STICS* input data, a special effort was made to incorporate agricultural practices, as they are key variables for the modeling of farmers' actions and of the link between their economic behavio (modelled by *AROPAj*) and the biophysical functioning of a field (modelled by *STICS*). The expert statements that determine all the *STICS* inputs are detailed in **?**. Through the linkage with the *STICS* model, *AROPAj* is provided with crop N-response curves (for yield and N-losses to the environment), thereby integrating the pedoclimatic conditions of each European farm-type. Moreover, thanks to this linkage, equivalence coefficients of N fertilizer were provided for manure-sourced N.

3.3.2 *PaSim*

The *PaSim* model feeds *AROPAj* through an external way, in the sense that there is no feedback between the two models, unlike the case of the *AROPAj-STICS* linkage (see Section **??**).

Running the *PaSim* model calls for basic settings of vegetation type settings such as intensively managed productive grassland (rye-grass/white clover mixture) as well as different input combi-

nations including: [4] (i) three climate scenarios, in line with the work of **?** (see also **?**) with a 0.5° x 0.5° spatial resolution; (ii) soil properties database, in which medium soil, with the 0.25 deg mean percent sand/silt/clay/organic/gravel, has been extracted from the Harmonized World Soil Database (source: the *US* Department of Agriculture); (iii) three types of nitrogen management generated by the *CAPRI EU*-wide agricultural sector model, which are region dependent, ranging from 0 to 0.35 t N/ha.

[4] Aghajanzadeh-Darzi, P., Laperche, S., Martin, R., Jayet, P.A., (2015), Vulnerability of grassland systems to future climate variability: Impact assessment across the European Union, submitted.

4 Architecture of the modelling framework

The structure of the *AROPAj* model and its last developments have been briefly presented in several papers (**????**). We focus here on the *LP* model core, delineated in variables, functions or routines, and parameters. In this section, we aim at describing the production set and the objective function related to each economic unit's behavior. We provide comprehensive lists of activities and parameters covering all eligible items for all economic units. As explained below, we use ''flags" when certain activities and parameters are not relevant for some units.

We consider that each economic unit refers to a farm group, implicitly viewed as a cluster of homogeneous farming systems. In doing so, the overall model is based on farm-groups, designed at the region scale, for all *EU* regions. An *AROPAj* region refers to a regional *FADN* unit. Any *AROPAj* version refers to a series of farm groups accurately defined by clustering techniques within each region.

The basic structure of the modelling framework uses linear programming (LP), popular over the years in agricultural economics. Let us consider the problem of farming systems that attempt to optimize their gross margins when their activities are subject to economic and technical constraints.

The model consists of a set of independent, mixed integer and linear-programming models. Each model describes the annual supply choice of a given 'farm-group' (denoted by k), representative of the behaviour of "real" farmers. The farm-group representation makes it possible to account for the wide diversity of technical constraints faced by European farmers. Each farm-group k is assumed to choose the supply level and the input demand (x_k) that maximizes total gross margin (π_k). In its most general expression, the generic model for farm-group k can be written as follows:

$$\begin{aligned} \max_{x_k} \quad & \pi_k(x_k) = \max \; g_k.x_k \\ & s.t. \quad A_k.x_k \leq z_k \qquad (\lambda_k) \\ & \qquad x_k \geq 0 \qquad (\mu_k) \end{aligned}$$

where x_k is the n-dimension vector of producing activities for farm-group k, and g_k is their n-dimension vector of valuations in the gross margin. The matrix A_k is of $m \times n$ dimensions and refers to input-output coefficients. The z_k m-dimension vector refers to right-hand side (RHS) capacities. Together, A_k and z_k define m constraints faced by farm-group k. We set λ_k as the m-dimension vector of associated Lagrange multipliers. Without loss of generality, the x_k vector stands for positive quantities, and the associated Lagrange vector is denoted by μ_k. The x_k activity vector has to be understood in a wide sense, knowing that the LP model may require additional intermediate variables for which objective valuations are usually null.

The components of x_k include, among others, the area and the production devoted to each crop, animal heads related to each animal category, milk and meat production as well as the quantity of purchased animal feed. The gross margin, g_k, contains series of elements corresponding to each production activity, which, for crops, gives per-ton revenue (price, when the production is partly or totally sold) and, when relevant, per-hectare support received, minus per-hectare variable costs. As the emphasis is on the farm-group level, each farm-group is assumed to be a price-taker. Thirty-two crop producing activities are permitted in the model and represent most of the European agricultural land related to arable land and pasture. Crop production can be sold at the market

price or used for animal feeding purposes (feed grains, forage, and pastures). As for livestock, thirty-one animal categories are represented in the model (27 for cattle plus one each for sheep, goats, swine, and poultry). Cattle categories depend on the age, the gender, the origin (on-farm or bought animals), the final output (dairy or meat), and *CAP* subsidies.

A comprehensive list of crops and animal categories included in the model is provided on Table **??**. This table also provides related index values (i for animals and j for crops) used in the "matrix generator" (see the `aro.mgg` file to which it is refereed hereafter), and the 2-character strings relating these index values to parameters in the data file `mgg.par`. Among other basic elements of the model, concentrated feed are of 4 categories referring to the n -index and to the 2-character strings `a1-a4` respectively denoting simple energy concentrate, simple protein concentrate, energy concentrated compound and protein concentrated compound. Parameter names referring to bought raw feed use the `ag` string.

The technically feasible production set is bounded by the constraints defined by A_k and z_k. Basic blocks of constraints include:

1. crop rotations and crop use (split in marketed output and on-farm consumption),
2. animal feeding,
3. internal relationship between bovine categories and livestock number limits (including key livestock adjustment allowing livestock to vary in a limited range (minus and plus $\pm 15\backslash$numbers in the corresponding animal categories),[5])
4. specific environmental balances and limits,
5. *CAP* instruments implementation,
6. resource endowments (land, quasi-fix livestock, quotas),
7. non previously affected constraints (denominated by "general").

All *LP* coefficients, i.e. valuation terms (g_k), matrix elements (A_k), and endowments (z_k) depend on parameters. Some of them relate to farm groups ; they are denoted by θ_k. Others refer to constant items or to common economic environment, e.g. related to the *CAP* ; they are denoted by ϕ. Any *LP* coefficient may be viewed like a function, i.e. $g_k(\theta_k, \phi)$, $A_{k_{mn}}(\theta_k, \phi)$, $z_k(\theta_k, \phi)$. These functions may be non linear, perhaps even non-analytic. A key point of the modeling chain is to construct the matrix generator, which allows for the construction of the *LP* and the opportunity to feed it with parameter estimates.

4.1 Overall technical structure of the model

Key-points:

Modularity, genericity

Nested partitioning: Farm group $\subset$ region $\subset$ *MS* $\subset$ *EU*

Key items: parameters, primal and dual variables

The model is built in such a way as that it facilitates its evolution and the addition of new blocks in accordance with new research problematics. That includes changes in farm group clustering and

[5]The discretionary level of the range is usually adopted as close to the yearly basic period of the AROPAj model divided by the average cow lifespan.

the introduction of new farm groups (e.g. in case of the *EU* enlargement). Other cases of model evolution are new environmental issues (nitrogen losses and *GHG* emissions), and refining of the production set in terms of input-output interaction (e.g. nitrogen to yield functions for crops). This requires a comprehensive parametrization of the model, and an extensible range of indexes.

Following notations mentioned above, parameters θ_k for all farm groups $1...k...K$ and parameters ϕ represent the comprehensive set of *AROPAj* parameters. The objective vector (g_k also denoted by *OBJ*), the "right hand side" vector (z_k also denoted by *RHS*), and the matrix A_k are composed of elements which all refer to these parameters.

The key items facilitating the presentation of the model structure are:

- *index*: affects *parameters*, *variables*, *functions*, *conditions*
- *parameter*: refers to data; their values are raw data used in the model through matrix *elements*; the way in which raw data is used by the model is described in Section **??**
- *variable*: compound of the *LP* solution
  - *primal*: related to an activity (e.g. a command variable of the *LP*); regarding the *LP* structure, primal variables are "columns", and *RHS* is viewed as one of them
  - *dual*: related to a constraint and labeled by constraint name; regarding the *LP* structure, dual variables are "rows", and *OBJ* is viewed as one of them
- *element*: intersects a row and a column, is a (possibly non-linear, even non-analytic) function of *parameters* and *indexes*
- *function*: of which variables are indexes, parameters and external data
- *condition*: may enter a "flag" activating one variable or a subset of variables
- *block*: a subset of activities and constraints characterizing either a part of the modeled farming system, or a policy (agricultural or environmental), or a category of environmental impacts; *LP* disaggregating into blocks is arbitrary but useful for model description; formally, blocks are defined by rectangle subsets of matrix elements; variables will be delineated in such a way that they are associated with only one block, allowing us to compose a partition of the *LP* matrix

4.2 Model blocks in detail

Key-points:

Up-to-date lists in line with the current running version

Comprehensive list built by a script applying on the kernel of the model

Below we display the comprehensive list of parameters and primal or dual variables characterizing the *AROPAj* model. Functions linking parameters and matrix elements, and "flag conditions" are not presented, for the sake of time and space. More importantly we display the date of list inception, which refers to the activation date of *FORTRAN* and *Linux* shell scripts which build the lists from the *AROPAj* core script.

Blocks necessarily interact, leading to overlapping activities and constraints. In the lists below, and in accordance with the *LP* matrix partitioning mentioned above, any parameter or any variable

is affected to one block, even if a lot of parameters (as well as activities) act in more than one block. The scheme **??** illustrates this construction. Only non-null elements are illustrated and listed. Some blocks are more easily delineated than others, in the sense that activities or constraints may play a predominant role in more than one block. And some parameters or variables are of common use without specific characterization. In this later case, items are assigned to the general block $\boldsymbol{G}$. Other blocks are "animal outputs and intermediate ratios for bovine" ($\boldsymbol{A}$), "crop rotations and outputs" ($\boldsymbol{C}$), "environmental impacts" ($\boldsymbol{E}$), "animal feeding" ($\boldsymbol{F}$), "nitrogen balance for crops and nitrogen sources" ($\boldsymbol{N}$), and "policies" ($\boldsymbol{P}$) including the *CAP*.

An additional block, $\boldsymbol{T}$, refers to "trans-farm group" constraints, which is acceptable thanks to performing software and workstation. This block does not generate additional primal activities in the *LP*, as long as constraints involve farm groups' activities subject to global bounds. These bounds appears in the model parametrization. This block is presently in progress.

	X_1	X_2	...	X_n	...	X_N		RHS
OBJ	G_1	G_2	...	G_n	...	G_N		
C_1	E_{11}	E_{12}	...	E_{1n}	...	E_{1N}	$\leq$	Z_1
C_2	E_{21}	E_{22}	...	E_{2n}	...	E_{2N}	$\leq$	Z_2
...	...	...	...	...	...	...	...	...
C_m	E_{m1}		...	E_{mn}	...	E_{mN}	$\leq$	Z_m
...	...	...	...	...	...	...	...	...
C_M	E_{M1}		...	E_{Mn}	...	E_{MN}	$\leq$	Z_M

Table 2. The *LP* matrix structure: a block refers to sub-matrix and related labels, objective and *RHS* compounds; X_n and C_m are respectively labels of activities and constraints; G_n, E_{mn}, and Z_m are matrix elements.

Inception date of the following item lists : Tue, 10 Jan 2023 17:28:13 +0100

4.2.1 Indices and parameters

A large part of parameters and primal or dual variables refers to one index or more:

k : *farm group index*
l : *farm sub-group index*
j : *crop index*
i : *animal index*
m : *constraint type index*
n : *constraint type index*
o : *index for organic source (effluent & urban waste)*

The comprehensive list of basic eligible parameters is split into the 7 blocks delimited above (i.e. $\boldsymbol{G}$, $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{E}$, $\boldsymbol{F}$, $\boldsymbol{N}$, $\boldsymbol{P}$, $\boldsymbol{T}$, see this section **??**).

- **the general block *G***

aik(i,k) : *purchase price animal (k€/head)*
cact : *coeff. / cost multiplier (1)*
cfxkk(k) : *fix cost (K€/farm)*
ctdie : *cont. of energy in "other variable costs"*
ctine : *cont. of energy in fertilizer making*
ejj(j) : *world price crop j (k€/t)*
ejk(j,k) : *world price or non guaranteed price (farm group) (k€/t)*
gjk(j,k) : *fertilizing costs (k€/ha)*
gjk0(j,k) : *preserved ferti. costs in case of recalculation (k€/ha)*
igene : *identifier of balances combination selected by the user*
ijb(j) : *identifier for BOUND LOWER introduction / variable J*
itest(k) : *identifier of policies combination selected by the user*
penrj(j,k) : *price of by-product related to crop j (k€/tDM)*
pii(i) : *relative change in animal price (yik et aik) (1)*
pjj(j) : *relative change of j-crop price (1)*
pjk(j,k) : *domestic price or guaranteed price / crop j (k€/t)*
popkk(k) : *farm group number k (1000)*
popul(k,l) : *sub-group farm number (by defaut l=1 popul=popkk (1000)*
rojk(j,k) : *shadow price of j-crop product (in case of endog.) (k€/t)*
pxenr : *energy relative price deviation*
saukk(k) : *UAA (farm group) (ha/farm)*
saukl(k,l) : *UAA (farm sub-group) (ha/farm)*
trans : *corr. transmission coeff. cereal price / feeding (1)*
uii(i) : *FADN sourced conversion coeff. in LU equiv. (LSU/head)*
vjk(j,k) : *other variable costs than fertilizers (k€/ha)*
yik(i,k) : *raw prod. / LS i-animal (not milk, eggs) (k€/head)*
yjj(j) : *price deviation / EU price (>0)*
yjk(j,k) : *price deviation within cereals (>0)*
xjk0(j,k) : *pre-calculated areas [intr. D.Leclere 22 Aug 2011] (ha)*
wokok : *flag for work estimation (>0) (-)*
wokgv(k) : *wages paid per worker (kE/AWU)*
wokup(k) : *unpaid work (AWU)*
wokut(k) : *unpaid work (yearly duration) (h/y)*
wokvp(k) : *paid work (AWU)*
wokvt(k) : *paid work (yearly duration) (h/y)*
wokwg(k) : *wages paid (kE)*
wokwp(k) : *unpaid + paid work (AWU)*
wokwt(k) : *total work (unpaid + paid - yearly duration) (h/y)*

Let us focus on two parameters: *itest* and *igene*. The *itest* k-parameter is dedicated to policies, and the *igene* parameter is dedicated to balance calculation (e.g. animal feed category sums). They refer to the case when certain sub-blocks are activated, such as specific CAP tools (*itest*) or specific environmental balances (e.g. GHG emissions, or nitrogen pollutants). This offers a wide range of combinations, since these parameters are the sum of 2 to i powers, i.e. $\sum_{i \in \mathscr{I}} 2^i$ when $\mathscr{I}$ refers to the user-selected compounds. We should mention that this allows up to 2^{31} possibilities.

Some other parameters may be of a binary type (0/1), e.g. *endrg* which activates the calculation of nitrogen intakes and crop yields when N to yield functions are available.

- **the animal block *A***

dik(i,k) : *animal lifetime i (equiv years time) (yr/head)*
dstkk(k) : *livestock adjustment coeff. (exponen. adjust.) (0<-)*
hik(i,k) : *animal number (head/farm)*
plaic : *relative change of milk price (1)*
plait(k) : *milk price (€/l)*
renf1(k) : *renewal rate for 1-yr bovine female (head/head)*
renf2(k) : *renewal rate for 2-yr bovine female (head/head)*
renm1(k) : *renewal rate for 1-yr bovine male (head/head)*
renm2(k) : *renewal rate for 2-yr bovine male (head/head)*
repro(k) : *bovine reproduction ratio (head/head)*
rlait(k) : *cow milk yield (kl/VL)*
xii(i) : *relat. prod. coeff. milk, eggs (other than bovine) (1)*
xik(i,k) : *prod. milk, eggs (other than bovine) (k€/head)*

- **the crop block *C***

bcekk(k) : *intercept in cereal / crops econometr. rel. (ha/farm)*
bockk(k) : *intercept in oils / cereals econometr. rel. (ha/farm)*
bovkk(k) : *intercept in oils / crops econometr. rel. (ha/farm)*
ctfkk(k) : *rel. fodders - pasture: econometr. intercept (ha/farm)*
dmas(j,k) : *variation coeff. / fodder yield (-)*
fcekk(k) : *number of farms producing cereals (1000)*
fjk(j,k) : *number of farms producing crops or crop clusters (1000)*
fockk(k) : *number of farms producing oilseeds and cereals (1000)*
fovkk(k) : *number of farms producing oilseeds (1000)*
renrj(j,k) : *by-product yield / crop j (tDM/ha/an)*
tenrj(j,k) : *correlating by-prod and prod yields / crop j (tDM/tRM)*
rjk(j,k) : *crop yield (FADN pre-estimated or calibrated) (t/ha)*
rmas(j,k) : *indicative forrage yield (calibration) (t/ha)*
scekk(k) : *sharing coeff. binding cereals and UAA (1)*
sffkk(k) : *sharing coeff. binding fodders and grasslands (1)*

sjk(j,k) : *eligible part of crop area (ha/ha SAU)*
sockk(k) : *sharing coeff. binding oilseeds and cereals (1)*
sovkk(k) : *sharing coeff. binding oilseeds and UAA (1)*

- **the environment block *E***

zcc01 : *(<EF1_2>) EF N2O- Ag soils- Dir.em. ManAppl (kgN2O-N/kgN)*
zcc02(i) : *(<NEX(i)>) Qty N excreted per animal (kg N excr/head/yr)*
zcc03 : *(<FRACGASM>) Qty N excr. volat. (kgNH3-N&NOx-N / kgNexcr)*
zcc04 : *(<FRACFUELM>) Qty N excr. utilized fuel (kgN/kgN)*
zcc05 : *(<FRACPRP>) Qty N direct deposit on soil / anim (kgN/kgN)*
zcc06 : *(<EF3PASTU>) EF N2O Ag soils- ani produc (kgN2O-N/kgNman)*
zcc07(i) : *(<MSPASTU(i)>) Qty N excr dep dir on soil (kgexcr/kgexcr)*
zcc08 : *(<EF4>) EF N2O agsol atm.redep (kgN2O-N/kgNH3-N&NOx-N)*
zcc09 : *(<EF5>) EF N2O Ag soils-Ind em / leaching (kgN2O-N/kgN)*
zcc10 : *(<FRACLEACH>) Qty N leaching (kg N.(kg N)-1)*
zcc11 : *(<EF3SOLID>) EF N20 manure manag syst SOLID (kgN2O-N/kgN)*
zcc12 : *(<EF3DRYLO>) EF N20 manure manag syst DRYLO (kgN2O-N/kgN)*
zcc13 : *(<EF3LIQUI>) EF N20 manure manag syst LIQUI (kgN2O-N/kgN)*
zcc14 : *(<EF3ANAER>) EF N20 manure manag syst ANAER (kgN2O-N/kgN)*
zcc15 : *(<EF3OPENP>) EF N20 manure manag syst OPENP (kgN2O-N/kgN)*
zcc16 : *(<EF3ANADI>) EF N20 manure manag syst ANADI (kgN2O-N/kgN)*
zcc17 : *(<EF3BURNF>) EF N20 manure manag syst BURNF (kgN2O-N/kgN)*
zcc18(i) : *(<MSSOLID(i)>) Frac. MS manure on syst SOLID (%)*
zcc19(i) : *(<MSDRYLO(i)>) Frac. MS manure on syst DRYLO (%)*
zcc20(i) : *(<MSLIQUI(i)>) Frac. MS manure on syst LIQUI (%)*
zcc21(i) : *(<MSANAER(i)>) Frac. MS manure on syst ANAER (%)*
zcc22(i) : *(<MSOPENP(i)>) Frac. MS manure on syst OPENP (%)*
zcc23(i) : *(<MSANADI(i)>) Frac. MS manure on syst ANADI (%)*
zcc24(i) : *(<MSBURNF(i)>) Frac. MS manure on syst BURNF (%)*
zcc25(i) : *(<B0(i)>) Max CH4 produc capa manu prod anim i (m3/kg-dm)*
zcc26 : *(<MCFPASTU>) CH4 conv fact syst PASTU(/clim)(kgCH4/kg-dm)*
zcc27 : *(<MCFSOLID>) CH4 conv fact syst SOLID(/clim)(kgCH4/kg-dm)*
zcc28 : *(<MCFDRYLO>) CH4 conv fact syst DRYLO(/clim)(kgCH4/kg-dm)*
zcc29 : *(<MCFLIQUI>) CH4 conv fact syst LIQUI(/clim)(kgCH4/kg-dm)*
zcc30 : *(<MCFANAER>) CH4 conv fact syst ANAER(/clim)(kgCH4/kg-dm)*
zcc31 : *(<MCFOPENP>) CH4 conv fact syst OPENP(/clim)(kgCH4/kg-dm)*
zcc32 : *(<MCFANADI>) CH4 conv fact syst ANADI(/clim)(kgCH4/kg-dm)*
zcc33 : *(<MCFBURNF>) CH4 conv fact syst BURNF(/clim)(kgCH4/kg-dm)*
zcc34 : *(<EF1_1>) EF Agr soils - dir emis synt fert (kgN2O-N/kgN)*
zcc35 : *(<FRACGASF>) Qty of N intake volatilised (kgN/kgN)*

zcc36 : *(<EF1_3>) EF Agr soils -dir. em N-fix crops (kgN2O-N/kgN)*
zcc37(j) : *(<RESBF/CROPBF(j)>) Fraction resid. / yield (kg/kg)*
zcc38(j) : *(<FRACNCRBF(j)>) Fraction N / plant (kgN/kgdm)*
zcc39(j) : *(<FRACDM(j)>) Dry matter ration / plant (kgdm/kg)*
zcc40 : *(<EF1_4>) EF Agr soils - dir. em crop resid (kgN2O-N/kgN)*
zcc41 : *(<FRACBURN>) Fraction of burnt resid. (kgres/kgres)*
zcc42 : *(<FRACFUEL>) Fract resid. util. c combust (kg_res/kg_res)*
zcc43 : *(<FRACCNST>) Fract resid. util in cstction (kgres/kgres)*
zcc44 : *(<FRACFOD>) Fract resid. util c fodder (kg_res/kg_res)*
zcc45 : *(<FRACR>) Fract resid. util c product (kg_res/kg_res)*
zcc46 : *(<EFRIZ>) EF for CH4 rice production (gCH4/m2)*
zcc47 : *(<FLAG4Dc>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc48 : *(<FLAG4Dd>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc49(i) : *(<ASHCONTENT(i)>) Content ash / manure (kgash/kgmanure)*
zcc50 : *(<FLAG4A>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc51(i) : *(<Ym(i)>) CH4 conver (gross ener in feed conv to CH4) (1)*
zcc52 : *(<FLAG4Da>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc53 : *(<FLAG4De>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc54 : *(<FLAG4Df>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc55 : *(<FLAG4Dg>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc56 : *(<FLAG4Bn>) Flag variable Tier1a, Tier1b, Tier 2 (0-4)*
zcc59(j,k) : *indexing crop / zcc34*
zcc60(j,k) : *constant emission factor N2O related to N*
zcc61(i) : *<MCFPASTU(i)> CH4 conv. fact. / manure manag. sys. PASTU*
zcc62(i) : *<MCFSOLID(i)> CH4 conv. fact. / manure manag. sys. SOLID*
zcc63(i) : *<MCFDRYLO(i)> CH4 conv. fact. / manure manag. sys. DRYLO*
zcc64(i) : *<MCFLIQUI(i)> CH4 conv. fact. / manure manag. sys. LIQUI*
zcc65(i) : *<MCFANAER(i)> CH4 conv. fact. / manure manag. sys. ANAER*
zcc67(i) : *<MCFANADI(i)> CH4 conv. fact. / manure manag. sys. ANADI*
zcc69(i) : *<MCFOTHER(i)> CH4 conv. fact. / manure manag. sys. OTHER*
zcc70(i) : *<MCFDAILY(i)> CH4 conv. fact. / manure manag. sys. DAILY*
zcc71(j) : *modulation of zcc01 (Bouwman) by crop (-)*
zcc72 : *<EF3SOLIDDRYLOT> Em. Fact. solid storage + dry lot (0-4)*
zcc73 : *<EF3OTHER> EF N20 manure manag syst OTHER (kgN2O-N/kgN)*
zcc74 : *<EF3COMPOST> Emission factor for composting (kgN2O-N/kgN)*
zcc75 : *<EF3DAILY> Emission factor for daily spread (kgN2O-N/kgN)*
zcc76(i) : *<MCFCOMPOST> CH4 conv. fact. / manure manag. sys. COMPOST*
zcc77(i) : *<MCFSOLIDDRY> CH4 conv. fac. / man.manag. sys. SOLID.DRY.*
zcc78(i) : *<MSSOLIDDRYLOT> Frac. MS manure sys. SOLID STORAGE DRYLOT*
zcc79(i) : *<MSCOMPOST> Frac. MS manure sys. COMPOSTING*
zcc80(i) : *<MSDAILY> fraction manure sys. DAILY SPREAD*

zcc81(i) : *<MSOTHER> fraction manure sys. OTHER*
zcc82(i) : *<MCFBURNF> CH4 conv. fact. / manure manag. sys. BURNFuel*
zcc83 : *<MCFSOLIDDRYLOT> CH4 con fact syst SOLIDDRY (kgCH4/kg-dm)*
zcc84 : *<MCFCOMPOST> CH4 conv fact syst COMPOSTING (kgCH4/kg-dm)*
zcc85 : *<MCFDAILY> CH4 conv fact syst DAILY SPREAD (kgCH4/kg-dm)*
zcc86(i) : *<FRACgasANAER(i)> anaerobic lagoon / fraction manure sys.*
zcc87(i) : *<FRACgasLIQUI(i)> system Liquid / fraction manure sys.*
zcc88(i) : *<FRACgasSOLIDDRYLOT(i)> Solid stor/ fraction manure sys.*
zcc89(i) : *<FRACgasPASTU(i)> system Pasture / fraction manure sys.*
zcc90(i) : *<FRACgasOTHER(i)> Other system / fraction manure sys.*
zcc91(i) : *<FRACgasANADI(i)> Digesters / fraction manure sys.*
zcc92(i) : *<FRACgasDAILY(i)> Daily Spread / fraction manure sys.*
zcc93(i) : *<FRACgasCOMPOST(i)> syst. Compost. / fraction manure sys.*
zcc94(i) : *<FRACgasBURNF(i)> Burned fuel-waste / fraction manure sy.*
zcc95(i) : *<FRACleachANAER(i)> N loss / runoff manag. manure / zcc86*
zcc96(i) : *<FRACleachLIQUI(i)> N loss / runoff manag. manure / zcc87*
zcc97(i) : *<FRACleachSOLIDDRYLOT(i)> N loss / runoff man. m. / zcc88*
zcc98(i) : *<FRACleachPASTU(i)> N loss / runoff manag. manure / zcc89*
zcc99(i) : *<FRACleachOTHER(i)> N loss / runoff manag. manure / zcc90*
zcc100(i) : *<FRACleachANADI(i)> N loss / runoff manag. manure / zcc91*
zcc101(i) : *<FRACleachDAILY(i)> N loss / runoff manag. manure / zcc92*
zcc102(i) : *<FRACleachCOMPOST(i)> N loss / runoff manag. man. / zcc93*
zcc103(i) : *<FRACleachBURNF(i)> N loss / runoff manag. manure / zcc94*
a1no3(k) : *coeff subsoil NO3 / geographical area unit in the region*
a2no3(k) : *coeff livestock / geographical unit in the NO3 conc.*
a3no3(k) : *estim IDPR in the regression related to NO3 concentr.*
a4no3(k) : *coeff IDPR in the NO3 level regression*
b1no3(k) : *intercept regression linking aver. NO3 conc. / explic.*
b2no3(k) : *crops / geographical area (related to hydrological zone)*
b3no3(k) : *UAA / geographical area (related to hydrological zone)*
c1no3(k) : *NO3 conc. threshold (related to hydrological zone) (mg/l)*
ano3(j,k) : *corr / emission factor NO3*
bno3(j,k) : *constant emission factor NO3*
cno3(j,k) : *corr / emission factor N03 related to irrig*
anh3(j,k) : *corr emission factor NH3*
bnh3(j,k) : *constant emission factor NH3 j k*
cnh3(j,k) : *corr / emission factor NH3 related to irrig*
cn2o(j,k) : *corr / emission factor N2O related to irrig*
bouw1 : *intercept in the Bouwman relation (N_N2O / N)*
bouw2 : *corr. fact. in the Bouwman relation (N_N2O / N)*
cesjk(j,k) : *yearly C stock qtity per crop (t C /ha)*

cfme(i) : *conversion ratio methane animal manure (%)*
fch41 : *intercept related to methane em. / simple feed*
fch42 : *correlation effect related to methane em. / raw feed*
fch43 : *intercept related to methane em. / mix feed*
fch44 : *correlation effect related to methane em. / mix feed*
fech(i) : *yearly CH4 emission factor related to manure (kg)*
fedf(i) : *enteric CH4 emission factor by defaut,IPCC (kg/head)*
fen2o : *direct N2O emission factor, IPCC*
gesjk(j,k) : *yearly C stock qtity / additionnal pasture (t C /ha)*
pcch4 : *CH4 calorific power (kcal/kg CH4)*
txcch : *GHG emissions : tax / CH4 (k€/tCO2)*
txcno : *GHG emissions : tax / N2O (k€/tCO2)*
prgch : *GWP CH4 (kgCO2 / kgCH4)*
prgn : *GWP N2O (kgCO2 / kgN2O)*
zjk(j,k) : *variable costs specific to pestici. (part of vjk) (k€/ha)*
zsttt : *factor multiplying pesticide cost accounts when >0 (-)*

It must be noticed that the *i*-indexed parameters, `zcc61(i)` up to `zcc70(i)`, were added in 2015, and used for estimating the *GHG* emissions related to the *V5* versions of *AROPAj* (*EU* - 27, 2007 - 2009). These parameters replace, respectively, `zcc26` up to `zcc33`.

- **the feed block *F***

brac(n,i,k) : *raw energy / concentrated feed intakes (kcal/kgDM)*
brag(i,k) : *raw energy / raw feed intakes (kcal/kgDM)*
clpdc : *PD needs for milk / intercept for dairy cows (kg/hd)*
clpdp(k) : *PD needs per ton of milk / coef for dairy cows (kg/t)*
clufc : *PD needs for milk / intercept for dairy cows (kg/hd)*
clufp(k) : *PD needs per ton of milk / coef for dairy cows (kg/t)*
digm(i) : *average digestibility intake, IPCC, anim. i (%)*
ebac(n,i,k) : *ingestion coeff. / energy content / concentrates (ce/t)*
ebag(i,k) : *ingestion coeff. / raw matter deed (ce/t)*
ebes(j) : *raw energy content of crop (kcal/kg DM)*
ebi(i,k) : *ingestion capacity by animal i (ce/head)*
ebj(j,k) : *ingestion coeff. / crop j (ce/t)*
ed(j) : *j-crop digestibility (1)*
eda1 : *feed concentrate digestibility (type 1 simpl energy) (1)*
eda2 : *feed concentrate digestibility (type 2 simpl protein) (1)*
eda3 : *feed concentrate digestibility (type 3 compound ener) (1)*
eda4 : *feed concentrate digestibility (type 4 compound prot) (1)*
edag : *raw feed digestibility (1)*
emac(n,i,k) : *metabolisable energy in concentrated feed graniv (Mcal/t)*

emag(i,k) : *metabolisable energy in raw feed for granivore (Mcal/t)*
enac(n,i,k) : *digestibility of concentrated feed per animal (1)*
enag(i,k) : *digestibility of raw feed per animal (1)*
emi(i,k) : *yearly requirement of metabolisable energy (Mcal/head)*
emj(j,k) : *metabolisable energy in crop j farm k (Mcal/t)*
eni(i,k) : *yearly proteins requirement per animal i farm k (t/head)*
fglai : *flag for adjusted feeding parameters*
pac(n,i,k) : *bought conc. feed price (incl. proteins) (k€/t)*
pag(i,k) : *bought raw feed price (k€/t)*
palac : *relative change in bought conc. feed price (1)*
palag : *relative change in raw feed price (1)*
pdib(i,k) : *digest. protein requirement of i ruminant (kg/head)*
pdie(j,k) : *N equiv. by j crop in PDIE (kg/t)*
pdin(j,k) : *N equiv. by j crop in PDIN (kg/t)*
peac(n,i,k) : *N equiv. by n concentr. feed in PDIE / anim. i (kg/t)*
peag(i,k) : *N equiv. by raw feed in PDIE / anim. i (kg/t)*
pnac(n,i,k) : *PDIN N-equivalent / concentr. feed (kg/t)*
pnag(i,k) : *PDIN N-equivalent / raw feed (kg/t)*
smi(i,k) : *dry matter requirement / animal i (tDM/head)*
smj(j,k) : *dry matter ratio in j-crop (1)*
smac(n,i,k) : *dry matter ratio in concentr. feed / i-animal (1)*
smsag(i,k) : *dry matter ration in raw feed / i-animal (1)*
tconv : *conversion ratio of i-animal digestibility in CH4*
trp(j,k) : *ratio of j-cereal content proteins / energy*
ufac(n,i,k) : *UF content in concentr. feed / i-animal (UFL_eq/t)*
ufag(i,k) : *UF content in raw feed / i-animal (UFL_eq/t)*
uflj(j,k) : *UFL content in j-crop (UFL/t)*
uflv(i,k) : *UFL requirement / UFV for i-animal (UFL_V/t)*
ufvj(j,k) : *UFV content in j-crop (UFV/t)*

- **the nitrogen block *N***

apolt(o,k) : *input N org LT / N mineral equiv*
aqolt(o,k) : *LT effect N org equivalent N mineral*
benjk(j,k) : *N need of crop (yield) (tN /(t.ha))*
casti : *test for benjk parameter use (0/1)*
coo(o,i,j,k) : *mineralisation fact. dejection o animal i / j*
cpro(o,j,k) : *mineralisation fact. UOR o / crop j*
endrg : *indicator for yield calculation activation (0/1/2) (-)*
enj(j,k) : *fraction of nitrogen matter in crops (feed) (t/t)*
epd(o,i,k) : *spreadable proportion of manure type o / animal i*
flf(o,i,k) : *part solid/liquid/droppings / effluent (t/t)*

fmfu(i) : *average mineral fraction in manure, IPCC (%)*
fned(o,i,k) : *fraction N in effluent o from animal i (kgN/kgo)*
fnedo(o,k) : *fraction N in UOR type o*
fnes(j,k) : *fraction N / fertilizers for crop j (kgN /kg fert)*
fnesc(j,k) : *fraction N / ferti. regul. for crop j (kgN /kg fert)*
fnesf(j,k) : *fraction N / ferti. starter for crop j (kgN /kg fert)*
mgesd : *indentifier 0/1 activating manure spreading block (-)*
yexjk(j,k) : *calculated N input requirement (crop model) (tN/ha)*
pexne : *price of exceeding N / spreadable limit*
proo(o,k) : *UOR type o available for Farm Group k (t RawMat)*
prixn : *relative change in price of fertilizer N-content (-)*
prxn(j,k) : *j-crop fertilizer price / farm group k (k€/t)*
prxnc(j,k) : *regular fertilizer price / crop j / farm group k (k€/t)*
prxnf(j,k) : *starter fertilizer price / crop j / farm group k (k€/t)*
qengf(j,k) : *starter fertilizer quantity brought to j-crop (t/ha)*
qidej(o,i,k) : *o-manure quantity excret. by i-animal (t ou m3/head)*
volat(o,k) : *N-volatil. coeff. reletd to o-manure/UOR storage (tN/tN)*
wjk(j,k) : *irrigation costs (k€/ha)*

- **the policy block *P***

aidkk(k) : *lump-sum farm payment (k€/farm)*
cerjk(j) : *caloric content per 100g of j-crop (kcal/hg)*
cerik(i) : *caloric content per 100g of i-animal (meat) (kcal/hg)*
calai : *caloric content per 100g of milk (kcal/hg)*
calorac(n) : *caloric content per 100g of concentrated feed (kcal/hg)*
calorag : *caloric content per 100g of (hay) raw feed (kcal/hg)*
capro : *protein / calorie ratio requested for diet (adefinir)*
proteac(n) : *caloric yield from protein per 100g of feed (kcal/hg)*
proteag : *caloric yield from protein per 100g of raw feed (kcal/hg)*
berjk(j) : *caloric yield from protein per 100g of j-crop (kcal/hg)*
berik(i) : *caloric yield from protein per 100g of i-animal (kcal/hg)*
cplai : *caloric yield from protein per 100g of milk (kcal/hg)*
derik(i) : *food content per i-animal (t/head)*
acerj : *ACE energy crop subsidy non set-aside / CAP*
aenrj(j) : *ACE energy crop subsidy non set-aside / CAP per crop j*
aidkk(k) : *lump-sum farm payment (k€/farm)*
aii(i) : *subsidy per animal ("small" herd) (k€/head)*
ajj(j) : *subsidy per crop (k€/t)*
amah : *threshold defining small herbivore producer (LU/farm)*
amaj : *threshold small prod. (young bovine premium) (LU/farm)*
angel : *min. area ratio set-aside allowing subsidy (1988) (0<-<1)*

ave : *subsidy / calves (k€/head)*
bii(i) : *animal premium (McSh - extensif. animal) (k€/head)*
bjj(j) : *per crop area premium (McSh hard wheat) (k€/ha)*
bjk(j,k) : *net "subsidy" / area (type XD wood fallow) (k€/ha)*
cerea : *threshold McSh CAP cereal yield (4.6 t/Ha) (t/ha)*
chgpa : *Livestock limit 1 (grass premium) (LU/ha)*
chgpb : *Livestock limit 2 (grass premium) (LU/ha)*
chgpc : *Livestock limit 3 (grass premium) (LU/ha)*
chgpd : *Livestock limit 4 (grass premium) (LU/ha)*
chgpe : *Livestock limit 5 (grass premium) (LU/ha)*
chgpf : *Livestock limit 6 (grass premium) (LU/ha)*
chgpg : *Livestock limit 7 (grass premium) (LU/ha)*
chgph : *Livestock limit 8 (grass premium) (LU/ha)*
chgpi : *Livestock limit 9 (grass premium) (LU/ha)*
chgtj : *Livestock limit young bovine premium (LU/ha)*
chgtk : *Livestock limit superpremium young bovine (LU/ha)*
chgui : *Livestock limit herbivores / non subsid. prem. (LU/ha)*
chgus : *Livestock limit grass super-premium (LU/ha)*
cii(i) : *subsid top up / national complement / animal (k€/head)*
cjj(j) : *subsid top up / national complement / vegetal (k€/ha)*
cjk(j,k) : *revenue / ha out of FEOGA (type XD wood fallow) (k€/ha)*
cqlai : *reduct. (>0) milk quota without prem. red. (1-cqlai) (-)*
crlai : *top up milk (k€/t)*
cupp : *coeff. min. set-aside part to get set-as prem Msh (0<-<1)*
cupq : *coeff. fix set-aside part (0<-<1)*
cvpp : *coeff. maximal set-aside to get premium Msh (0<-<1)*
djj(j) : *floor price McSharry reform / SCOP (k€/t)*
faxou : *max cumulative premimum for grass production (k€)*
ienrj(j) : *energy use identifier (val. 1 in case)*
yxdej(j,k) : *upper limit of epandable org. N / crop area unit (tN/ha)*
nhpgp : *option number for piecewise linear. subs. for grass (1)*
ojj(j) : *add. subsid. per hectare*yield of crop j (k€/t)*
olea : *ref. oilseeds yield (2.36 t/ha) (t/ha)*
penrr(j) : *relative price change of by-product / crop j (-)*
pesac : *tax/subsid. (tax if >0) on concentr. feed (GHG) (k€/t)*
pesag : *tax/subsid. (tax if >0) on raw feed (GHG) (k€/t)*
pgms : *McSh set-aside subsid. (block / CAP cereal yield) (k€/t)*
piac(n,i) : *tax/subsid. on n-concentr. feed / anim. i (k€/t)*
piag(i) : *tax/subsid. on raw feed / animal i (k€/t)*
pges(j) : *subsidy(<0) / tax(>0) ferti. crop j (k€/tN)*
pgesn : *subsidy(<0) / tax(>0) fertilizers (k€/tN)*

pgeso : *subsidy(<0) / tax(>0) organic Nitrogen intakes (k€/tN)*
pies(i) : *subsidy(<0) / tax(>0) animal i (GHG) (k€/head)*
pjes(j) : *subsidy(<0) / tax(>0) crop j (GHG) (k€/ha)*
pjr(j) : *reference price before CAP reform basing djj(j) (k€/t)*
plaid : *relative change of milk price on 2nd market (1)*
plaiw(k) : *milk price on 2nd market (€/l)*
prges : *GHG em. tax/subsidy on CO2 equiv. (k€/tCO2)*
qenrj(j,k) : *quota holding crop j final energy use (t/F)*
prgpa : *premium 1 livestock limit 1 (grass subsidy) (k€/ha)*
prgpb : *premium 2 livestock limit 2 (grass subsidy) (k€/ha)*
prgpc : *premium 3 livestock limit 3 (grass subsidy) (k€/ha)*
prgpd : *premium 4 livestock limit 4 (grass subsidy) (k€/ha)*
prgpe : *premium 5 livestock limit 5 (grass subsidy) (k€/ha)*
prgpf : *premium 6 livestock limit 6 (grass subsidy) (k€/ha)*
prgpg : *premium 7 livestock limit 7 (grass subsidy) (k€/ha)*
prgph : *premium 8 livestock limit 8 (grass subsidy) (k€/ha)*
prgpi : *premium 9 livestock limit 9 (grass subsidy) (k€/ha)*
prgte : *superpremium young bov. limit 1 (young bov subs.) (k€/ha)*
prgtj : *young bov subs. LS limit 1 (young bov. subs.) (k€/ha)*
prgtl : *milk cow subs. LS limit 1 (milk cow subsidy) (k€/ha)*
prgtn : *suckler cow subs. LS limit 1 (suckler cow subs.) (k€/ha)*
prjml : *male calve subsidy 8 days / milk herd (k€/head)*
pqlai : *milk quota price in case of quota market (k€/t)*
prlai : *milk premium / Agenda 2000 / under initial quota (k€/t)*
psgel(k) : *set-aside premium / Luxembourg decoupling (k€/ha)*
psngl(k) : *non set-aside premium L Luxembourg decoupling (k€/ha)*
psrpa(k) : *grassland premium / Luxembourg decoupling (k€/ha)*
psrta(k) : *arable land premium Luxembourg decoupling (k€/ha)*
pssap(k) : *single farm-dependent area payment (k€/ha)*
psuni : *single area payment (k€/ha)*
puges : *premium(<0) / tax(>0) LU (GHG) (k€/LU)*
qdfau(k) : *sugarbeet frac. A+B/TOT in case of missing estim. quota*
qpcab(k) : *min frac. of C sugarbeet related to A and B*
qfo(n,k) : *upper limit per support category (Lux. CAP reform) (k€)*
qgelu(k) : *set-aside quota / Luxembourg decoupling (ha)*
qjk(j,k) : *vegetal quota (kt)*
qlait(k) : *milk quota (kl)*
qlaiw(k) : *milk quota 2nd market (kl)*
qopes(k,l) : *GHG emission quota (sub-group k,l)*
qorkk(k) : *barley quota (kt)*
qpsth(k) : *lower bound / grasslands (INSEA / C storage)*

qtalk(k) : *reference milk yield / young bovine premium (t/VL/an)*
qtete : *young bov. number limit / young bov. premium (heads)*
qtola : *milk quota limit / young bov. premium (t/an)*
rafou : *ratio grassland / UAA required for grass subsidy (1)*
rcehm(k) : *region ref. yield of cereal but maize (ir. or not) (t/ha)*
rceja(k) : *reference yield / fallow (t/ha)*
rcemg(k) : *region reference yield maize (irrig or not) (t/ha)*
redes : *single reduction coef. on quota C/GHG*
rqlai : *exceeding milk quota / initial quota Agenda 2000 (1)*
rwlai : *variation / milk quota 2nd market / quota B (1)*
quank(k) : *financial quantum (k€/farm)*
sbett(k) : *average sugarbeet area (under quota) (ha/farm)*
scop(k,l) : *reference COP area limiting set-aside subsidy (ha/farm)*
senrj(k) : *upper bound of total area in energy crops (ha/farm)*
rii(i) : *slaughter premium (Agenda 2000 = option 14) (k€/head)*
sii(i) : *threshold "small" herd (national) (head/farm)*
sjc(j,k) : *reference crop area accounting for C storage (ha)*
spre : *additional premium if LS load < chgus / amateur (k€/LU)*
ndivb : *min number of crops for 1st diversification threshold*
ndivc : *min number of crops for 2nd diversification threshold*
sdmax(k) : *max limit for crop diversification (ha)*
sdmin(k) : *min limit for crop diversification (ha)*
sudiv(k) : *subsidy for crop diversification*
thrss(k) : *crop area limit accounting for crop diversification*
tabkk(k) : *ratio binding sugarbeet quotas A and B (1)*
tan : *equiv. cereal "230t" in COP (non-subs. set-aside) (t)*
tegkk(k) : *fertilizer tax (cox, cozy()) (1)*
timou : *UAA threshold for grass subsidy (ha)*
tjj(j) : *j-crop tax (see tjk(j,k)) (1)*
tjk(j,k) : *marketed prod. j-crop tax (price * (1+t)) (1)*
tlaid : *"superlevy" / milk price 2nd market (1)*
ton : *equiv. cereal "92t" in COP (threshold "amateur") (t)*
txn2o : *direct N2O loss tax (k€/tN)*
txnh3 : *direct NH3 loss tax (k€/tN)*
txno3 : *direct N03 loss tax (k€/tN)*
u(i,j,k) : *LS load i-animal / j-fodder area (head/ha)*
ugbou : *lower bound of herd for grass subsidy (LSU)*

- **the water block *W***

ywajk(j,k) : *calculated water input requirement (crop model) (m3/ha)*
pxeau(j,k) : *water price / crop j / farm group k (k€/m3)*

wjk0(j,k) : *preserved water costs in case of recalculation (k€/ha)*
xsikk(k) : *shadow price of k- irrig water limit case endog. (k€/m3)*
xsitt : *common shadow price of water limit in case endog (k€/m3)*
wirig(k) : *water irrigation limit k-scale (m3)*
wirtt : *water irrigation limit for the whole set of k-GT (m3)*

- **the "trans-farm group" block *T***

qymin(j) : *Min j-crop Qty to be marketed at domestic price (t)*
qycal : *Min calories Qty to be exported from agr.sys. (teqcere)*
qene(n) : *Min bio-energy typ1 provided 1+2G from ag.sys. (tgazoleq)*
qymce : *Min cereal Qty to be marketed at domestic price (t)*
qmxnm : *Max Qty of mineral nitrogen (t)*
qmxge : *Max Qty of GHG emissions (soil-C exclud.) (tCO2eq)*
eenrj(n,j,k) : *biomass transformed into energy / crop j farm k (tbd))*

4.2.2 Activities operating in the model

Activities designate the command variables (also named by primal variables) against which the optimization problem is solved.

The comprehensive list of activities also includes those activities which were incorporated in past AROPAj versions but are now possibly obsolete. We provide units in which variables are expressed. At the farm level, main units are metric tons (t), livestock units (LU), animal head numbers ($head$), hectares (ha), 1000 liters (kl), metric tons of CO_2eq (tCO_2eq), 1000 € ($k€$). At the farm group level, numerical figures have to be multiplied by f, which is implicitly equal to 1000 and relates to the representativity of the farm group. In other words, when the number of farms represented by a farm group is 954, the population enters in the model with the value 0.954, $f = 1000$.

The animal and plant activities on which the model is based are listed in Table **??** with the labels used in the files. It should be added that the links between these two groups of activities exist in different forms, in particular through animal feed (by distinguishing the cereals and fodder produced and consumed on the farm on the one hand, raw feed or concentrates purchased on the other hand) and by nitrogen inputs (distinguishing between fertilizers of industrial origin and organic amendments sourced from manure).

- **the general block *G***

aokeu(k) : not if CONDITION *work paid - wages per farm (kE)*
aokwu(k) : not if CONDITION *work unpaid - working units per farm (AWU)*
aoket : not if CONDITION *work paid (kE)*
aokut : not if CONDITION *work unpaid (AWU)*
efo(n,k,l) : *generic variable of "deficit or excedent" (CAP) (var)*
excdt(n) : *generic variable "deficit or excedent or total" (var)*
ief(n,k,l) : not if CONDITION *generic binary variable*

zf(k,l) : not if CONDITION *farm number k l (f)*

- **the animal block *A***

ani(i,k,l) : *animal number cat. i (total) (heads)*
ant(i) : not if CONDITION *aggregated i-herd (heads)*
xolai(k,l) : not if CONDITION *dairy milk production (f*kl)*
xolat : *total dairy milk production (f*kl)*
xugb(k,l) : not if CONDITION *livestock (per sub-group) (f*LU)*
xugbk(k) : *livestock (per group) (f*LU)*
xugbt : *total livestock (f*LU)*
yolai(k,l) : not if CONDITION *dairy milk product. entitling premium Agenda2000 (f*kl)*

- **the crop block *C***

aut(j) : *aggregated on-farm use of j-crop prod. (f*t)*
c(i,j,k,l) : not if CONDITION *on-farm j-prod. to feed i-animal (f*t)*
dctr : *trade balance defic. / sugarbe (if demand info.) (f*k€)*
excr : *trade balance excess / sugarbe (if demand info.) (f*k€)*
x(j,k,l) : not if CONDITION *standard j-crop area (out of CAP) (f*ha)*
xtk(j,k) : *aggregated j-crop area (per k) (f*ha)*
xtt(j) : *aggregated j-crop area (f*ha)*
y(j,k,l) : not if CONDITION *marketed j-crop prodution (farm) (f*t)*
y0(j,k) : not if CONDITION *marketed j-crop prodution (farm) exluding energy (f*t)*
ye(o,j,k) : not if CONDITION *marketed j-crop prodution (farm) for energy type O (f*t)*
yc(j,k) : not if CONDITION *marketed j-crop by-product (f*t)*
yb0(j,k) : not if CONDITION *marketed j-crop by-product excluding bioenergy (f*t)*
yb(o,j,k) : not if CONDITION *marketed j-crop by-product for energy type O (f*t)*
ycj(j) : *aggregated j-crop by-product*
yi(j) : *marketed j-crop production (domestic price) (f*t)*
ym(j) : *marketed j-crop prod. (out of quota / worldprice) (f*t)*

- **the environment block *E***

emcdt(n) : *generic counterpart of GHG balance (tCO2)*
epcdt(n) : *balance N-pollutant emission (3 cat.: N2O NO3 NH3)*
qzz(j,k) : not if CONDITION *utilized amount of pesticide on area j (k EUR)*

- **the feed block *F***

acgta(m) : not if CONDITION *aggregated feed (conc. bought / cat.) (f*t)*
b(n,i,k,l) : *concentrate feed cat. n bought for i-herd (f*t)*
agt(i,k,l) : not if CONDITION *raw feed bought for i-herd (f*t)*
dcgta : *aggregated feed expendit.per cat. (conc. or raw) (f*k€)*
f(i,j,k,l) : not if CONDITION *j-fodder area used by i-animal (f*ha)*

xfour(k,l) : not if CONDITION *fodder area (includ. fod. maize non COP subsid.) (f*ha)*

- **the nitrogen block *N***

po(j,k,o) : not if CONDITION *UOR quantity of type o used by crop j (t)*
qen(j,k) : not if CONDITION *quantity of regular fertilizer used on j-crop area (t)*
q(i,j,k,o) : not if CONDITION *quantity of manure type o / i-herd sourced (t ou m3)*
qon(j,k) : not if CONDITION *useful quant. N from manure + UOR spread on area j (t N)*
qonlt(k) : not if CONDITION *average quantity of org. N brought on long term per ha*
xo(o,i,k) : not if CONDITION *exceeding unused organic N type o i-anim. sourced*

- **the policy block *P***

a(n,i,k,l) : not if CONDITION *subsidized herbivore number (animal extensif.) (f*head)*
aig(i,k,l) : not if CONDITION *exceeding "small" herd threshold (heads)*
aip(i,k,l) : not if CONDITION *herd number under the "small" herd threshold (heads)*
aml(i,k,l) : not if CONDITION *male calves from milk bovine herd (heads)*
amv(i,k,l) : not if CONDITION *male calves from breeding cows (heads)*
dolat : *fictive deviation / "qolai"*
dwlat : *fictive deviation / "qwlai" 2nd market*
egcdt(n) : *generic balance excess (support cat. in CAP)*
foh(n,k,l) : not if CONDITION *total fodder related to grass subsidy (f*ha)*
fph(n,k,l) : not if CONDITION *grasslands related to grass subsidy category (f*ha)*
h(j,k,l) : not if CONDITION *mark. out of quota or guarant. price (except sugar) (f*t)*
her(n,k,l) : not if CONDITION *binaries / subsid. herbivores (extensif. animal) (-)*
igdlg(k,l) : not if CONDITION *binary related to set-aside luxemb. reform*
igdlh(k,l) : not if CONDITION *binary related to set-aside luxemb. reform*
igelg(k,l) : not if CONDITION *binary related to set-aside 1988*
igelh(k,l) : not if CONDITION *binary related to set-aside 1988*
iph(n,k,l) : not if CONDITION *fodder related to grass subsidy*
ivegn(k,l) : not if CONDITION *variable used in exclusion principle McSharry*
ivegp(k,l) : not if CONDITION *variable used in exclusion principle McSharry*
qta(n,k,l) : not if CONDITION *subsid. young bov. + (suckler + ref. dairy cow) (f*head)*
qtb(n,k,l) : not if CONDITION *young bovine numb. related to premium (f*head)*
qtc(n,k,l) : not if CONDITION *suckler cows rel. to prem. (young bov. prem.) (f*head)*
qti(n,k,l) : not if CONDITION *binary related to premium (young bovine premium) (-)*
qtx(n,k,l) : not if CONDITION *fodders related to premium (young bovine premium) (f*ha)*
soc(o,k) : not if CONDITION *deviation of area for carbon sequestration 1=+/2=- (ha)*
ugh(n,k,l) : *herbivore number related to CAP (extensif. animal) (f*LU)*
ugo(n,k,l) : not if CONDITION *herbivore number related to CAP (grass subsidy) (f*LU)*
xa(j,k,l) : not if CONDITION *compensated COP area (rel. quant. < "92t") (f*ha)*
xb(j,k,l) : not if CONDITION *compensated COP area (quant. exceeding "92t") (f*ha)*
xc(j,k,l) : not if CONDITION *compensated COP area (quant. exceeding "230t") (f*ha)*

xarab(k) : not if CONDITION *cultivated areas except fodders (f*ha)*
xdivb(k) : not if CONDITION *eligible crop area: diversific. subs. small (f*ha)*
xdivc(k) : not if CONDITION *eligible crop area: diversific. subs. inter (f*ha)*
xdivd(k) : not if CONDITION *eligible crop area: diversific. subs. big (f*ha)*
xdive(k) : not if CONDITION *relaxing area variable for diversific. subs. (f*ha)*
idivc(k) : not if CONDITION *binary var related to XDIVC*
idivd(k) : not if CONDITION *binary var related to XDIVD*
idivb(k) : not if CONDITION *binary var related to XDIVB*
idive(k) : not if CONDITION *binary var related to XDIVE*
xd(j,k,l) : not if CONDITION *compensated COP / fix set-aside (non rotat.) (f*ha)*
xe(j,k,l) : not if CONDITION *agro-energy area out of set-aside but with ACE*
xea(j,k,l) : not if CONDITION *sub-cat. agro-energy area out of set-aside but with ACE*
xeb(j,k,l) : not if CONDITION *sub-cat. agro-energy area out of set-aside but with ACE*
xec(j,k,l) : not if CONDITION *sub-cat. agro-energy area out of set-aside but with ACE*
xed(j,k,l) : not if CONDITION *sub-cat. agro-energy area out of set-aside but with ACE*
xf(j,k,l) : not if CONDITION *agro-energy area on set-aside without ACE*
xfa(j,k,l) : not if CONDITION *agro-energy area on set-aside McSh replacing XA*
xfb(j,k,l) : not if CONDITION *agro energy area on set-aside McSh replacing XB*
xfl(j,k,l) : not if CONDITION *agro-energy area on set-aside "Luxembourg"*
xgelo : *aggregated set-aside whatever CAP options (f*ha)*
xolqm(k,l) : not if CONDITION *sale (>0) purchase (<0) milk quota (rent)*
xpgdl(k,l) : not if CONDITION *grassla / single prem./ha "x49kl-lux" set-asi (f*ha)*
xrgdl(k,l) : not if CONDITION *arable la single prem./ha "x49kl-lux" set-aside (f*ha)*
xsgdl(k,l) : not if CONDITION *area complenting XRGDL "x49kl-lux" set-aside (f*ha)*
xsgel(k,l) : not if CONDITION *threshod area completing "x42kl" 1988 set-aside (f*ha)*
itk(j,k) : not if CONDITION *crop index related to aggregated j-crop area (per k)*
nttjk(k) : not if CONDITION *number of crops cultivated (per k) in the LP solution*
xwlai(k,l) : not if CONDITION *dairy milk production / 2nd market (farm) (f*kl)*
xwlat : *dairy milk production/ 2nd market (f*kl)*

- **the water block *W***

awj(j,k) : not if CONDITION *quantity of irrigation water per crop (m3/ha)*
exwkk(k) : not if CONDITION *balance irrig water use at k farm level (m3)*

- **the trans-farm group block *T***

4.2.3 The objective function and constraints

The objective function and the set of constraints are listed below. The "not if CONDITION" string refers to a flag depending on parameters (explicitly referenced in the flag or through functions, not listed here: all these intermediary functions are parts of the *aro.mgg* code transformed into *FORTRAN* code and operating upstream of the use of the *GAMS* solver). Some constraints

may be obsolete, for different reasons (change in CAP, old version of the model), but we keep on the principle "what has worked should continue to work." The suffix l is maintained to one value, i.e. 1, except when farm groups are disaggregated into subgroups. It must be noticed that constraints may be activated, not only when flags apply, but also when some parameters are non null. At last, the mathematical expression of constraints uses "elements" crossing activities, which are expressed by 3 uppercase letters (possibly omitted when the element value is equal to 1). For the sake of paper length, the definitions of such elements are not detailed in this paper, except for the "GHG" and "nitrogen" blocks (see Sections **??** and **??**). Similarly, the water block will be added, expected in 2017 (**??**, {not still detailed}).

- **the general block G**

obj : *objective function (f*k€)* max obj

$$
\begin{aligned}
&+\sum_{i,k,l} CPN * ani(i,k,l) + \sum_{j,k,l} COX * x(j,k,l) + \sum_{j,k,l} COX * xa(j,k,l) \\
&+\sum_{j,k,l} COX * xb(j,k,l) + \sum_{j,k,l} COX * xc(j,k,l) + \sum_{j,k,l} COX * xd(j,k,l) \\
&+\sum_{j,k,l} COX * xe(j,k,l) + \sum_{j,k,l} COX * xf(j,k,l) - \sum_{n,i,k,l} OAC * b(n,i,k,l) \\
&-\sum_{i,k,l} OAG * agt(i,k,l) + \sum_{j,k,l} POI * y(j,k,l) + \sum_{j,k} POC * yc(j,k) \\
&+\sum_{j,k,l} POE * h(j,k,l) + \sum_{k,l} POL * xolai(k,l) + \sum_{k,l} PWL * xwlai(k,l) \\
&+\sum_{k,l} CPU * zf(k,l) - \sum_{n} OES * emcdt(n) + \sum_{k,l} POQ * xolqm(k,l) \\
&-\sum_{j,k} CQE * qen(j,k) - \sum_{j,k} POW * awj(j,k) - \sum_{n} FEG * excdt(n) \\
= obj
\end{aligned}
$$

nb(k,l) : not if CONDITION *refers to farm group number (= per default) (1000)*

$$zf(k,l) = fzt$$

sci(j) : *balance / marketed j-crop by-product (f*t)*

$$ycj(j) - \sum_{k} yc(j,k) = 0.0$$

si(j) : *balance / domestic market-priced j-crop product (f*t)*

$$yi(j) - \sum_{k,l} y(j,k,l) = 0.0$$

sm(j) : *world price sale or out of domestic quota (f*t)*

$$-\sum_{k,l} h(j,k,l) + ym(j) = 0.0$$

sx(j) : *aggregated j-crop area*

$$\sum_k xtk(j,k) - xtt(j) = 0.0$$

sxk(j,k) : *aggregated j-crop area (whatever the CAP applying)*

$$\begin{aligned}&\sum_l x(j,k,l) + \sum_l xa(j,k,l) + \sum_l xb(j,k,l)\\&+\sum_l xc(j,k,l) + \sum_l xd(j,k,l) + \sum_l xe(j,k,l)\\&+\sum_l xf(j,k,l) - xtk(j,k) = 0.0\end{aligned}$$

terre(k,l) : not if CONDITION *land endowment limit (f*ha)*

$$\sum_j TXJ * xtk(j,k) - TZT * zf(k,l) \leq 0.0$$

tat(i) : not if CONDITION *aggregated i-animal activity (heads)*

$$ant(i) - \sum_{k,l} ani(i,k,l) = 0.0$$

tugb(k,l) : not if CONDITION *total livestock per sub-group (obsolete) (f*LU)*

$$xugb(k,l) - \sum_i TUG * ani(i,k,l) = 0.0$$

tugbk(k) : *total livestock per group (f*LU)*

$$\sum_l xugb(k,l) - xugbk(k) = 0.0$$

tugbt : *total livestock (f*LU)*

$$xugbt - \sum_k xugbk(k) = 0.0$$

tugh(k,l) : not if CONDITION *total herbivore livestock (f*LU)*

$$\sum_n ugh(n,k,l) - \sum_{i1} TUG * ani(i1,k,l) = 0.0$$

tuh(n,k,l) : not if CONDITION *aggregated herbiv. liv. (extensif. subsid.) (f*LU)*

$$ugh(n,k,l) - \sum_{i1} TUG * a(n,i1,k,l) = 0.0$$

tun(i,k,l) : not if CONDITION *aggregated i-herbivore herd (f*head)*

$$\sum_n a(n,i,k,l) - ani(i,k,l) = 0.0$$

fx1(j,k) : not if CONDITION *fixed areas constrain 1*

$$-xtk(j,k)+\sum_{l} FY1*zf(k,l) \leq 0.0$$

fx2(j,k) : not if CONDITION *fixed areas constrain 2*

$$xtk(j,k)-\sum_{l} FY2*zf(k,l) \leq 0.0$$

fx3(j,k) : not if CONDITION *fixed areas constrain 3 limite inf prairie*

$$-xtk(j,k)+\sum_{l} FY1*zf(k,l) \leq 0.0$$

wokeu(k) : not if CONDITION *work paid (f*kE/farm)*

$$-aokeu(k)+\sum_{l} WEK*zf(k,l) = 0.0$$

wokwu(k) : not if CONDITION *work unpaid (f*AWU/farm)*

$$-aokwu(k)+\sum_{l} WUK*zf(k,l) = 0.0$$

woket : not if CONDITION *work paid (f*kE)*

$$-aoket+\sum_{k} aokeu(k) = 0.0$$

wokwt : not if CONDITION *work unpaid (f*AWU)*

$$-aokut+\sum_{k} aokwu(k) = 0.0$$

- **the animal block *A***

d(m,i,k,l) : *amplitude limit ref. to i-animal number adjust. (f*head)*

$$DIA*ani(i,k,l)-DIZ*zf(k,l) \leq 0.0$$

das(m,k,l) : *adjustment links within animal cat. (bo ta ve) (f*head)*

$$\sum_{i} DSA*ani(i,k,l)-DSZ*zf(k,l) \leq 0.0$$

cpani : not if CONDITION *gross animal product subvent includ? (igene+2**4) (f*k€)*

$$+\sum_{i,k,l} CPN*ani(i,k,l)-\sum_{n,i,k,l} OAC*b(n,i,k,l)-\sum_{i,k,l} OAG*agt(i,k,l)$$
$$+\sum_{k,l} POL*xolai(k,l)+\sum_{k,l} PWL*xwlai(k,l)+\sum_{n} CEX*excdt(n)$$
$$+\sum_{k,l} POQ*xolqm(k,l) = 0.0$$

dem(m,k,l) : *bovine demographic block (heads)*

$$+\sum_i DMO * ani(i,k,l) + \sum_i DML * aml(i,k,l) + \sum_i DMV * amv(i,k,l) \leq 0.0$$

polai(k,l) : not if CONDITION *cow milk production (f*l)*

$$xolai(k,l) + xwlai(k,l) - \sum_{i1} ROL * ani(i1,k,l) \leq 0.0$$

- **the crop block *C***

inc(j) : *aggregated on-farm reuse of j-cereal*

$$\sum_{i,k,l} c(i,j,k,l) - aut(j) = 0.0$$

rmx(m,k) : *crop categ. rotations (f*ha)*

$$\sum_j ROX * xtk(j,k) - \sum_l ROZ * zf(k,l) \leq 0.0$$

roj(j,k) : not if CONDITION *crop limits and rotations (f*ha)*

$$\sum_{j1} ROT * xtk(j1,k) \leq 0.0$$

roper(k) : not if CONDITION *perennial crops limit (f*ha)*

$$\sum_j ROP * xtk(j,k) - \sum_l ROR * zf(k,l) \leq 0.0$$

s(j,k) : not if CONDITION *j-crop product sharing (market vs on-farm) (f*t)*

$$-RRR * xtk(j,k) + \sum_{i,l} RRF * f(i,j,k,l) + \sum_l y(j,k,l)$$
$$+\sum_l h(j,k,l) + \sum_{i,l} RRC * c(i,j,k,l) = 0.0$$

sc(j,k) : not if CONDITION *available by-product outputs (f*t)*

$$yc(j,k) - REN * xtk(j,k) = 0.0$$

sbtt(k) : not if CONDITION *sugarbeet production sharing*

$$-\sum_j RRA * xtk(j,k) + \sum_{j,l} CNC * y(j,k,l) + \sum_{j,l} CND * h(j,k,l) = 0.0$$

- **the environment block *E***

bz(j,k) : not if CONDITION *Pest expenditure on crop J in farm K (k EUR)*

$$-qzz(j,k) + BEZ * xtk(j,k) = 0.0$$

bilxn : not if CONDITION *organic N balance referring to spreading*

$$\sum_{n} BXA * excdt(n) + \sum_{i,j,k,o} QJM * q(i,j,k,o) + \sum_{j,k,o} QJO * po(j,k,o) = 0.0$$

cpest : *total expenditure in pesticides*

$$+\sum_{j,k} qzz(j,k) - \sum_{n} CVZ * excdt(n) = 0.0$$

ems(m) : *GHG balances per category (IPCC) (tCO2)*

$$\begin{aligned}
&\sum_{n} GES * emcdt(n) - \sum_{i,k,l} AFS * ani(i,k,l) - \sum_{j,k} CFS * xtk(j,k) \\
&- \sum_{n,i,k,l} CSB * b(n,i,k,l) - \sum_{i,j,k,l} CSC * c(i,j,k,l) - \sum_{i,j,k,l} CSF * f(i,j,k,l) \\
&- \sum_{i,k,l} CSG * agt(i,k,l) - \sum_{j,k} CDC * qen(j,k) - \sum_{j,k} CDD * qon(j,k) \\
&- \sum_{j,k,o} CDP * po(j,k,o) - \sum_{o,k} CSX * soc(o,k) = 0.0
\end{aligned}$$

epn(m) : *balances of N pollut. emis. originated by crops man.*

$$\begin{aligned}
&\sum_{n} PNE * epcdt(n) - \sum_{j,k} PNX * xtk(j,k) - \sum_{j,k} PNQ * qen(j,k) \\
&- \sum_{j,k} PNW * awj(j,k) - \sum_{j,k} PNO * qon(j,k) - \sum_{j,k,o} PNP * po(j,k,o) = 0.0
\end{aligned}$$

fbcar : not if CONDITION *balance / soil carbon storage (in progress) (tC/yr)*

$$+\sum_{j,k} CCX * xtk(j,k) + \sum_{o,k} GGX * soc(o,k) - \sum_{n} GE4 * excdt(n) = 0.0$$

qges : *GHG emission quota (cumul / k l) (f*tCO2zq)*

$$\sum_{n} QES * emcdt(n) - \sum_{k,l} ZES * zf(k,l) \leq 0.0$$

qprai(k) : not if CONDITION *minimal limit / grasslands (INSEA) / carbon storage*

$$-\sum_{j} QPA * xtk(j,k) + \sum_{l} QPZ * zf(k,l) \leq 0.0$$

- **the feed block *F***

eq(m,i,k,l) : *i-herd feeding of type m (kg_or_t/head)*

$$\begin{aligned}
&BES * ani(i,k,l) - \sum_{j} VAF * f(i,j,k,l) - \sum_{j} VCF * c(i,j,k,l) \\
&- \sum_{n} VAC * b(n,i,k,l) - VAG * agt(i,k,l) \leq 0.0
\end{aligned}$$

falcg(m) : not if CONDITION *bought feed quantity per herd cat. and per type m*

$$\sum_{n,i,k,l} FAC * b(n,i,k,l) + \sum_{i,k,l} FAG * agt(i,k,l) + \sum_{n} FAX * excdt(n) = 0.0$$

fcgta : *total of aggregated (marketed) expenditure for feed*

$$\sum_{n,i,k,l} DAC * b(n,i,k,l) + \sum_{i,k,l} DAG * agt(i,k,l) - dcgta = 0.0$$

- **the nitrogen block *N***

apnlt(k) : not if CONDITION *calculation of average per ha organic spreading*

$$APA * qonlt(k) - \sum_{i,j,o} APQ * q(i,j,k,o) - \sum_{j,o} APP * po(j,k,o) = 0.0$$

ba(j,k) : not if CONDITION *N requirement to reach given yield (kg N)*

$$-FNE * qen(j,k) - FNO * qon(j,k) + BEN * xtk(j,k) \leq 0.0$$

cnvol : not if CONDITION *sum of mineral fertilizer inputs*

$$+\sum_{j,k} qen(j,k) - \sum_{n} CVO * excdt(n) = 0.0$$

cnmin : *sum of mineral N intakes in terms of equiv. mineral N*

$$+\sum_{j,k} CVJ * qen(j,k) - \sum_{n} CVN * excdt(n) = 0.0$$

cnorg : not if CONDITION *sum of organic N intakes in terms of equiv. mineral N*

$$+\sum_{j,k} qon(j,k) - \sum_{n} CVR * excdt(n) = 0.0$$

ei(o,i,k) : not if CONDITION *o-type manure provided by i-herd*

$$\sum_{j} q(i,j,k,o) + xo(o,i,k) - \sum_{l} QID * ani(i,k,l) = 0.0$$

epro(o) : not if CONDITION *use of UOR disposal*

$$\sum_{j,k} po(j,k,o) - \sum_{k,l} QRO * zf(k,l) \leq 0.0$$

eqj(j,k) : not if CONDITION *balance spreading - source of available N (tN)*

$$qon(j,k) - \sum_{i,o} QJD * q(i,j,k,o) - \sum_{o} QJE * po(j,k,o) \leq 0.0$$

- **the policy block *P***

eld(m,k) : *CAP 2013 - crop diversification (ha)*

$$CED*xdivc(k)+CEE*xdivd(k)+CEB*xdivb(k)$$
$$+CEF*xdive(k)+CEA*xarab(k)+\sum_{j}CEL*xtk(j,k)$$
$$+\sum_{l}CEM*zf(k,l)+CIC*idivc(k)+CID*idivd(k)$$
$$+CIB*idivb(k)+CIF*idive(k)+\sum_{j}TII*itk(j,k)$$
$$+TIN*nttjk(k)\leq tiz$$

ild(j,k,m) : *CAP 2013 - crop diversification*

$$XIT*xtk(j,k)+TIT*itk(j,k)\leq 0.0$$

epa(j,k) : not if CONDITION *environmental limit to organic N spreading (tN)*

$$\sum_{i,o}QJM*q(i,j,k,o)+\sum_{o}QJO*po(j,k,o)-OXD*xtk(j,k)\leq 0.0$$

epno3(k) : *environmental limit to organic N spreading (NO3/l)*

$$\sum_{n}ENO*excdt(n)+ENU*xugbk(k)-\sum_{l}ENZ*zf(k,l)\leq 0.0$$

fog(m) : *FEOGA (export refunds not included) (f*k€)*

$$\sum_{n}GOG*egcdt(n)+\sum_{k,l}CAD*zf(k,l)+\sum_{n,i,k,l}CPA*a(n,i,k,l)$$
$$+\sum_{i,k,l}CPE*aip(i,k,l)+\sum_{i,k,l}CPM*aml(i,k,l)+\sum_{i,k,l}CPP*ani(i,k,l)$$
$$+\sum_{n,k,l}CQB*qtb(n,k,l)+\sum_{n,k,l}CQC*qtc(n,k,l)+\sum_{n,k,l}CPH*ugh(n,k,l)$$
$$+\sum_{j,k,l}FGX*x(j,k,l)+\sum_{j,k,l}FXA*xa(j,k,l)+\sum_{j,k,l}FXA*xea(j,k,l)$$
$$+\sum_{j,k,l}FXA*xfa(j,k,l)+\sum_{j,k,l}FXB*xb(j,k,l)+\sum_{j,k,l}FXB*xeb(j,k,l)$$
$$+\sum_{j,k,l}FXB*xfb(j,k,l)+\sum_{j,k,l}FXC*xc(j,k,l)+\sum_{j,k,l}FXC*xec(j,k,l)$$
$$+\sum_{j,k,l}FXD*xd(j,k,l)+\sum_{j,k,l}FXD*xed(j,k,l)+\sum_{j,k,l}FXE*xe(j,k,l)$$
$$+\sum_{n,k,l}CFO*foh(n,k,l)+\sum_{n,k,l}CFE*efo(n,k,l)+\sum_{j,k,l}FGI*y(j,k,l)$$
$$+\sum_{n}FEX*excdt(n)+\sum_{k,l}YOL*yolai(k,l)+\sum_{k,l}FWL*xwlai(k,l)$$
$$+\sum_{k,l}FDL*xrgdl(k,l)+\sum_{k,l}PDL*xpgdl(k,l)+\sum_{k}SDI*xdivb(k)$$
$$+\sum_{k}SDI*xdivc(k)+\sum_{k}SDI*xdivd(k)=0.0$$

fbges : not if CONDITION *public support and tax balance / GHG em. regul. (f*k€)*

$$+\sum_{i,k,l} GAN * ani(i,k,l) + \sum_{n,i,k,l} GAC * b(n,i,k,l) + \sum_{i,k,l} GAG * agt(i,k,l)$$
$$+\sum_{j,k} GAX * xtk(j,k) + \sum_{j,k} GAE * qen(j,k) - \sum_{n} GEX * excdt(n)$$
$$-\sum_{n} OES * emcdt(n) = 0.0$$

fbn2o : not if CONDITION *public balance regarding support and tax / N2O*

$$\sum_{n} PNH * epcdt(n) - \sum_{n} GNH * excdt(n) = 0.0$$

fbno3 : not if CONDITION *public balance regarding support and tax / NO3*

$$\sum_{n} PNI * epcdt(n) - \sum_{n} GNI * excdt(n) = 0.0$$

fbnh3 : not if CONDITION *public balance regarding support and tax / NH3*

$$\sum_{n} PNJ * epcdt(n) - \sum_{n} GNJ * excdt(n) = 0.0$$

gdl(m,k,l) : not if CONDITION *refers to set-aside decoupling*

$$GD1 * xsgdl(k,l) + GD2 * xrgdl(k,l) + GD7 * xpgdl(k,l)$$
$$+\sum_{j} GD3 * x(j,k,l) + \sum_{j} GDF * xfl(j,k,l) + GD4 * igdlg(k,l)$$
$$+GD5 * igdlh(k,l) - GD6 * zf(k,l) \leq 0.0$$

gdu(k,l) : not if CONDITION *set-aside / Lux. agreem. (igdlh + igdlg = 1)*

$$igdlg(k,l) + igdlh(k,l) = 1.0$$

gfo(n,k,l) : not if CONDITION *Luxembourg decoupling reform (cap limits)*

$$-egcdt(n) - GOZ * zf(k,l) \leq 0.0$$

gel(m,k,l) : *1988 set-aside policy*

$$GL1 * xsgel(k,l) + \sum_{j} GL2 * x(j,k,l) + \sum_{j} GL2 * xa(j,k,l)$$
$$+\sum_{j} GL2 * xb(j,k,l) + \sum_{j} GL2 * xc(j,k,l) + \sum_{j} GL2 * xd(j,k,l)$$
$$+\sum_{j} GL2 * xe(j,k,l) + \sum_{j} GL2 * xf(j,k,l) + GL3 * igelg(k,l)$$
$$+GL4 * igelh(k,l) \leq 0.0$$

gli(k,l) : not if CONDITION *refers to set-aside 1988 (igelh + igelg = 1)*

$$igelg(k,l) + igelh(k,l) = 1.0$$

icgta(m) : *aggregated bought feed of type m (19 cat.) (t)*

$$\sum_{n,i,k,l} CAC * b(n,i,k,l) + \sum_{i,k,l} CAG * agt(i,k,l) - acgta(m) = 0.0$$

pap(i,k,l) : not if CONDITION *threshold referring to "small" herd (heads)*

$$+aip(i,k,l) - OVE * zf(k,l) \leq 0.0$$

pas(i,k,l) : not if CONDITION *balance "big" herd exceed. the threshold (heads)*

$$aip(i,k,l) + aig(i,k,l) - ani(i,k,l) = 0.0$$

pfoui(k,l) : not if CONDITION *CAP grass subsidy accounting for subsid. fodders (-)*

$$+\sum_{n} iph(n,k,l) = 1.0$$

pft(m,k,l) : *CAP grass subsidy redeployment*

$$\sum_{n} PTO * foh(n,k,l) + \sum_{n} PTP * fph(n,k,l) - \sum_{j} PTX * x(j,k,l)$$
$$+\sum_{n} PTH * ugh(n,k,l) - \sum_{n} PTU * ugo(n,k,l) = 0.0$$

pfx(i,k,l) : *CAP grass subsidy (follows) (f*ha)*

$$\sum_{n} PFO * foh(n,k,l) + \sum_{n} PFP * fph(n,k,l) + \sum_{n} PFI * iph(n,k,l)$$
$$+\sum_{n} PFU * ugo(n,k,l) + \sum_{n} PFE * efo(n,k,l) + \sum_{n} PFD * ief(n,k,l)$$
$$+PFZ * zf(k,l) \leq pfs$$

pjt(m,k,l) : *CAP young bov. premium / redeploy. jbv+VN+VLref (f*LU)*

$$\sum_{n} PTA * qta(n,k,l) + \sum_{n} PTB * qtb(n,k,l) + \sum_{n} PTC * qtc(n,k,l)$$
$$+\sum_{n} PTF * qtx(n,k,l) + PTR * xfour(k,l) + \sum_{i} PTN * ani(i,k,l)$$
$$+\sum_{n} PTI * qti(n,k,l) + PTZ * zf(k,l) = pts$$

pjx(m,k,l) : *CAP young bovine premium (f*LU)*

$$\sum_{n} PXA * qta(n,k,l) + \sum_{n} PXB * qtb(n,k,l) + \sum_{n} PXC * qtc(n,k,l)$$
$$+\sum_{n} PXF * qtx(n,k,l) + \sum_{n} PXI * qti(n,k,l) - PXZ * zf(k,l) \leq 0.0$$

polat : *balance / cow milk under quota 1st market (f*l)*

$$+xolat - \sum_{k,l} xolai(k,l) = 0.0$$

psh(m,k,l) : *CAP herbivore extensification (f*LU)*

$$\sum_{n} PRH * her(n,k,l) + \sum_{n} PRU * ugh(n,k,l) + PRX * xfour(k,l) \leq 0.0$$

psi(k,l) : not if CONDITION *CAP McSh excluding binaries "extensif. herbivores"*

$$\sum_{n} her(n,k,l) = 1.0$$

pve(m,k,l) : *CAP McSh if tsy(2)=1 (f*t)*

$$\begin{aligned} &+\sum_{j} PVA * xa(j,k,l) + \sum_{j} PVA * xfa(j,k,l) + \sum_{j} PVA * xea(j,k,l) \\ &+\sum_{j} PVB * xb(j,k,l) + \sum_{j} PVB * xfb(j,k,l) + \sum_{j} PVB * xeb(j,k,l) \\ &+\sum_{j} PVC * xc(j,k,l) + \sum_{j} PVC * xec(j,k,l) + \sum_{j} PVD * xd(j,k,l) \\ &+\sum_{j} PVD * xed(j,k,l) + PVN * ivegn(k,l) + PVP * ivegp(k,l) \\ &+PVZ * zf(k,l) \leq 0.0 \end{aligned}$$

pvi(k,l) : not if CONDITION *CAP McSh (exclud. constr. refer. to IVEGP et IVEGN)*

$$ivegn(k,l) + ivegp(k,l) = 1.0$$

pwlat : *balance milk cow prod. out of quota 2nd market (f*l)*

$$+xwlat - \sum_{k,l} xwlai(k,l) = 0.0$$

qbtb(k,l) : not if CONDITION *refers to sugarbeet quotas A and B (f*t)*

$$\sum_{j} QAB * y(j,k,l) = 0.0$$

qrj(j,k) : not if CONDITION *energy destin. limit on fallow / set-aside and ACE*

$$\sum_{l} xe(j,k,l) + \sum_{l} xf(j,k,l) - \sum_{l} QQE * zf(k,l) \leq 0.0$$

qyj(j,k) : not if CONDITION *CAP energy destinated domestic production*

$$-\sum_{l} y(j,k,l) + \sum_{l} RYE * xe(j,k,l) + \sum_{l} RYE * xf(j,k,l) \leq 0.0$$

qolai(k,l) : not if CONDITION *cow milk quota (f*l)*

$$xolai(k,l) + dolat + xolqm(k,l)$$
$$-QOL * zf(k,l) \leq 0.0$$

qolqt : not if CONDITION *balance / milk quota market (sold if > 0)*

$$-\sum_{k,l} xolqm(k,l) + \sum_{n} QOT * excdt(n) = 0.0$$

qoyl(m,k,l) : not if CONDITION *condition on payment of milk subsidy (f*l)*

$$yolai(k,l) - QOX * xolai(k,l) - QOZ * zf(k,l) \leq 0.0$$

quant(k,l) : not if CONDITION *financial quantum (itest+2**9) (f*k€)*

$$+\sum_{n} QUX * efo(n,k,l) + QUZ * zf(k,l) - \sum_{n} FEG * excdt(n) = 0.0$$

qve(n,k,l) : *crop marketed prod. quota (tsy(8)=1.) (f*t)*

$$\sum_{j} QQY * y(j,k,l) + \sum_{j} QQF * xe(j,k,l) + \sum_{j} QQF * xf(j,k,l)$$
$$+\sum_{j} QQH * h(j,k,l) - QQZ * zf(k,l) \leq 0.0$$

qwlai(k,l) : not if CONDITION *cow milk quota / 2nd market (f*l)*

$$xwlai(k,l) + dwlat - QWL * zf(k,l) \leq 0.0$$

regbt : *special sugarbeet regime (before the 2007 ref.)*

$$+\sum_{j,k,l} TBE * y(j,k,l) + dctr - excr = 0.0$$

snrjm(k) : not if CONDITION *maximum area for energy - fallow / set-aside and ACE*

$$\sum_{j,l} xe(j,k,l) + \sum_{j,l} xf(j,k,l) - \sum_{l} SMX * zf(k,l) \leq 0.0$$

snrjt : not if CONDITION *total area / energy - fallow / set-aside and ACE*

$$\sum_{j,k,l} xe(j,k,l) + \sum_{j,k,l} xf(j,k,l) - \sum_{n} SMT * excdt(n) = 0.0$$

sglo : *aggregation of set-aside areas whatever the CAP*

$$\sum_{j1,k} xtk(j1,k) - xgelo = 0.0$$

sxe(j,k,l) : not if CONDITION *aggregated energy crop area out of set-aside*

$$xea(j,k,l)+xeb(j,k,l)+xec(j,k,l)$$
$$+xed(j,k,l)-xe(j,k,l)=0.0$$

sxf(j,k,l) : not if CONDITION *aggregated energy crop area on set-aside*

$$xfa(j,k,l)+xfb(j,k,l)+xfl(j,k,l)$$
$$-xf(j,k,l)=0.0$$

txsoc(k,l) : not if CONDITION *computation of land deviation in fav. of C storage (f*ha)*

$$\sum_j TXT*xtk(j,k)-\sum_o TXX*soc(o,k)-TZX*zf(k,l)=0.0$$

tfhme(k,l) : not if CONDITION *total fodder except subsidized fodder maize (f*ha)*

$$xfour(k,l)-\sum_j FOU*x(j,k,l)=0.0$$

- **the water block W**

bw(j,k) : not if CONDITION *water requirement to reach given yield (kg N)*

$$-FWA*awj(j,k)+BEW*xtk(j,k)=0.0$$

cwate : not if CONDITION *sum of irrigation water consumptions*

$$+\sum_{j,k} awj(j,k)-\sum_n CWA*excdt(n)=0.0$$

cwakk(k) : not if CONDITION *sum of irrigation water consumptions k farm*

$$+\sum_j awj(j,k)-CWK*exwkk(k)=0.0$$

walim(k) : not if CONDITION *water limit k-farm scale (m3)*

$$-\sum_l WAL*zf(k,l)+\sum_j awj(j,k)\leq 0.0$$

waltt : not if CONDITION *water limit for weighted k-GT setz (km3)*

$$+\sum_{j,k} awj(j,k)\leq wtt$$

- **the "trans-farm group" block T**

minyi(j) : not if CONDITION *j-crop Qty marketed subject to threshold (t)*

$$yi(j)\geq qnj$$

minen(n) : not if CONDITION *bioenergy subject to threshold (teqenergy)*

$$+\sum_{o,j,k} QCY * yb(o,j,k) + \sum_{o,j,k} QEY * ye(o,j,k) \geq qnr$$

sje(j,k) : not if CONDITION *excluding energy and non-ene. uses of mark. crop (tDM)*

$$y0(j,k) + \sum_{o} ye(o,j,k) - \sum_{l} y(j,k,l) \leq 0.0$$

sce(j,k) : not if CONDITION *excluding energy and non-ene. uses of by-products (tDM)*

$$yb0(j,k) + \sum_{o} yb(o,j,k) - yc(j,k) \leq 0.0$$

mincl : not if CONDITION *diet calories subject to threshold (teqcereal)*

$$\sum_{j} YCV * yi(j) + \sum_{j} YCV * ym(j) + \sum_{i} YCA * ant(i)$$
$$+\sum_{m} YAL * acgta(m) + YLK * xolat \geq qcl$$

minpr : not if CONDITION *protein quantity subject to threshold (teqcereal)*

$$\sum_{j} YPV * yi(j) + \sum_{j} YPV * ym(j) + \sum_{i} YPA * ant(i)$$
$$+\sum_{m} YPL * acgta(m) + YPK * xolat \geq qpl$$

minyc : not if CONDITION *cereal Qty marketed subject to threshold (t)*

$$\sum_{j} YMC * yi(j) \geq qnc$$

maxcn : not if CONDITION *maximal threshold of mineral nitrogen (t)*

$$\sum_{j,k} qen(j,k) \leq qxn$$

maxqg : not if CONDITION *maximal GHG limit (tCO2eq)*

$$\sum_{n} QES * emcdt(n) \leq qgr$$

4.2.4 *CAP* implementation

In great part, the *AROPAj* model originates from modeling tools aiming at *CAP* impact assessments. Mathematical programming supplies effective tools for this end when the implementation of new policy tools in economic models has to rely on non-time dependent data. Due to the wide and complex range of *CAP* tools used throughout the decades (until the mid-2000's), the *CAP* block grew immensely.

We keep all previously implemented policy tools operational in *AROPAj*. In reference to real policies, we need to introduce series of binary variables acting as cap and threshold mechanisms. This makes the presentation of sub-matrices related to each *CAP* policy tool incorporated into the *AROPAj* model incredibly time- and space-convuming. Below, we present a selection of such sub-matrices, like the one presented above in Table **??**.

Table 3. Policy tools implemented in the *AROPAj* model.

policy (i):	2^i	tool description
(1):	2	set-aside 1988
(2):	4	MacSharry vegetal (focused on compensated areas)
(3):	8	discriminating thresholds / herd size (part subsidied)
(4):	16	animal extensif. (herbiv/ herd subsidied if load < threshold)
(5):	32	grass subsidy (areas subsidied if *LS*-load < threshold)
(6):	64	young male bovine and suckler cow subsidy (see 2**13)
(7):	128	subsidy to male calves from milk herd
(8):	256	crop marketed production quota
(9):	512	financial quantum
(10):	1024	variation in relative prices and costs
(11):	2048	policy toward farm afforestation support
(12):	4096	regulation policy against greenhouse gas emissions
(13):	8192	Santer package / bovines (see 2**6 ⇒ 6 OR 13)
(14):	16384	Agenda 2000 / bovines (see 2**6 ⇒ 6 OR 14)
(15):	32768	milk quota market (!!! quota renting !!!)
(16):	65536	Decoupling Luxemburg 2003 (option 14 basement)

Most of the policy tools are activated through the parameter `itest(k)`. This is summarized in Table **??**. The addition of the i-policy to a policy set for simulation means that the `itest(k)` has to be additionally increased by 2^i. In relation to the comprehensive, detailed, and mathematical forms of the objective and constraints (Section **??**), some differences may appear. This is the case when the objective is directly impacted by *CAP* subsidies, whereas the above subsidies impact support balances and therefore indirectly impact the objective function. This only a question of presentation, aiming at alleviating the formal description of *CAP* sub-matrices.

Set-aside implemented by the 1988 *CAP*

This tools functions as a binomial tariff. When the part of arable land set aside exceeds the threshold designed by the *CAP*, land set-aside is proportionally subsidized. We extend the policy tool in the sense that the per hectare premium may be indexed per farm group.

This policy tool is activated when the `itest(k)` parameter increases by 2. The *LP* set-aside activity refers to `x(42,k,l)`. Added parameters are the premium (`pjk(42,k)`), which acts as a price multiplied by one quantity unit) and the threshold (`angel`). A variables cost may be used (`vjk(42,k)`).

The sub-matrix related to this policy involves 2 binary variables, one real variable and 3 constraints as described on Table **??**.

Table 4. *CAP* tool (1): set-aside, 1988.

	x(42,k,l)	xsgel(k,l)	igelg(k,l)	igelh(k,l)	x,xa-c(j^=,k,l)		RHS
obj	-c(42,k)+pgl(k)				...		
gel(1,k,l)		1.		-99999.	...	<=	
gel(2,k,l)	angel - 1.	-1.			angel	<=	
gel(3,k,l)	1.		-99999.			<=	
gli(k,l)			1.	1.		=	1.

Crop part of the 1992 MacSharry reform

This *CAP* option stylizes the decoupling propositions related to crops, which were adopted in 1992. Added parameters are reference yields: (`rcehm(k)` for cereals except maize, `rceja(k)` for set-aside, `rcemg(k)` for maize), `olea` and `cerea` for other reference yields; thresholds in terms of cereal equivalent quantities (*ton* related to "small producer" threshold and `tan` related to non subsidized set-aside, that has finally not been implemented in the *EU*), set-aside thresholds (`cupp`, `cvpp`); and series of parameters referring to prices and quantity or area support (`ajj(j)`, `bjj(j)`, `pgms`, `djj(j)`).

The sub-matrix related to this set of policy tools implies the introduction of binary variables (`ivegn(k)`, `ivegp(k)`) and real variables referring to subsidized crops (`xa(j,k,l)`, `xb(j,k,l)`, `xc(j,k,l)`, for $j \leq 72$; $j = 46$ refers to subsidized set-aside, $j = 47$ refers to non subsidized set-aside, $j = 72$ refers to fodder maize). Series of constraints (`pve(m,k,l)`, `pvi(k,l)`) are built compactly as described in Table **??**.

"Fixed" set-aside, in addition to "rotating" set-aside, are incorporated in the model by means of additional parameters and *LP* variables, which are not shown in this subsection.

Livestock extensification

Subsidies are provided depending on the ratio of bovine livestock to hectare. We describe here a policy in which premium and super-premium are proposed to breeders. The system combines premium per *i*-animal head (`bii(i)`) and super-premium (`spre`) with *LS* threshold (`amah`) and *LS* loads per hectare (`chgui`, `chgus`). Added *LP* variables refer to bovine policy categories (`a(m,i,k)`, `ugh(m,k)`) and binary variables (`her(m,k,l)`).

The *CAP* related sub-matrix is described on Table **??**.

Grassland premium

This policy tool was implemented when some stylization questions arose regarding the livestock extensification proposed by the *CAP* reform. In contradistinction from the previous "livestock extensification" option, `nhpgp` levels of *LS* per hectare loads are taken into account (`chgpa, chgpb, ..., chgpi`) in line with `nhpgp` subsidy levels (`prgpa, prgpb, ..., prgpi`), with `nhpgp`= 9. Thresholds accounting for areas and *LS* are also introduced (`rafou, faxou, timou`). We add `nhpgp`+1 series of binary (`iph(m,k,l)`) and real (`foh(m,k,l), fph(m,k,l), ugo(m,k,l)`) *LP* variables referring to the `nhpgp` levels of extensification. Other added variables refer to threshold

Table 5. *CAP* tool (2): MacSharry reform, 1992, for crop compensation.

```
Reference yields are renamed rref(j,k):
          rref(j,k) = rcehm(k) for cereals
                      rcemg(k) for maize
                      rceja(k) for subsidized set-aside
A support function is used for facilitating the reading of the matrix:
          comp(j,k) = rref(j,k)*ajj(j) for cereals, oilseeds and proteins
                      rref(j,k)*ajj(j)*olea/cerea for oilseeds
                      rref(j,k)*pgms for subsidized set-aside
The LP variable "zf(k,l)" (equal to "popul(k,l)") refers to the farm number represented by the k-farm group.
"scop" refers to area dedicetd to cereals, oilseeds and proteins crops.
See additional stuff referring to pve & PVI for CAP option 11 (fix set-aside and forestry eligibility)
-------------------------------------------------------------------------------------------------------------
             |x(j,k,l)|  xa(j,k,l)  |  xb(j,k,l)  |  xc(j,k,l) |ivegn(k)|ivegp(k)| zf(k,l)  | y(j,k,l) |   | RHS
-------------------------------------------------------------------------------------------------------------
 obj         |-c(j,k) | -c(j,k)+comp(j,k)+bjj(j)  |   -c(j,k)  |        |        |          |  djj(j)  |
-------------------------------------------------------------------------------------------------------------
 pve(1,k,l)  |        |  rref(j,k)  |             |            |        |        |  - ton   |          | < |
-------------------------------------------------------------------------------------------------------------
 pve(2,k,l)  |        |-cupp is j#46|             |            |        |        |          |          | < |
             |        |else 1-cupp  |             |            |        |        |          |          |   |
-------------------------------------------------------------------------------------------------------------
 pve(3,k,l)  |        |             |cupp if j#46 |            |        |        |          |          | < |
             |        |             |else cupp-1  |            |        |        |          |          |   |
-------------------------------------------------------------------------------------------------------------
 pve(4,k,l)  |        |      1.     |             |            |        | -400.  |          |          | < |     (mod. 11)
-------------------------------------------------------------------------------------------------------------
 pve(5,k,l)  |        |             |      1.     |      1.    | -400.  |        |          |          | < |     (mod. 11)
-------------------------------------------------------------------------------------------------------------
 pve(6,k,l)  |        |  rref(j,k)  |  rref(j,k)  |            |        |        |  -
tan   |          | < |     (mod. 11)
-------------------------------------------------------------------------------------------------------------
 pve(7,k,l)  |        |             |             |cupp if j#47|        |        |          |          | < |
             |        |             |             |else cupp-1 |        |        |          |          |   |
-------------------------------------------------------------------------------------------------------------
 pve(8,k,l)  |        | if scop=0: -cvpp if j#46  |            |        |        |          |          | < |     (mod. 11)
             |        | if scop=0: else 1-cvpp    |            |        |        |          |          |   |
             |        |                           |            |        |        | scop>0:  |          |   |
             |        |   if scop>0:  1 if j=46   |            |        |        |-cvpp*scop|          |   |
-------------------------------------------------------------------------------------------------------------
 pvi(k,l)    |        |             |             |            |   1.   |   1.   |          |          | = | 1.
-------------------------------------------------------------------------------------------------------------
```

Table 6. *CAP* tool (4): livestock extensification.

```
The k-index related to farm group is omitted
xfour denotes the fodder area
Added constraints : psh(1-6,k,l), psi(k,l), tuh(1-4,k,l), tguh(k,l), tun(i,k,l)
-------------------------------------------------------------------------------------------------------------------------
            |a(1,i,|a(2,i,|a(3,i,|a(4,i,|her(1,|her(2,|her(3,|her(4,|ugh(1,|ugh(2,|ugh(3,|ugh(4,|ani(i,)|xfour(|    RHS
-------------------------------------------------------------------------------------------------------------------------
------
 obj        |      |bii(i)|      |bii(i)|      |      |      |      |      |      | spre | spre |       |      |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(1,k,l) |      |      |      |      |      |      |      |-amah |      |      |      |  1.  |       |      | < |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(2,k,l) |      |      |      |      |      |      |      |      |      |      |  1.  |      |      |       |-
chgui| < |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(3,k,l) |      |      |      |      |      |      |      |      |      |      |      |  1.  |      |       |-
chgus| < |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(4,k,l) |      |      |      |      |      |-9999.|      |      |      |  1.  |      |      |      |       |      | < |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(5,k,l) |      |      |      |      |      |      |-9999.|      |      |      |  1.  |      |      |       |      | < |
-------------------------------------------------------------------------------------------------------------------------
------
 psh(6,k,l) |      |      |      |      |-9999.|      |      |      |  1.  |      |      |      |      |       |      | < |
-------------------------------------------------------------------------------------------------------------------------
------
 psi(k,l)   |      |      |      |      |  1.  |  1.  |  1.  |  1.  |      |      |      |      |      |       |      | = | 1.
-------------------------------------------------------------------------------------------------------------------------
------
 tugh(k,l)  |      |      |      |      |      |      |      |      |  1.  |  1.  |  1.  |  1.  |-
uii(i)|      | = |
-------------------------------------------------------------------------------------------------------------------------
------
 tuh(1,k,l) |-uii(i|      |      |      |      |      |      |      |  1.  |      |      |      |      |       |      | = |
-------------------------------------------------------------------------------------------------------------------------
------
 tuh(2,k,l) |      |-uii(i|      |      |      |      |      |      |      |  1.  |      |      |      |       |      | = |
-------------------------------------------------------------------------------------------------------------------------
------
 tuh(3,k,l) |      |      |-uii(i|      |      |      |      |      |      |      |      |  1.  |      |       |      | = |
-------------------------------------------------------------------------------------------------------------------------
------
 tuh(4,k,l) |      |      |      |-uii(i|      |      |      |      |      |      |      |      |  1.  |       |      | = |
-------------------------------------------------------------------------------------------------------------------------
------
 tuh(i,k,l) |  1.  |  1.  |  1.  |  1.  |      |      |      |      |      |      |      |      |      |  -
1.  |      | = |
-------------------------------------------------------------------------------------------------------------------------
------
```

implementation and to exclusion conditions (`ief(m,k,l)` as binaries and `ugh(k,l), efo(n,k,l)` as reals).

The *CAP*-related sub-matrix is described in Table **??**.

Subsidizing young male bovines and suckler cows

Another complex *CAP* tool was designed for sets of bovine premiums. We now focus on the instrument that impacts the bovine herd, and more specifically the young males and suckler cows eligible for subsidizing. Loading ratio and thresholds apply again. This is stylized via ratios `chgtj, chgtk`, premiums `prgtj, prgte, prgtn`, and number thresholds `qtalk, qtete, qtola, amah`. Thresholds refer to dairy production (dairy cow and milk production). Binary variables (`qti(n,k,l)`) and real variables (`qta(n,k,l), qtb(n,k,l), qtc(n,k,l), qtx(n,k,l)`) are added to the *AROPAj* model.

The *CAP*-related sub-matrix is presented in Table **??**.

The 1992 set-aside policy

Compulsory set-aside schemes were a major introduction of the MacSharry *CAP* reform. We present here the stylization involving rotations and fixed set-aside. Farm afforestation and other perennial productions may be also priced through the parameter `bjk(48,k)` (the 48-index referring to `fp` among the list of land activities). This package completes the 2-*CAP* package presented above (see Table **??**). We need to add set-aside ratio (`cupq`) and the *LP* activity `xd(j,k,l)` ($j = 48$ accounting for fix set-aside).

The *CAP*-related sub-matrix is described in Table **??**.

Part of the Luxembourg decoupling scheme

After the 1992 MacSharry reform, which initiated a first attempt of decoupling (more precisely transferring support from the consumer to the taxpayer), the 2003 Luxembourg agreement completed this transferral, leading to a more or less complete decoupling scheme characterized by a single farm (or single area unit) payments. We consider here that decoupling is complete when the entire *CAP* support is provided by a single payment (as denoted by the support parameter `psuni`). A degree of complexity is introduced when supports that differentiate crops and grasslands are chosen by *MS*. In addition, the implementation of the Luxembourg agreement in *AROPAj* requires some complex combination of the new and previous decoupling schemes when partial re-coupling was chosen by certain *MS*. This is not fully detailed here.

The *CAP*-related sub-matrix is described in Table **??**.

The sugar beet regime (before and after 2006)

The sugar beet support changed shortly after the Luxembourg agreement was implemented within the *EU*. Before this change, the sugar regime referred to "quotas", namely "A" with high price and low tax, "B" with high price (like "A") and high tax, "C" sold at the world price. In addition, a complex system of support was applied to this regime in the case of bio-ethanol use. Quotas "A", "B" and "C" respectively refer to $j = 22$, $j = 23$ and $j = 24$. Common parameters for

Table 7. *CAP* tool (5): a menu of grassland premiums related to *LS* load per hectare.

```
Some indexes are partly omitted.
-----------------------------------------------------------------------------------------------------------------------
--
             | foh(n,) | fph(n,) | ugo(n,) | iph(n,) |x(71-2|x(81-2|ugh|efo(1,|efo(2,|ief(1,|ief(2,|ief(3,|-
zf(k,l)|     RHS
-----------------------------------------------------------------------------------------------------------------------
-----------
 obj         | prgpa-i |         |         |         |      |      |   |  -1. |      |      |      |      |        |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfoui(k,l)  |         |         |         |    1.   |      |      |   |      |      |      |      |      |        | = | 1.
-----------------------------------------------------------------------------------------------------------------------
-----------
 pft(1,k,l)  |    1.   |         |         |         |  -1. |  -1. |   |      |      |      |      |      |        | = |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pft(2,k,l)  |         |    1.   |         |         |      |  -1. |   |      |      |      |      |      |        | = |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pft(3,k,l)  |         |         |    1.   |         |      |      |-1.|      |      |      |      |      |        | = |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(1,k,l)  |         |         |         |     3.  |      |      |   |      |      |  1.  |  1.  |  1.  |        | < | 3.
             |         |         |         |0 if n=1 |      |      |   |      |      |      |      |      |        |   |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(2-,k,l) |         |  -chgpa |  1.     |         |      |      |   |      |      |      |      |      |        | < |
     .       |         |     .   |     .   |         |      |      |   |      |      |      |      |      |        | < |
   nhpgp+1   |         |   -chgp-|       1.|         |      |      |   |      |      |      |      |      |        | < |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(11-,k,l)|1.       |         |         |-999.    |      |      |   |      |      |      |      |      |        | < |
     .       |   .     |         |         |  -999.  |      |      |   |      |      |      |      |      |        | < |
   nhpgp+11  |       1.|         |         |   -999.|      |      |   |      |      |      |      |      |        | < |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(21-,k,l)|         |1.       |         |-999.    |      |      |   |      |      |      |      |      |        | < |
     .       |         |    .    |         |  -999 . |      |      |   |      |      |      |      |      |        | < |
   nhpgp+21  |         |       1.|         |   -999.|      |      |   |      |      |      |      |      |        | < |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(31-,k,l)|         |         |1.       |-999.    |      |      |   |      |      |      |      |      |        | < |
     .       |         |         |    .    |  -999.  |      |      |   |      |      |      |      |      |        | < |
   nhpgp+31  |         |         |       1.|   -999.|      |      |   |      |      |      |      |      |        | < |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(41,k,l) | prgpa-i |         |         |         |      |      |   | -1.  |  1.  |      |      |      | -
faxou | < |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(42,k,l) |         |         |         |  -saukl |      |      |   |      |      |-
999. |      |      | timou  | < |
             |         |         |         |0 if n=1 |      |      |   |      |      |      |      |      |        |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(43,k,l) |         |         |    -1.  |         |      |      |   |      |      |      |-
999. |      | ugbou  | < |
             |         |         |0 if n=1 |         |      |      |   |      |      |      |      |      |        |
-----------------------------------------------------------------------------------------------------------------------
-----------
 pfx(44,k,l) |         |    -1.  |         |         |      |      |   |      |      |      |      |-
999. | rafou* | < |
             |         |0 if n=1 |         |         |      |      |   |      |      |      |      |      | saukl* |   |
-----------------------------------------------------------------------------------------------------------------------
-----------
```

Table 8. *CAP* tool (6): subsidizing young bovine males and suckler cows under certain conditions.

```
For simplification, we use the function: vlref = uii(35)*qlait(k)/qtalk(k).
"xfour()" denotes the fodder area.
"zf()" denotes the farm number related to the farm group.
-----------------------------------------------------------------------------------------------------------------------
            | qta(n,) |   qtb(n=,)  |   qtc(n=,) | qti(n=,)| qtx(n=,)| xfour | ani(21| ani(23| ani(36|  zf()  |     RHS
            | 1 2 3 4 |1   2   3  4 |1   2  3  4 | 1 2 3 4 | 1 2 3 4 |       |       |       |       |        |
-----------------------------------------------------------------------------------------------------------------------
--------
 obj        |         |0 prgtj      |0    prgtn  |         |         |       |       |       |       |        |
            |         |     prgtj+te|            |         |         |       |       |       |       |        |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjt(1,k,l) |   -1.   |             |            |         |         |       |uii(21)|uii(23)|uii(36)| vlref  | = |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjt(2,k,l) |         |     -1.     |            |         |         |       |uii(21)|uii(23)|       |        | = |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjt(3,k,l) |         |             |     -1.    |         |         |       |       |       |uii(36)|        | = |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjt(4,k,l) |         |             |            |         |   -1.   |   1.  |       |       |       |        | = |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjt(5,k,l) |         |             |            |    1.   |         |       |       |       |       |        | = | 1.
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(1-4,k,l)|1.      |             |            |-999.    |         |       |       |       |       |        | < |
            |    .    |             |            |    .    |         |       |       |       |       |        | < |
            |      1.|             |            |    -999.|         |       |       |       |       |        | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(5-8,k,l)|         |1.           |            |-999.    |         |       |       |       |       |        | < |
            |         |      .      |            |    .    |         |       |       |       |       |        | < |
            |         |           1.|            |    -999.|         |       |       |       |       |        | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(9-12,k,l|         |             |1.          |-999.    |         |       |       |       |       |        | < |
            |         |             |     .      |    .    |         |       |       |       |       |        | < |
            |         |             |          1.|    -999.|         |       |       |       |       |        | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(13-16,k,|         |             |            |-999.    |1.       |       |       |       |       |        | < |
            |         |             |            |    .    |    .    |       |       |       |       |        | < |
            |         |             |            |    -999.|       1.|       |       |       |       |        | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(17,k,l) | 0 1 0 0 |             |            |         | -chgtj  |       |       |       |       |        | < |
            |         |             |            |         |n=2, ou 0|       |       |       |       |        |   |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(18,k,l) | 0 0 1 0 |             |            |         | -chgtk  |       |       |       |       |        | < |
            |         |             |            |         |n=3, ou 0|       |       |       |       |        |   |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(19,k,l) | 0 0 0 1 |             |            |         |         |       |       |       |       |  amaj  | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(20,k,l) |         |  0  1  1  0 |            |         |         |       |       |       |       |  qtete | < |
-----------------------------------------------------------------------------------------------------------------------
--------
 pjx(21,k,l) |         |             |            |0 qlait 0|         |       |       |       |       |  qtola | < |
-----------------------------------------------------------------------------------------------------------------------
--------
```

Table 9. *CAP* tool (11): adding fixed set-aside to the rotating set-aside.

```
See also the CAP tool "2".
We used functions: comp(j,k) = rref(j,k)*ajj(j)             if j cereals or protein crops
                               rref(j,k)*ajj(j)*olea/cerea  if j oilseeds
                               rref(j,k)*pgms               if j = 46 (subsidized set-aside)
                   rref(j,k) = rcehm(k) if j cereals,
                               rcemg(k) if maize,
                               rceja(k) if set-aside eligible to premium
-------------------------------------------------------------------------------------------------------------------
            |  xa(j,k,l) |  xb(j,k,l) | xc(j,k,l) |     xd(j,k,l)    |ivegn(k)|ivegp(k)| zf(k,l)  | y(j,k,l) |     RHS
-------------------------------------------------------------------------------------------------------------------
-----
 obj        | -c(j,k)+comp(j,k)+bjj(j)|  -c(j,k)  | -c(j,k)+comp(j,k)|        |        |          |  djj(j)  |
            |                         |           |    +bjk(j,k)     |        |        |          |  djj(j)  |
-------------------------------------------------------------------------------------------------------------------
-----
 pve(4,k,l) |      1.    |           |           |                  |        | -
400.  |          |          | < |      (cf mod. 2)
-------------------------------------------------------------------------------------------------------------------
-----
 pve(5,k,l) |            |      1.    |    1.     |         1.       | -400.  |        |          |          | < |      (cf mod. 2)
-------------------------------------------------------------------------------------------------------------------
-----
 pve(6,k,l) | rref(j,k)  | rref(j,k)  |           |     rref(j,k)    |        |        |  -
tan   |          | < |      (cf mod. 2)
-------------------------------------------------------------------------------------------------------------------
-----
 pve(8,k,l) | if scop=0: -cvpp if j#46|           |scop=0: -cvpp j#48|        |        |          |          | < |      (cf mod. 2)
            | if scop=0: else 1-cvpp  |           |scop=0:selse1-cvpp|        |        |          |          |   |
            |                         |           |                  |        |        | scop>0:  |          |   |
            | if scop>0:   1 if j=46  |           |scop>0: 1 if j=48 |        |        |-cvpp*scop|          |   |
-------------------------------------------------------------------------------------------------------------------
-----
 pve(9,k,l) |            |            |           |   cupq if j#48   |        |        |          |          | < |
            |            |            |           |   else cupq-1    |        |        |          |          |   |
-------------------------------------------------------------------------------------------------------------------
-----
 pvi(k,l)   |            |            |           |                  |   1.   |   1.   |          |          | = | 1.
-------------------------------------------------------------------------------------------------------------------
-----
```

Table 10. *CAP* tool (16): from the 2003 decoupling scheme referring to set-aside toward a unique *SFP* (`psuni`).

```
gfo(n,k,l): CAP balance per support category (for n<=13, n^=2, n^=11)
j=49 refers to new set-aside area
-----------------------------------------------------------------------------------------------------------------------
---
            | igdlg(k,l)  | igdlh(k,l) | xpgdl(k,l)| xrgdl(k,l)| xsgdl(k,l)| x(j,k,l) |xfl(j,k,l)| zf(k,l) |egcdt(n)|     RHS
-----------------------------------------------------------------------------------------------------------------------
------------
 obj        |             |            |           |           |           |          |          |         |    1   |
-----------------------------------------------------------------------------------------------------------------------
------------
 fog(n)     |             |            |           |           |           |          |          |         |        |
 n = ...    |             |            |           |           |           |          |          |         |  -
1   |
 n = 14     |             |            |           |psngl+psrpa|           | psgel(k) |          |         |  -
1   |
 n = 15     |             |            |psngl+psrta|           |           |          |          |         |  -
1   |
 n = 16     |             |            |           |           |           |          |          |  psuni  |  -
1   |
 n = ...    |             |            |           |           |           |          |          |         |  -
1   |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(1,k,l) |qgelu(k)*0.99|            |           |           |           |-1 if j=49|   -
1    |         |        | < |
            | *popul(k,l) |            |           |           |           |          |          |         |        |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(2,k,l) |             |  - 99999.  |           |           |     1     |          |          |         |        | < |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(3,k,l) |  - 99999.   |            |     1     |     1     |           |          |          |         |        | < |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(4,k,l) |             |            |           |           |           | 1 if j=49|     1    |-
qgelu(k)|        | < |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(5,k,l) |             |            |     1     |     1     |     1     | 1 if j=49|     1    |-
saukk(k)|        | < |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdl(6,k,l) |             |            |     1     |           |           |    -
1   |          |         |        | < |
            |             |            |           |           |           |if j=81,82|          |         |        |   |
-----------------------------------------------------------------------------------------------------------------------
------------
 gdu(k,l)   |      1      |      1     |           |           |           |          |          |         |        | = | 1
-----------------------------------------------------------------------------------------------------------------------
------------
 gfo(n,k,l) |             |            |           |           |           |          |          |-qfo(n,k)|  -
1   | < |
-----------------------------------------------------------------------------------------------------------------------
------------
```

sugar beet refer to $j = 21$. There exist some additional constraints for sugar beet production in the "C" quota scenarios as well as the potential energy regime, but they are not delineated here.

This regime was progressively abandoned and replaced by a substantial direct subsidy. In this case, the parameter `tabkk(k)` is of 0-value and quotas are suppressed (i.e. they are valuated at very high value, 99999). The sugar beet production refers to the `x(21,k,l)` area and to the `y(22,k,l)` production.

The *CAP*-related sub-matrix is described in Table **??**.

Table 11. *CAP* tool: the sugar support regime before 2007.

```
Quota may appear through area or quantity, depending of data delivered by the MS to the EU.
We use here the quantity parameter (qdfau(k)).
-------------------------------------------------------------------------------
            | x(j,k,l)   |  zf(k,l)   |  y(22,k,l) | y(23,k,l)  | dctr | excr |     RHS
-----------------------------------------------------------------------------------------
 qve(22,k,l)|            | - qdfau(k) |      1     |            |      |      | < |
-----------------------------------------------------------------------------------------
 qbtb(k,l)  |            |            |  tabkk(k)  |    - 1     |      |      | = |
-----------------------------------------------------------------------------------------
 regbt      |            |            |-pjk(21,k,l)|-pjk(21,k,l)|   1  |  -1  | = |
            |            |            | * tjk(22)  | * tjk(23)  |      |      |   |
-----------------------------------------------------------------------------------------
```

Greening of the CAP and crop diversification (introduced in 2013)

========= in progress ==========

The *CAP* evolves to better include environmental aspects such as protection of biodiversity (so-called greening the *CAP*). This is supported through incentives in favor of crop diversification. Diversification in itself depends on the number of crops actually planted in the farm. It is subsidized when this number satisfies thresholds prescribed in terms of arable area and total utilized agricultural area. Typically, subsidy benefits (i) small farms (e.g. arable land of less than 10 ha), (ii) medium farms (arable land between 10 and 30 ha) with respect to at least two planted crops and the most planted under 75% of arable land, (iii) bigger farms with respect to four conditions: (a) at least three crops planted, (b) the most planted of them under 75% of arable land, (c) the sum of the two most planted crops under 95% of arable land, (d) as an exception to (b) and (c) when the most important crop is temporary meadow or fallow, knowing that the second crop in area is under 75% of the rest of arable land except if this crop is temporary meadow or fallow. At least, there is no condition in terms of arable land coverage in two cases: (1) the total area devoted to fallow, temporary meadow and legume is over 75% of arable land, (2) the total area devoted to grasslands (permanent and temporary meadow) and rice is greater than 75% of the utilized agricultural area.

We consider that maize, beet and protein fodders are part of agricultural area, and that each of them counts as an individual crop. A part of the policy designed under *CAP* greening is implemented in the model (except the percentage limits), as shown by Table **??**.

It must be noticed that representative *AROPAj* farm groups possibly plant a number of crops lower or equal than the total number of crops planted by represented individual farms, which is a priori greater than any individual number. We can expect that implementation of CAP greening incentives may lead *AROPAj* to over estimate the subsidized area as well as the number

Table 12. *CAP* greening initiated in 2013 (17).

```
Added variables:
 xtk(j,k): total j-crop area in k-farm group
 itk(j,k): binary variable which equals 1 when xtk(j,k) is strictly positive
 nttjk(k): number of crops cultivated in the LP solution (of quasi-integer value but inserted as a real variable in the LP)
 xarab(k): arable area to be affected to a size category
 xdivb(k), xdivc(k), xdivd(k), xdive(k): eligible crop area according size and relaxing area, for diversification subsidizing
           xdivb < sdmin; sdmin < xdivc < sdmax; xdivd > sdmax; xdiv(b-d) subsidized under condition; xdive: relaxing variable
 idivb(k), idivc(k), idivd(k), idive(k): binary variable related to diversification limits
Added constraints:
 eld(m,k): crop diversification block
 ild(j,k,m): crop numbering block
Added parameters:
 sdmax(k), sdmin(k): min or max area limits
 sudiv(k): subsidy for crop diversification (k€/ha)
 ndivb, nidvc: numbers of crops requested according to the arable area size
 thrss(k): min limit for crop area when accounting for diversification
* ELEMENT refers to the matrix element label utilized in the AROPAj code
-------------------------------------------------------------------------------------------------------------------------------
------
            |egcdt|  xtk |xdivb |xdivc |xdivd |xdive |xarab | idivc | idivd | idivb | idive |itk(j,k)| nttjk | zf(k,l) |       RHS
            |(19) |(j,k) | (k)  | (k)  | (k)  | (k)  | (k)  |  (k)  |  (k)  |  (k)  |  (k)  |        |  (k)  |         |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 obj        | -1  |      |      |      |      |      |      |       |       |       |       |        |       |         |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 fog(19)    |  1  |      |sudiv |sudiv |sudiv |      |      |       |       |       |       |        |       |         | = |
            |     |      | (k)  | (k)  | (k)  |      |      |       |       |       |       |        |       |         |   |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 ELEMENT*            CEL   CEB    CED    CEE    CEF    CEA     CIC     CID     CIB     CIF     TII      TIN      CEM           TIZ
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(1,k)   |     |  -1  |      |      |      |      |   1  |       |       |       |       |        |       |         | < |
            |     |j elig|      |      |      |      |      |       |       |       |       |        |       |         |   |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(2,k)   |     |      |   1  |   1  |   1  |   1  |  -1  |       |       |       |       |        |       |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(13,k)  |     |      |   1  |      |      |      |      |       |       |       |       |        |       |-
sdmin(k)| < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(3,k)   |     |      |      |   1  |      |      |      |       |       |       |       |        |       |-
sdmax(k)| < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(4,k)   |     |      |      |  -1  |      |      |      |       |-99999 |-99999 |-
99999 |        |       | sdmin(k)| < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(10,k)  |     |      |      |      |  -1  |      |      |-99999 |       |-99999 |-
99999 |        |       | sdmax(k)| < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(5,k)   |     |      |      |   1  |      |      |      |-99999 |       |       |       |        |       |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(6,k)   |     |      |      |      |   1  |      |      |       |-99999 |       |       |        |       |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(11,k)  |     |      |   1  |      |      |      |      |       |       |-99999 |       |        |       |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(12,k)  |     |      |      |      |      |   1  |      |       |       |       |-
99999 |        |       |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(9,k)   |     |      |      |      |      |      |      |   1   |   1   |   1   |   1   |        |       |         | < |    1
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(7,k)   |     |      |      |      |      |      |      |       |       |       |       |   -
1   |   1   |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(8,k)   |     |      |      |      |      |      |      |       |       |       |       |0.999999|  -
1   |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(14,k)  |     |      |      |      |      |      |      | ndivb |       |       |       |        |  -
1   |         | < |
-------------------------------------------------------------------------------------------------------------------------------
-------------------
 eld(15,k)  |     |      |      |      |      |      |      |       | ndivc |       |       |        |  -
1   |         | < |
```

of subsidized farms. Nevertheless, this phenomenon does not prevent the model to show land cover change when CAP greening is implemented.

Other policy tools implemented in *AROPAj*

Among others, we introduced:

- "Agenda 2000", referring to the set of policies described above
- the "paquet Santer", referring to the set of policies described above
- "top-ups" (applied when East and Central European countries joined the *EU*)
- energy crops support when cropped on part of set-aside (not only sugar beet, shown above)
- many tax systems, on *GHG* emissions and/or *N*-losses, on (marketed) inputs (fertilizers), on products (for instance when accompanying the "A" and "B" sugar beet quotas)

4.2.5 Greenhouse gas emissions

Overview

Agricultural activities contribute directly to *GHG* emissions through five main processes (**?**): nitrous oxide (N_2O) emissions from agricultural soils; N_2O emissions from manure management; methane (CH_4) emissions from manure management; CH_4 emissions from enteric fermentation; CH_4 emissions from rice cultivation. N_2O emission from agricultural soils can be further disaggregated according to nitrogen inputs to soils: use of synthetic fertilisers, manure application, biological nitrogen fixation, and crop residues.

The *AROPAj GHG* accounting method is consistent with the approach recommended by the *IPCC* for the submission of national inventories to the *UNFCCC*. In this inventory-based approach, only emissions emitted within the agricultural sector are accounted for. For instance, emissions caused by the production of inputs and capital goods and the transport of food and feed products are not accounted for. This approach also ignores emissions related to the use of fossil fuel in agricultural production. The latter are indeed accounted for in the *IPCC* energy use category, and represent a relatively minor agricultural source compared to CH_4 or N_2O emissions. Furthermore, in accordance with international agreements on climate change, non-anthropogenic sources – e.g. N_2O background emissions by agricultural soils – are ignored.

In the model, each source of *GHG* emissions is linked to the relevant activity variables at the representative-farm level. The *GHG* budget is thus established at this resolution. This emissions accounting method closely resembles the method proposed by the *IPCC* guidelines (**?**), thereby facilitating comparison of emission inventories among countries. In general terms, the computation of *GHG* emissions relies on linear relationships between emissions and activity data through the use of emission factors. As listed in the table below, a total of 11 emission sources are computed within the model. Emissions of nitrous oxide are divided into eight subsources: direct agricultural soil emissions (4), indirect agricultural soil emissions (2), emissions from grazing animals (1) and manure management (1). Emissions of methane are disaggregated into three sub-sources: manure management, enteric fermentation, and rice cultivation. The first two sub-sources are further disaggregated into six animal categories (dairy cattle, non-dairy cattle, sheep, goats, swine, and poultry). This level of disaggregation facilitates comparisons with the *GHG* inventories as reported in the national communications to the *UNFCCC*.

Table 13. Summary of *GHG* emission sources accounted for in the model.

Emission sources	Activity variable	Comments
N_2O agricultural soils		
Direct emissions		
Use of synthetic fertilisers	Quantity of N synthetic fertilizer used on the farm	Fixed or endogenous per ha N application rate
Manure application	Total nitrogen excreted by (housed) animals	Country- and animal-specific *N* excretation rate
Crop residues	Crop output	Crop- and country-specific *N* content
Biological *N* fixation	*N*-fixing crop output	Crop- and country-specific N content
Animal production	Total *N* excreted by grazing animals	Country- and animal-specific N excretation rate
Indirect emissions		
Atmospheric deposition	Total *N* application	Sum of above *N* inputs
Leaching and run-off	Total *N* application	Sum of above *N* inputs
CH_4 manure management	Total *N* excreted by (housed) animals	Country- and animal-specific shares of solid and liquid waste management systems
CH_4 enteric fermentation	Feed energy intake (cattle) or animal numbers (other categories)	Based on energy content of animal feed or country- and animal-specific emission factors
CH_4 rice area	Rice area	Country-specific per ha emission factors

Whenever available in the respective national communications to the *UNFCCC*, country-specific information is used. Otherwise, the *IPCC* default values are used. All emission factors are converted into carbon dioxide (CO_2) equivalent by using the respective Global Warming Potentials (GWP).

N_2O emissions from agricultural soils depend upon total nitrogen inputs. Depending on the version of the model, the per-hectare nitrogen application rate could fluctuate with input and output prices, or remain fixed for a given crop and a given representative farm. In the latter case, quantities of nitrogen applied to soils are driven by the optimal crop area mix. For each farm type k, per-hectare fertilizer expenditures for each crop are estimated from *FADN* data. For each crop and each country, two fertilizers are chosen among the commercial fertilizers listed in *FAOSTAT* and Eurostat databases. These databases cover the most commonly used fertilisers in each country. In addition, a mass ratio between the two fertiliser types is computed based on current standard agricultural practices for each crop. The prices and nitrogen content of the two fertiliser types are taken from the *FAOSTAT* and *Eurostat* databases. They are weighted according to their content-to-mass ratios in order to derive a representative composite fertilizer, thereby allowing the computation of per-hectare nitrogen applied to each crop at each farm type. It is important to note that this approach relies on constant per-hectare nitrogen inputs for each crop and each farm type.

Emission factors, as well as volatilization and leaching parameters are taken from each member state's national communication to the *UNFCCC*. As for biological fixation and nitrogen in crop residues, the values of relevant parameters-–such as nitrogen content, crop/residue ratio, and dry matter fraction—are also taken, when available, from the national communications or the IPCC defaults.

Methane emissions from both enteric fermentation and manure management depend on the energy content of feed intake for each animal category. In the simplest form of methane inventories, *IPCC* recommends to use average energy requirements for each animal category to derive methane

emissions. In short, this implies a constant energy intake for any given animal category, and therefore constant emission factors on a per-head basis. In this case, animal numbers are the only driver of methane emissions. The chosen approach is more general and more flexible. In the model, animal feeding is endogenous. The total energy intake by each animal category can thus be derived from the optimal quantity and composition of feed. Emissions are therefore computed using the (animal-category dependent) share of total energy intake by animal category lost as methane. As a result, methane emissions are driven, not only by animal numbers, but also by the composition of animal feeding.

Manure can be either applied to crops, deposited directly on soils by grazing animals, or stored/treated using different management systems. The total production of manure-related nitrogen is computed as the product of nitrogen content of manure – defined for each animal category – and the corresponding animal numbers. Nitrogen excretion average rates for each animal category are taken from the national communications or the *IPCC* defaults. Because of the lack of available data at regional or farm-type level, the shares of manure applied to crops, deposited on grassland, and handled under all management systems are also taken from the national communications, which only provide information at the country level. The country average share is applied to each farm type.

Examples of *GHG* emission assessment and abatement reduction costs are displayed in Section **??**.

We use this emission calculation method as our default because it presents balanced estimates for CH_4 and N_2O. Another means of computation is incorporated into *AROPAj* for nitrogen to yield functions and for nitrogen input to N-loss functions, as used in the studies presented in Section **??**.

We attempted the assessment of soil Carbon storage in the framework of the European *INSEA* project. This was based on estimates referring to average values per crop category (i.e. arable crops and grasslands) over the *EU*. The parameters value need to be updated and disaggregated in accordance with soil and climate conditions (this refers to the `FBCAR` constraint in the environment block, {**E**}, see section **??**).

Detailed calculation based on the IPCC guidelines

As outlined in Section **??**, we mainly refer to the so-called Tables 4 entitled "Sectoral Report for Agriculture" and sourced by national reports of *GHG* inventories and submitted to the *UNFCCC* (see the "Table 4" entry in Table **??**).

The Average CH_4 Conversion Rate (`Ym`) refers to the fraction of gross energy in feed converted to methane. The data is expressed for each of the six animal categories considered (e.g. dairy cattle, non-dairy cattle, sheep, goats, swine and poultry).

The CH_4 producing potential (`Bo`) represents the maximum methane producing capacity for manure.

The fractions of manure handled under systems, include: Anaerobic lagoon (`MSANAER`), Liquid system (`MSLIQUI`), Solid storage (`MSSOLID`), Dry lot (`MSDRYLO`), Pasture range paddock (`MSPASTU`), Other (`MSOTHER`) and Daily spread (`MSDAILYS`). These systems are divided by climate type.

The indicator `MSANADI` is not taken into consideration. Therefore, it equals to 0 for all Member States.

Because the indicators `MSOPENP` and `MSBURNF` were not in the national inventories submitted, they have been replaced by `MSOTHER` and `MSDAILYS` respectively.

For the following systems, i.e.: Anaerobic lagoon, Liquid system, Solid storage, Dry lot, Pasture range paddock, Other and Daily spread, the study uses the sum of parameters for each category of animals. This sum equals to 1 for each Member State. In the case of Italy, Spain and Portugal, we have considered both climates (cool and temperate). Therefore, we have calculated the weighted mean. In the case of Malta, the study uses the values for Cyprus.

Initially, the Methane Conversion Factor (`MCF`) was calculated without taking into account the animal categories. However, we have decided to create a `MCF` parameter by taking into account the animal type. The same goes for the fractions of manure handled under different systems ; in particular, `MCFOPENP` and `MCFBURNF` were replaced by `MCFOTHER` and `MCFDAILYS`. In the case of Italy, Spain and Portugal, we use the weighted mean.

The Nitrogen excretion (`NEX`) is expressed in kg $N/head/yr$ for each animal category.

Table 14. Relationships between the *IPCC* parameters and the *UNFCCC* inventories ("Table 4" refers to the *UNFCCC* report).

IPCC Parameters	Source: National Reports of *GHG* Inventories (*UNFCCC*)
Average CH_4 Conversion Rate (`Ym`)	Table 4.A (Enteric Fermentation)
CH_4 producing potential (Bo)	Table 4.B(a)s1 (CH_4 emissions from Manure Management)
Fractions of manure handled under systems: Anaerobic lagoon (`MSANAER`), Liquid system (`MSLIQUI`), Solid storage (`MSSOLID`), Dry lot (`MSDRYLO`), Pasture range paddock (`MSPASTU`), Other (`MSOTHER`), Daily spread (`MSDAILYS`)	Table 4.B(a)s2 (CH_4 emissions from Manure Management)
Methane Conversion Factor (`MCF`)	Table 4.B(a)s2 (CH_4 emissions from Manure Management)
Nitrogen excretion (`NEX`)	Table 4.B(b) (N_2O Emissions from Manure Management)
Emission factor per animal waste management system	Table 4.B(b) (N_2O Emissions from Manure Management)
Emission factor for rice production (`EFRIZ`)	Table 4.C (Rice Cultivation)
Emission factor for Direct Soil Emissions, Pasture, Range and Paddock Manure, and Indirect Emissions	Table 4.Ds1 (Agricultural Soils)
Fraction crop burned, Fraction crop used for fuel, Fraction of synthetic fertilizer N applied to soils that volatilizes as NH_3 and NO_x, Fraction of N from manure that volatilizes, Fraction used as product, Fraction of N input to soils that is lost through leaching and run-off	Table 4.Ds2 (Agricultural Soils)
Residue/Crop product ratio	Table 4.F (Field Burning of Agricultural Residues)
Dry Matter Fraction	Table 4.F (Field Burning of Agricultural Residues)

Other indicators included are: the emission factor per animal waste management system for anaerobic lagoon, liquid system, solid storage, dry lot and other; the emission factor for rice production and the emission factor for direct soil emissions; pasture, range and paddock manure; and indirect emissions.

The residue/crop product ratio, the dry matter fraction and the ash content are also included. The animal type highlights six animal categories used in the model (e.g. 1 – dairy cattle, 2 – non dairy cattle, 3 – sheep, 4 – goats, 5 – swine, 6 - poultry).

Relationships between the *IPCC* parameters and the *UNFCCC* inventories are summarized in Table **??**.

Computation of emissions in the various versions

Items related to the *GHG* computation block in the model are detailed below. Let us recall that a *LP* element is the result of crossing one activity (i.e. a primal variable) and one constraint (i.e. a dual variable).

The balance sheet refers to 21 emission lines, plus a carbon storage one. Constraints are labeled with `EMS` and activities with `EMCDT`.

primal var.	dual var.	GHG emission balance	(unit)
EMCDT(01)	EMS(01)	N2O - Agr soils - Direct em. synth fertilizers	(tCO2)
EMCDT(02)	EMS(02)	N2O - Agr soils - Direct em. manure application	(tCO2)
EMCDT(03)	EMS(03)	N2O - Agr soils - Direct em. by N fixing crops	(tCO2)
EMCDT(04)	EMS(04)	N2O - Agr soils - Direct em. by crops residues	(tCO2)
EMCDT(05)	EMS(05)	N2O - Agr soils - Animal production (pasture,...)	(tCO2)
EMCDT(06)	EMS(06)	N2O - Agr soils - Indirect em. N atmosph. redepos	(tCO2)
EMCDT(07)	EMS(07)	N2O - Agr soils - Indirect em. N from leaching	(tCO2)
EMCDT(08)	EMS(08)	N2O - Manure management	(tCO2)
EMCDT(09)	EMS(09)	CH4 - Manure management (Dairy cattle)	(tCO2)
EMCDT(10)	EMS(10)	CH4 - Manure management (Non-dairy cattle)	(tCO2)
EMCDT(11)	EMS(11)	CH4 - Manure management (Sheep)	(tCO2)
EMCDT(12)	EMS(12)	CH4 - Manure management (Goats)	(tCO2)
EMCDT(13)	EMS(13)	CH4 - Manure management (Swine)	(tCO2)
EMCDT(14)	EMS(14)	CH4 - Manure management (Poultry)	(tCO2)
EMCDT(15)	EMS(15)	CH4 - Rice cultivation	(tCO2)
EMCDT(16)	EMS(16)	CH4 - Enteric fermentation (Dairy cattle)	(tCO2)
EMCDT(17)	EMS(17)	CH4 - Enteric fermentation (Non-dairy cattle)	(tCO2)
EMCDT(18)	EMS(18)	CH4 - Enteric fermentation (Sheep)	(tCO2)
EMCDT(19)	EMS(19)	CH4 - Enteric fermentation (Goats)	(tCO2)
EMCDT(20)	EMS(20)	CH4 - Enteric fermentation (Swine)	(tCO2)
EMCDT(21)	EMS(21)	CH4 - Enteric fermentation (Poultry)	(tCO2)
EMCDT(22)	EMS(22)	C - Stockage carbone sols	(tCO2)

The writing of the *GHG* emission balance equations possibly refers to all crop and animal activities, as shown below (see also above the meaning of listed activities):

```
   SUM(N)       GES * EMCDT(N)
 - SUM(I,K,L)   AFS * ANI(I,K,L)
 - SUM(J,K)     CFS * XTK(J,K)
 - SUM(N,I,K,L) CSB * B(N,I,K,L)
 - SUM(I,J,K,L) CSC * C(I,J,K,L)
 - SUM(I,J,K,L) CSF * F(I,J,K,L)
 - SUM(I,K,L)   CSG * AGT(I,K,L)
 - SUM(J,K)     CDC * QEN(J,K)
 - SUM(J,K)     CDD * QON(J,K)
 - SUM(J,K,O)   CDP * PO(J,K,O)
 .EQ. 0.0
```

The objective function (`OBJ`) may be impacted when *GHG* emissions are priced, as follows:

```
  ...
 - SUM(N)       OES * EMCDT(N)
 - SUM(N)       FEG * EXCDT(N)
  ...
```

Matrix elements may be either directly computed through parameter calls or through functions. Major parameters displayed below belong to the "environment category" described above (see

section **??**). The functions used are displayed hereafter. Let us note that comment lines use "c" as 1st character of the line.

<u>matrix "ELEMENTS"</u>:

```
c:E: afs           : EMS(M) x ANI(I,K,L)
      afs = afsy(m,i,k)
c:E: ges           : EMS(M) x EMCDT(N)
      ges = gesy(m,n)
c:E: cfs           : EMS(M) x (X(J,K,L), XA-D(J,K,L))
      cfs = cfsy(m,j,k)
c:E: csb           : EMS(16-21) x B(N,I,K,L)
      csb = csby(m,n,i,j,k)
c:E: csc           : EMS(16-21) x C(I,J,K,L)
      csc = cscy(m,n,i,j,k)
c:E: csf           : EMS(16-21) x F(I,J,K,L)
      csf = csfy(m,n,i,j,k)
c:E: csg           : EMS(16-21) x AGT(I,K,L)
      csg = csgy(m,n,i,k)
c:E: cdc           : EMS(M) x QEN(J,K)
      cdc = 0.
     0IF (m.eq.1.and.zcc59(j,k).le.0.) cdc=zcc34*(1-zcc35)
     0IF (m.eq.1.and.zcc59(j,k).gt.0.) cdc=zcc59(j,k)
     0IF (m.eq.6) cdc=zcc08*zcc35
     0IF (m.eq.7) cdc=zcc09*zcc10
     0IF (m.eq.1) cdc=cdc*44./28.*prgn*(fnes(j,k)+fnesc(j,k))
c:E: cdd           : EMS(M) x QON(J,K)
      cdd = 0.
     0IF (m.eq.6) cdd=zcc08*zcc35
     0IF (m.eq.7) cdd=zcc09*zcc10
c:E: cdp           : EMS(M) x PO(J,K,O)
      cdp = 0.
     0IF (m.eq.1.and.zcc59(j,k).le.0.) cdp=zcc34*(1-zcc35)
     0IF (m.eq.1.and.zcc59(j,k).gt.0.) cdp=zcc59(j,k)
c deja pris en compte avec QON pour ems 6 et 7
     0IF (m.eq.1) cdp=cdp*44./28.*prgn*fnedo(o,k)
...
c:E: oes           : OBJ x EMCDT(N)
      oes = prges
...
c:E: feg           : (OBJ QUANT(K,L)) x EXCDT(N)
      feg = 0.
     0IF (n.eq.1) feg = 1.
     0IF (n.eq.13.and.pexne.eq.0.) pexne = 0.000001
     0IF (n.eq.13) feg = - pexne
     0IF (n.eq.6) feg = 1.
```

<u>functions used in *GHG* balance "elements"</u>:

```
c:F: afsy()        : afs                                    (tCO2/head/an)
      function afsy(im,ii,ik)
      afsy=0.

c     Descr: N2O - Agr soils - Direct em. manure application
c     Source: Equations 4.20 4.24 Good Practices
      if (im.eq.2) then
         if (zcc52.eq.1.) then
            afsy =zcc01*(zcc02(ii)/1000.)*(1.-zcc03)*(1.-zcc04
     s           -zcc05)*prgn*44./28.
            if (ii.eq.55) afsy = afsy/1000.
         endif
      endif
```

```
      if (im.eq.5) then
         if (zcc53.eq.1.) then
            afsy = zcc06*(zcc02(ii)/1000.)*zcc07(ii)*prgn*44./28.
            if (ii.eq.55) afsy = afsy/1000.
         endif
      endif

c     Descr: N2O - Agr soils - Indirect em. N atmosph. redepos (manure
C     applicat)
c     Source: Equations 4.20 4.21 4.31 4.32 Good Practices
      if (im.eq.6) then
         if (zcc54.eq.1.) then
            afsy = zcc08*(zcc02(ii)/1000.)*zcc03*prgn*44./28.
            if (ii.eq.55) afsy = afsy/1000.
         endif
      endif

c     Descr: N2O - Agr soils - Indirect em. N from leaching (manure
C     application)
c     Source: Equations 4.20 4.21 4.35 4.36 Good Practices
      if (im.eq.7) then
         if (zcc55.eq.1.) then
            afsy = zcc09*(zcc02(ii)/1000.)*(1-zcc04)*zcc10
     s           *prgn*44./28.
            if (ii.eq.55) afsy = afsy/1000.
         endif
      endif

c     Descr: N2O - Manure management
c     Source: Equation 4.18 Good Practices
      if (im.eq.8) then
         if (zcc56.eq.1.) then
            afsy = (zcc11*zcc18(ii)*(zcc02(ii)/1000.)
     s           +zcc12*zcc19(ii)*(zcc02(ii)/1000.)
     s           +zcc13*zcc20(ii)*(zcc02(ii)/1000.)
     s           +zcc14*zcc21(ii)*(zcc02(ii)/1000.)
     s           +zcc15*zcc22(ii)*(zcc02(ii)/1000.)
     s           +zcc16*zcc23(ii)*(zcc02(ii)/1000.)
     s           +zcc17*zcc24(ii)*(zcc02(ii)/1000.))*prgn*44./28.
         endif
            if (ii.eq.55) afsy = afsy/1000.
      endif

c     Descr: CH4 enteric fermentation (young animals)
c     Source: Table A-2 p4.32 RM
c
      if (im.eq.17) then
          if (ii.le.14) then
            afsy = 0.033*prgch
          endif
      endif
      return
      end

c:F: cfsy()        : cfs                                        (tCO2/ha)
      function cfsy(im,ij,ik)
      cfsy = 0.

c     Descr: N2O - Agr soils - Direct em. synth fertilizers
c     Source: equations 4.20 4.21 4.22 Good Practices IPCC
      if (im.eq.1.and.zcc59(ij,ik).eq.0.) then
          cfsy=zcc34*(1-zcc35)*fazoty(ij,ik)*prgn*44./28.
      endif
      if (im.eq.1.and.zcc59(ij,ik).gt.0.) then
```

```
          cfsy=zcc59(ij,ik)*(1-zcc35)*fazoty(ij,ik)*prgn*44./28.
      endif
      if (im.eq.1.and.(endrg+mgesd).gt.0.) cfsy=0.

c     Descr: N2O - Agr soils - Direct em. by N fixing crops
c     Source: equations 4.20 4.26 4.27 Good practices IPCC
      if (im.eq.3) then
c     Selection des N-fixing crops
         if (fnxing(ij).eq.1.) then
c     Tier 1a eq 4.25 Good Practice
            if (zcc47.eq.1.) then
               cfsy=zcc36*2.*zcc38(ij)*rdy(ij,ik)*prgn*44./28.
            endif
c     Tier 1b eq 4.25 Good Practice
            if (zcc47.eq.2.) then
               cfsy=zcc36*(1+zcc37(ij))*zcc39(ij)
     s              *zcc38(ij)*rdy(ij,ik)*prgn*44./28.
            endif
         endif
      endif

c     Descr: N2O - Agr soils - Direct em. by crops residues
c     Source: equations 4.20 4.28 4.29 Good Practices IPCC
      if (im.eq.4) then
c     Tier 1a eq 4.28 Good Practice
         if (zcc48.eq.1.) then
            cfsy=zcc40*2.*zcc38(ij)*(1-zcc45)*(1-zcc41)
     s           *rdy(ij,ik)*prgn*44./28.
         endif
         if (zcc48.eq.2.) then
c     Tier 1b eq 4.29 Good Practice
            cfsy=zcc40*(1+zcc37(ij))*zcc39(ij)
     s           *zcc38(ij)*(1-zcc41-zcc42-zcc43
     s           -zcc44)*rdy(ij,ik)*prgn*44./28.
         endif
      endif

c     Descr: N2O - Agr soils - Indirect em. N atmosph. redepos (synth fert)
c     Source: equations 4.20 4.31 Good Practices IPCC
      if (im.eq.6) then
          cfsy=zcc08*zcc35*fazoty(ij,ik)*prgn*44./28.
      endif

c     Descr: N2O - Agr soils - Indirect em. N from leaching (synth fert)
c     Source: equations 4.20 4.34 Good Practices IPCC
      if (im.eq.7) then
          cfsy=zcc09*zcc10*fazoty(ij,ik)*prgn*44./28.
      endif
      if (im.eq.7.and.(endrg+mgesd).gt.0.) cfsy=0.

c     Descr: CH4 - Rice Production
c     Source: equations 4.34 Good Practices IPCC
c     Comme: prch : kg CO2.(kg CH4)-1
      if (im.eq.15) then
         if (ij.eq.19) then
            cfsy = zcc46*prgch/100.
         endif
      endif

c stockage carbone sols
      if (im.eq.22) then
         cfsy=-cesjk(ij,ik)*44./12.
      endif
      return
```

```
      end

c:F: csby()         : csb                                    (tCO2/t_alim)
      function csby(im,in,ii,ij,ik)
      csby=0.

c     Descr: CH4 - Manure management
c     Source: Equations 4.15 4.16 4.17 Good Practices
c     Comme: 0.67 kg.(m)-3 entre en dur
      if (zcc61(ii).eq.0.) zcc61(ii)=zcc26
      if (zcc62(ii).eq.0.) zcc62(ii)=zcc27
      if (zcc63(ii).eq.0.) zcc63(ii)=zcc28
      if (zcc64(ii).eq.0.) zcc64(ii)=zcc29
      if (zcc65(ii).eq.0.) zcc65(ii)=zcc30
      if (zcc69(ii).eq.0.) zcc69(ii)=zcc31
      if (zcc67(ii).eq.0.) zcc67(ii)=zcc32
      if (zcc70(ii).eq.0.) zcc70(ii)=zcc33
      if (im.ge.09.and.im.le.14) then
         if (im.eq.animalt(ii,8)) then
            csby = vsob(in,ii,ik)*zcc25(ii)
     s            *(zcc61(ii)*zcc07(ii)
     s             +zcc62(ii)*zcc18(ii)
     s             +zcc63(ii)*zcc19(ii)
     s             +zcc64(ii)*zcc20(ii)
     s             +zcc65(ii)*zcc21(ii)
     s             +zcc69(ii)*zcc22(ii)
     s             +zcc67(ii)*zcc23(ii)
     s             +zcc70(ii)*zcc24(ii))*0.67*prgch
         endif
      endif

c     Descr: CH4 - Enteric fermentation
c     Source: Gilles Le Moguedec (zcc50=1.)
c     Descr: CH4 - Enteric fermentation (zcc50=2.)
c     Source: Equations 4.14 Good Practices
c     Comme: 55.65 MJ/kg CH4 entre en dur
      if (im.ge.16.and.im.le.21) then
         if (im.eq.animalt(ii,15)) then
            if (zcc50.eq.1.) then
               csby = fmethane(1,in,ii,ij,ik)*prgch
            endif
            if (zcc50.eq.2.) then
               csby = ((brac(in,ii,ik)*smac(in,ii,ik))
     s                *(4.185/1000.)*zcc51(ii))*prgch/55.65
            endif
         endif
      endif
      return
      end

c:F: cscy()         : csc                                 (t CO2.(t alim-1))
      function cscy(im,in,ii,ij,ik)
      cscy=0.

c     Descr: CH4 - Manure management
c     Source: Equations 4.15 4.16 4.17 Good Practices
c     Comme: 0.67 kg.(m)-3 entre en dur
c     Descr: CH4 - Enteric fermentation (zcc50=2.)
c     Source: Equations 4.14 Good Practices
c     Comme: 55.65 MJ/kg CH4 entre en dur
      if (zcc61(ii).eq.0.) zcc61(ii)=zcc26
      if (zcc62(ii).eq.0.) zcc62(ii)=zcc27
      if (zcc63(ii).eq.0.) zcc63(ii)=zcc28
      if (zcc64(ii).eq.0.) zcc64(ii)=zcc29
```

```
      if (zcc65(ii).eq.0.) zcc65(ii)=zcc30
      if (zcc69(ii).eq.0.) zcc69(ii)=zcc31
      if (zcc67(ii).eq.0.) zcc67(ii)=zcc32
      if (zcc70(ii).eq.0.) zcc70(ii)=zcc33
      if (im.ge.09.and.im.le.14) then
         if (im.eq.animalt(ii,8)) then
            cscy = vsoc(ij,ii,ik)*zcc25(ii)
     s            *(zcc61(ii)*zcc07(ii)
     s             +zcc62(ii)*zcc18(ii)
     s             +zcc63(ii)*zcc19(ii)
     s             +zcc64(ii)*zcc20(ii)
     s             +zcc65(ii)*zcc21(ii)
     s             +zcc69(ii)*zcc22(ii)
     s             +zcc67(ii)*zcc23(ii)
     s             +zcc70(ii)*zcc24(ii))*0.67*prgch
         endif
      endif

c     Descr: CH4 - Enteric fermentation
c     Source: Gilles Le Moguedec
      if (im.ge.16.and.im.le.21) then
         if (im.eq.animalt(ii,15)) then
           if (zcc50.eq.1.) then
              cscy = fmethane(2,1,ii,ij,ik)*prgch
           endif
           if (zcc50.eq.2.) then
              cscy = ((ebes(ij)*smj(ij,ik))
     s             *(4.185/1000.)*zcc51(ii))*prgch/55.65
           endif
        endif
      endif
      return
      end

c:F: csfy()         : csf                                  (t CO2.(t alim-1))
      function csfy(im,in,ii,ij,ik)
      csfy=0.

c     Descr: CH4 - Manure management
c     Source: Equations 4.15 4.16 4.17 Good Practices
c     Comme: 0.67 kg.(m)-3 entre en dur
      if (zcc61(ii).eq.0.) zcc61(ii)=zcc26
      if (zcc62(ii).eq.0.) zcc62(ii)=zcc27
      if (zcc63(ii).eq.0.) zcc63(ii)=zcc28
      if (zcc64(ii).eq.0.) zcc64(ii)=zcc29
      if (zcc65(ii).eq.0.) zcc65(ii)=zcc30
      if (zcc69(ii).eq.0.) zcc69(ii)=zcc31
      if (zcc67(ii).eq.0.) zcc67(ii)=zcc32
      if (zcc70(ii).eq.0.) zcc70(ii)=zcc33
      if (im.ge.09.and.im.le.14) then
         if (im.eq.animalt(ii,8)) then
            csfy = vsof(ij,ii,ik)*rdy(ij,ik)*zcc25(ii)
     s            *(zcc61(ii)*zcc07(ii)
     s             +zcc62(ii)*zcc18(ii)
     s             +zcc63(ii)*zcc19(ii)
     s             +zcc64(ii)*zcc20(ii)
     s             +zcc65(ii)*zcc21(ii)
     s             +zcc69(ii)*zcc22(ii)
     s             +zcc67(ii)*zcc23(ii)
     s             +zcc70(ii)*zcc24(ii))*0.67*prgch
         endif
      endif

c     Descr: CH4 - Enteric fermentation
```

```
c     Source: Gilles Le Moguedec
c     Descr: CH4 - Enteric fermentation (zcc50=2.)
c     Source: Equations 4.14 Good Practices
c     Comme: 55.65 MJ/kg CH4 entre en dur
      if (im.ge.16.and.im.le.21) then
         if (im.eq.animalt(ii,15)) then
           if (zcc50.eq.1.) then
              csfy = fmethane(2,1,ii,ij,ik)*rdy(ij,ik)*prgch
           endif
           if (zcc50.eq.2.) then
              csfy = ((ebes(ij)*smj(ij,ik)*rdy(ij,ik))
     s             *(4.185/1000.)*zcc51(ii))*prgch/55.65
           endif
        endif
      endif
      return
      end

c:F: csgy()         : csg                                  (t CO2.(t alim-1))
      function csgy(im,in,ii,ik)
      csgy=0.

c     Descr: CH4 - Manure management
c     Source: Equations 4.15 4.16 4.17 Good Practices
c     Comme: 0.67 kg.(m)-3 entre en dur
      if (zcc61(ii).eq.0.) zcc61(ii)=zcc26
      if (zcc62(ii).eq.0.) zcc62(ii)=zcc27
      if (zcc63(ii).eq.0.) zcc63(ii)=zcc28
      if (zcc64(ii).eq.0.) zcc64(ii)=zcc29
      if (zcc65(ii).eq.0.) zcc65(ii)=zcc30
      if (zcc69(ii).eq.0.) zcc69(ii)=zcc31
      if (zcc67(ii).eq.0.) zcc67(ii)=zcc32
      if (zcc70(ii).eq.0.) zcc70(ii)=zcc33
      if (im.ge.09.and.im.le.14) then
         if (im.eq.animalt(ii,8)) then
            csgy = vsog(ii,ik)*zcc25(ii)
     s          *(zcc61(ii)*zcc07(ii)
     s           +zcc62(ii)*zcc18(ii)
     s           +zcc63(ii)*zcc19(ii)
     s           +zcc64(ii)*zcc20(ii)
     s           +zcc65(ii)*zcc21(ii)
     s           +zcc69(ii)*zcc22(ii)
     s           +zcc67(ii)*zcc23(ii)
     s           +zcc70(ii)*zcc24(ii))*0.67*prgch
      endif

c     Descr: CH4 - Enteric fermentation
c     Source: Gilles Le Moguedec
c     Descr: CH4 - Enteric fermentation (zcc50=2.)
c     Source: Equations 4.14 Good Practices
c     Comme: 55.65 MJ/kg CH4 entre en dur
      if (im.le.16.and.im.le.21) then
         if (im.eq.animalt(ii,15)) then
            if (zcc50.eq.1.) then
               csgy = fmethane(3,1,ii,ij,ik)*prgch
            endif
            if (zcc50.eq.2.) then
            csgy=brag(ii,ik)*smsag(ii,ik)*(4.185/55650.)*zcc51(ii)*prgch
            endif
         endif
      endif
      return
      end
```

```
c:F: gesy()        : ges
      function gesy(im,in)
      gesy=0.
      if ((im.ge.1.and.im.le.22).and.(in.eq.im)) gesy=1.
      return
      end
```

By default parameter values are null, and positive values are displayed below.

```
zcc01 :  0.01250000 ::    zcc17 :  0.00700000 ::    zcc36 :  0.01250000
zcc03 :  0.10000000 ::    zcc26 :  0.01500000 ::    zcc40 :  0.01250000
zcc06 :  0.02000000 ::    zcc27 :  0.01500000 ::    zcc46 : 36.00000000
zcc08 :  0.01000000 ::    zcc28 :  0.01500000 ::    zcc47 :  2.00000000
zcc09 :  0.02500000 ::    zcc29 :  0.35000000 ::    zcc48 :  2.00000000
zcc10 :  0.30000000 ::    zcc30 :  0.90000000 ::    zcc50 :  2.00000000
zcc11 :  0.02117377 ::    zcc31 :  0.01000000 ::    zcc52 :  1.00000000
zcc12 :  0.02117377 ::    zcc32 :  0.10000000 ::    zcc53 :  1.00000000
zcc13 :  0.00103546 ::    zcc33 :  0.10000000 ::    zcc54 :  1.00000000
zcc15 :  0.00100000 ::    zcc34 :  0.01250000 ::    zcc55 :  1.00000000
zcc16 :  0.00100000 ::    zcc35 :  0.10000000 ::    zcc56 :  1.00000000
```

4.2.6 Nitrogen

Implementation of the N block and call to N to yield functions

The nitrogen block operates when the related flag parameter is activated (`endrg=1.`, `0.` by default). The part referring to manure is included when the related flag parameter is activated (`mgesd=1.`, `0.` by default). All parameters, activities and constraints relating the N block are listed above (see Sections **??** and **??**).

When this block is not activated, nitrogen comes only through fertilizer costs. When the block is activated, the basic elements refer to the nitrogen requirement of crops and to the nitrogen content of mineral and organic fertilizers. By default, parameters refer to mineral fertilizer prices and nitrogen content ratio of fertilizers. When manure accounts for nitrogen intakes and impacts the nitrogen balance for plants, manure-to-animal relations are parametrized.

In the case of crops for which farm-dependent nitrogen to yield functions are available, the market price ratio supplies the value of required nitrogen and its corresponding yield. These values optimize the plot gross margin and the farm gross margin when a part of crop and a part of fertilizer are marketed. In the case of on-farm use and non-market goods, shadow prices have to replace market prices (see Section **??** for discussion and quoted papers for demonstration).

Parameters related to these dose response functions are provided by external data (`lisscal.don` file called by `intern.f`). Organic nitrogen accounts for fertilizing with values expressed in mineral nitrogen equivalence and delivered by additional data (included in the parameter file, `mgg.par`). Optimal yields and nitrogen inputs in generic equivalent terms are calculated for any crop and farm group occurring in the `lisscal.don`. This is calculated through a subroutine called by the *FORTRAN* `intern.f` file.

Urban organic residues may be added in the process, as explained in section **??**. Related parameters and primal and dual variables are integrated in the description of the block. When no iterative *AROPAj* runs are realized, with the aim of assessing the "long term" impact of regular spreading of organic matter, manure and *UOR* as well are viewed as "short term" substitution sources of mineral fertilizer. Two categories of urban wastes are now referenced in the model: green waste compost (GWC) and urban sewage sludge (USS).

The comprehensive *MP* block

The bio-technical and economical parts detailed here refer to `N`, `E` and `P` blocks, which were described in Section **??**.

The technical side accounts for:

- the nitrogen balance required by each crop
- the organic nitrogen source related to livestock (per animal category)
- the organic nitrogen balance per *UOR* source
- the nitrogen balance referring to all organic sources (manure and *UOR*), in terms of equivalent mineral nitrogen
- the balance of organic nitrogen spread per crop
- the *N* pollutant balance per pollutant category (see section **??**)
- the average organic spreading computation aiming at implementing the long term impact of organic matter spreading

The economic side accounts for:

- organic spreading limit per unit of crop area
- implementation of various other environmental policies (regarding *GHG* and *N* pollutants)

It must be stated that several *LP* activities referring to nitrogen are defined at the farm level, while the yields and nitrogen inputs mentioned above refer to calculation per unit of area. This type of procedure allows us to maintain a basic linear structure for the model.

The primal and dual activities involved in the nitrogen block are listed below:

primal	dual	nitrogen block description	(unit)
po(j,k,o)		UOR quantity of type o used by crop j	(t/ha)
qen(j,k)		quantity of regular fertilizer used on j-crop area	(t/ha)
q(i,j,k,o)		quantity of manure type o / i-herd sourced	(t m3)
qon(j,k)		useful N from manure & UOR spread on area j	(tN/ha)
qonlt(k)		org. N average quantity brought on long term per	(tN/ha)
xo(o,i,k)		exceeding unused organic N type o i-anim. sourced	(tN)
xtk(j,k)		aggregated j-crop area	(ha)
ani(i,k)		animal number referring to category i	(heads)
epcdt(n)		balance N-pollutant emission (3 cat.: N2O NO3 NH3)	(tN)
	apnlt(k)	calculation of average per ha organic spreading :	(tN)
	ba(j,k)	N requirement to reach given yield	(tN)
	cnvol	sum of mineral fertilizer inputs	(tfert)
	cnorg	sum of organic N intakes in terms of equiv. min. N	(tN)
	ei(o,i,k)	o-type manure provided by i-herd	(tRM)
	epro(o)	use of UOR disposal	(tN)
	eqj(j,k)	balance spreading - source of available N	(tN)
	epn(m)	balances of N pollut. emis. originated by crops :	(N)

Nitrogen mineral fertilizers impact the objective function through the `qen(j,k)`, and there is the possibility for *UOR* pricing (negative in case of subsidizing). In addition, environmental policy tools can easily be introduced, such as taxes either referring to inputs (or outputs) or to *N*-losses (`epcdt(n)`).

The sub-matrix is displayed on Table **??**.

Table 15. Sub-matrix referring to technical parameters and *LP* variables operating for nitrogen.

```
The term "yexjk(j,k)" is the optimal amount of mineral N equivalent per hectare in case of N-to-yield function
We use the functions: qjmy(o,i,j,k) = fned(o,i,k)*coo(o,i,j,k)*(1-volat(o,k))
                      qjoy(o,j,k)   = fnedo(o,k)*cpro(o,j,k)*(1-volat(o,k))
-------------------------------------------------------------------------------------------------------------------
----------
          |  q(i,j,k,0)  | qonlt(k)  |qon(j,k)|    qen(j,k)     |  xtk(j,k)  | ani(i,k,l) | po(j,k,o)  |xo(o,i,k)| zf(k,l) |     RHS
-------------------------------------------------------------------------------------------------------------------
-------------------
 obj      |              |           |        |if(mgesd+endrg>1)|            |            |            |         |         |
          |              |           |        |  -prxnc(nj,nk)  |            |            |            |         |         |
          |              |           |        |   *(1+tegkk(k)) |            |            |            |         |         |
          |              |           |        |  -fnesc(nj,nk)  |            |            |            |         |         |
          |              |           |        |   * pgesn       |            |            |            |         |         |
          |              |           |        |else gjk(j,k)    |            |            |            |         |         |
-------------------------------------------------------------------------------------------------------------------
-------------------
 ei(o,i,k)|      1       |           |        |                 |            |     -
|           |    1    |         | < | 0
          |              |           |        |                 |            |qidej(o,i,k)|            |         |         |   |
          |              |           |        |                 |            | *epd(o,i,k)|            |         |         |   |
          |              |           |        |                 |            | *flf(o,i,k)|            |         |         |   |
-------------------------------------------------------------------------------------------------------------------
-------------------
 eqj(j,k) |      -       |           |    1   |                 |            |            |     -
|         |         | < | 0
          | qjmy(o,i,j,k)|           |        |                 |            |            |qjoy(o,j,k) |         |         |   |
-------------------------------------------------------------------------------------------------------------------
-------------------
 epro(o)  | fned(o,i,k)  |           |        |                 |            |            |     1      |         |    -
| < | 0
          |              |           |        |                 |            |            |            |         |proo(o,k)|   |
-------------------------------------------------------------------------------------------------------------------
-------------------
 ba(j,k)  |              |           |  - 1   |  - fnesc(j,k)   | benjk(j,k) |            |            |         |         | < | 0
          |              |           |        |                 | * rjk(j,k) |            |            |         |         |   |
          |              |           |        |                 |- aqolt(k)  |            |            |         |         |   |
          |              |           |        |                 |     or     |            |            |         |         |   |
          |              |           |        |                 | yexjk(j,k) |            |            |         |         |   |
          |              |           |        |                 |-qengf(j,k) |            |            |         |         |   |
          |              |           |        |                 | *fnesf(j,k)|            |            |         |         |   |
          |              |           |        |                 |- aqolt(k)  |            |            |         |         |   |
-------------------------------------------------------------------------------------------------------------------
-------------------
 epa(j,k) | qjmy(o,i,j,k)|     1     |        |                 |-yxdej(j,k) |            |qjoy(o,j,k) |         |         | < | 0
-------------------------------------------------------------------------------------------------------------------
-------------------
 apnlt(k) | -apolt(o,k)  | saukl(k,l)|        |                 |            |            |-
apolt(o,k) |         |         | < | 0
          | or 0 if o>10 |*popul(k,l)|        |                 |            |            |or 0 if o<10|         |         |   |
-------------------------------------------------------------------------------------------------------------------
-------------------
 ems()    |     ...                              see the GHG block
-------------------------------------------------------------------------------------------------------------------
-------------------
```

Calculation of nitrogen pollutant emissions

There is no default evaluation of nitrogen losses in the *AROPAj* model. I must be noted that N_2O emission can also be estimated within the *GHG* emission blockl, so as long as the estimates are *IPCC* sourced. Related *LP* variables are activities (`epcdt`) and constraints (`epn`). Emissions refer to crop areas when no N-to-yield function applies ; otherwise they refer to spread N inputs (`qen(j,k)`, `q(i,j,k,o)`, `po(j,k,o)`).

When nitrogen-to-yield functions are available and the complete nitrogen block is activated in the model, it is possible to assess N-losses in three forms, namely nitrous oxide (N_2O), ammonia (NH_3) and nitrate (NO_3). Such assessment capacity is due to the use of the *STICS* crop model simulations and regression calculation (see Section **??**). All that is needed to make a nitrogen-to-loss function operational is the ability to give the function through *AROPAj* parameters, knowing that the function may take any analytical form or point form.

The linear form is based on usual linear regression estimates when N crop intake is within the usual range of values ([50,400] kg$N$/ha). When the related nitrogen to yield function is estimated with the *STICS* model, six additional parameters for each crop and each farm group are needed. There is still no default emission function for the subset of farm groups and crops not referring to N-to-yield functions. But *EU* agricultural regions are well covered by these functions in terms of area as well as production (see Section **??**). This is particularly true in regions intersecting the Seine river basin (see Section **??**).

4.2.7 Water

The water block is operating in the model since 2017. Basically it runs like the nitrogen block when related to input quantities and crop yields.

Water-activities, in the sense of *AROPAj*, refer to irrigation. When the climatic dimension matters, natural resources (rainfalls) are taken into consideration through parameters of yield functions.

4.3 Non linearities within the *LP* framework

> Key-points:
>
> Non linearities within the *AROPAj* approach
> Advantage of keeping the *LP* framework
> Operational proxy solution
> Going further with loop computation

First, non linearities, *scricto sensu*, come from the use of integer (and usually binary) variables. This is mainly in the case of *CAP* tool implementation, when threshold conditions enter the policy. The solving process is carried out by solvers proposed by the software (e.g. *MIP* modules used by *GAMS*).

The major case to which we refer here is the substitution of *AROPAj* parameters by the so-called dose-response functions. This case occurs when a nitrogen-to-yield point is replaced by a nitrogen-to-yield function. When this function is linear, the *MP* farm group model is quadratic

instead of linear. The problem is better solved when the function is monotonously increasing and quasi-concave (see Section **??**, and Figure **??** which illustrates effects of replacing constant yields by response functions).

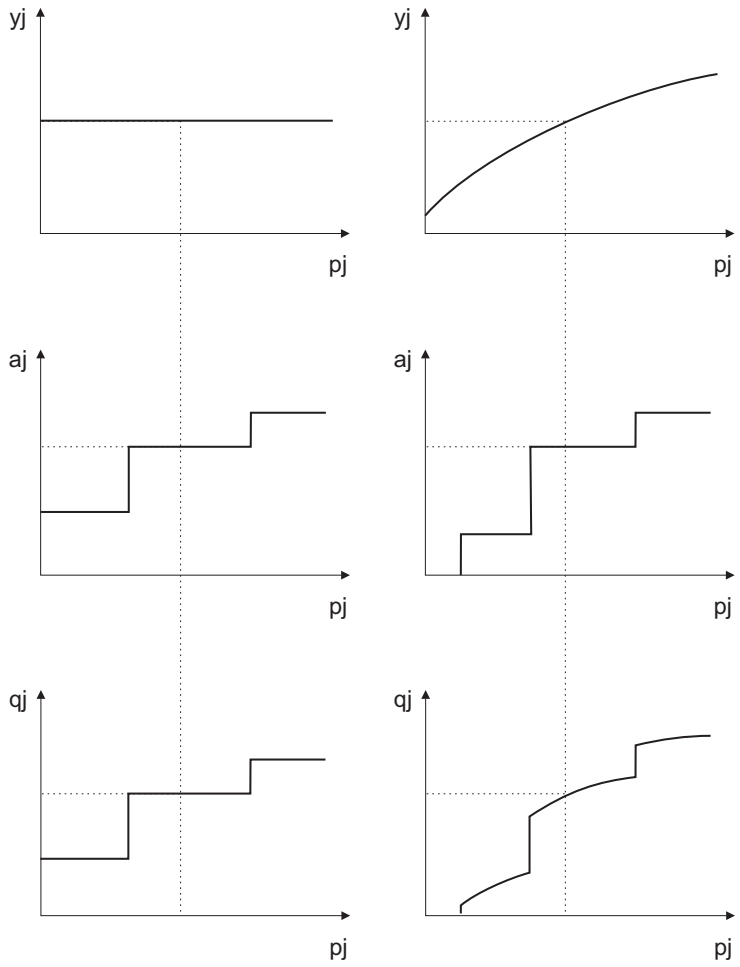


Figure 5. Stylized effects of point (on the left) by response function (on the right) replacement on the optimal yield (`yj`), allocated area (`aj`) and production (`qj`) related to the j crop when its price (`pj`) -and only this price- varies. The reference optimally calibrated LP solution is materialized by dotted lines.

This ensures that there is a unique solution that depends primarily on the nitrogen and crop price ratio. Some refinements are required when the farm group's optimal solution is made of non-marketed crop output (or of non-marketed nitrogen input). In other words, this occurs when the entire crop product feeds the livestock or when livestock manure is sufficient to supply the crop with the nitrogen required. In this case, market prices have to be replaced by relevant shadow prices which are themselves a part of the optimal solution.

The process of calculation of shadow prices is an iterative process that entails solving the model and re-estimating the shadow prices until a converging solution is found. This methodology extends and enhances the work done by **?**, which was previously developed for crop production entirely used on-farm as feedingstuff for livestock (e.g. cereal).

In this case, the optimal solution leads to a substituting of market prices with shadow prices in

the yield and nitrogen calculations. These shadow prices, which are required in order to calculate the yield and nitrogen prices, are provided by the linear programming model. The iterative process consists in using at each iteration shadow prices calculated in the previous iteration.

This two-step process starts with the given market prices, yields, and N-inputs that are computed at the plot level using the first order conditions. Then, these yields and N-inputs enter the linear programming model, which, once solved, provides the shadow prices (equal to market prices when optimal production is at least partly marketed). In the case of grass productions, where no market price exists, fodder prices are set at 0.000001 at the first iteration. The addition of multiple response functions and the complex interaction between farming and breeding activities might lead linear program solutions to be unstable. Finally, the full application to the *AROPAj* model appears to be satisfactory in that convergence or weak amplitude oscillation is obtained. Nevertheless, that needs a greater number of iterations than required when only marketed crop response functions were implemented.

5 Typology and farming system clustering

5.1 Clustering methodology

Key-points:

Regional representativity

From *FADN* farm types to *AROPAj* farm groups

Metrics for clustering

The information in the *FADN* database consists of surveys of individual farm samples. These individual farms are grouped by a clustering method to form a group of similar farms that present similar characteristics and thus can form a modeling unit for the *AROPAj* model: the *farm type* (also denominated by *farm group*).

The *FADN* data are presented at the regional level without any more localization information. Consequently, the farm types that are created belong to a specific region, that approximately corresponds to the *NUTS*-2 level according to the European classification (Nomenclature of Units for Territorial Statistics). For instance, in France the *NUTS*-2 classification corresponds to the Administrative Units.

Table 16. Organizational form under the *FADN* classification (simplified version when 2 digits).

Code	Type of farm
13	Cereals, Oleagineous, protein crops
14	Other field crops
2	Horticulture
3	Viticulture
411	Bovines milk
412	Bovines meat
42	Bovines livestock for meat
43	Bovines for milk and meat
44	Ovines, caprines and other herbivorous
50	Grain eater livestock
72	Livestock (mainly grain eater)
60	Various crops
71	Livestock (mainly herbivorous)
81	Field crops and herbivorous
82	Other livestock/field crop combinations

Within these regional units, the individually sampled farms are aggregated by a clustering method under the condition that each farm type must gather at least 15 different farms (to keep the information fully anonymous). The farms are grouped according to different key variables: organisational form, altitude, irrigation and economic size. Organisational form (*FADN* variable `A18`, also called *OTEX* or *FT*) refers to a *FADN* classification of the farms depending on the type of agricultural production, from field crops farms to intensive livestock farming. Table **??** presents the various values that this variable take (note that horticulture and viticulture are not included in *AROPAj*). There are three values that correspond to altitude zone (*FADN* variable `A41`): 1 for farms below 300 meters, 2 for farm between 300 and 600 meters, and 3 for farms above 600

meters. The irrigation variable was calculated as the ratio of irrigated crop areas divided by the total amount of crop areas. The economic size (*FADN* variable `A27`) is an integer between 1 and 10 corresponding to the size of the sampled farm.

A good clustering method is a method that maximizes the similarity within groups and the distance between different groups. We thus chose to perform the typical agglomerative hierarchical clustering method on these four variables. In order to do so, one must be able to calculate a distance between the farms. Each farm was placed on a four dimensional space, and the distance between two farms A and B could then be calculated according to the classical Euler formula:

$$Distance_{AB} = \sqrt{\sum_{i=1}^{N} (xA_i - xB_i)^2} \tag{1}$$

where N is the number of dimensions of the space (4 in our case) and xA_i and xB_i being the coordinates of the farms A and B in the dimension i. The calculation of the distance (a measure of similarity) is a subjective choice that belongs to the modeler, who can emphasize the importance of the variables of his choice.

Another important point is the transformation of the variables to a numerical equivalent that will be used to place the farms on each dimension. Table **??** presents the different variables that were used for the clustering and their arbitrary given value. We decided to put more weight to the organisational form since we intended to group the farms according to their type of agricultural production. The scale goes from 0 to 1200 with 0 being the more extensive (field crops) and 1200 being the more intensive (swine and poultry) type of farm. Irrigation and economic size were considered at a similar weight (scale ranging from 0 to 100) while altitude was considered as the less important variable (scale ranging from 10 to 30).

Table 17. Variables used by the hierarchical clustering procedure to create farm-types at the European Union level.

Organisational form

Code	Value
13	0
14	100
60	200
82	300
81	400
71	500
411	600
412	700
42	800
43	900
44	1000
50	1100
72	1200

Economic class

Code	Value
1	10
2	20
3	30
4	40
5	50
6	60
7	70
8	80
9	90
10	100

Irrigation

Irrigation ratio (%)	Value
0	0
10	10
20	20
30	30
40	40
50	50
60	60
70	70
80	80
90	90
100	100

Altitude

Code	Value
1	10
2	20
3	30

We now have the information to perform the clustering. The algorithm used was the following:

- **for** each region:
  - Place all farms on the four dimensional space. Each farm forms a unique farm-type at first.

- **while** there exists some farm types with less than 15 individual farms:
  - group the farm-type with the lowest number of individual farms to the closest corresponding farm-type using equation (**??**).
- recalculate the position of each farm-type in the space (i.e. the mean of the values of each farm belonging to the farm-type, for each dimension).

- repeat until each region is completed.

This algorithm was implemented with the *R GNU* software for the whole European Union. The types of output constructed are presented in Table **??** and Figures **??** and **??** for the French region Picardie, as an example. We can see that farm-types 10 and 11 present similar farms (same organisational form) but different economic size classes. Farm-type 14 gathers farms of different organizational forms because too few were surveyed to create a unique farm type.

Table 18. Typology created for the french region "Picardie".

farm-type	sampled farms	OTEX	mean economic class	mean altitude	ratio irrigation
10	50	13	7.98	1.08	0.00
11	61	13	9.11	1.00	0.00
12	21	14	8.00	1.00	0.00
13	45	14-60	9.31	1.00	7.88
14	35	43-50-81-411-412	8.29	1.17	0.00
15	50	81	9.14	1.00	0.00

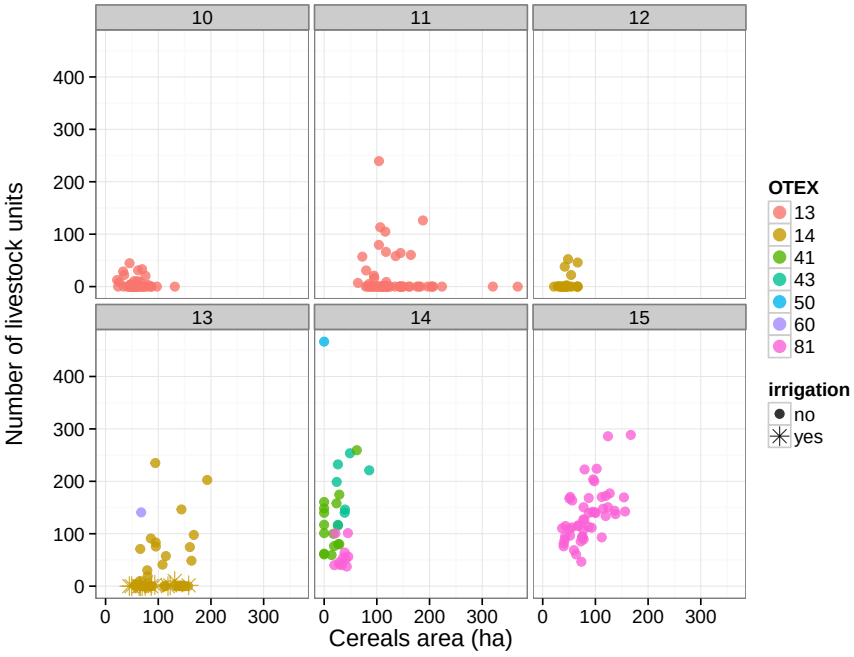


Figure 6. The 6 farm-types produced by the clustering method for the French region "Picardie" .

It has to be noted that the *AROPAj* farm type numbering is automatically processed by the algorithm. The *FADN* region index determines the order of farm clustering per country. When the

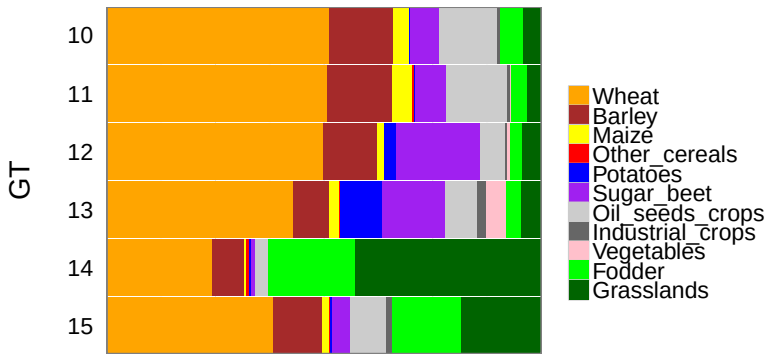


Figure 7. Types of crops grown by the 6 farm-types of the region "Picardie".

number becomes close to 99, the country is divided in two (or more) subcountries, and the farm type numbering begins again at 1 for this "new" country. For the use of the model, the *AROPAj* shell will automatically manage the calculation directories per sub-country and per farm type.

5.2 Farm groups throughout the *AROPAj* versions

Key-point:

AROPAj farm groups over the model versions

The *FADN* sample clustering into farm groups follows the evolution of the European Union size on the one hand, and the clustering criteria on the other hand. The key gap comes when the typology accounts for irrigation information and when we change the weigths associated to the different criteria in the clustering process (with the *V5* versions).

Table **??** delivers the typology, in terms of farm group number, at the *EU* country level for currently running versions of the model. Tables **??** and **??** delivers respective associated farm numbers and utilized agricultural areas, again per country. It must be noted that this data relates both on *AROPAj* activities and *FADN* farm types in the *AROPAj* typology. Some agricultural activities are not represented in the model (e.g. greenhouse productions and wine production, among a few others). Moreover, some *FADN* farm types may have *AROPAj* activities but nevertheless be excluded from the model. Tests are undertaken to gather information about missing areas with *AROPAj* activities: missing areas in the major part of countries are less than 2\and Greece where missing areas are close to 8\

Let us remark that even if the *FADN* is a noteworthy database and even if real agriculture may strongly change in some countries after their entrance in the *EU*, the figures in Table **??** show that farm sampling may be deficient in a few Member States. Thus, part of the *AROPAj* model's ''aggregation bias" is due to the *FADN*. We have to keep this in mind before attempting to identify bias in the *AROPAj* calibrating process (see Section **??**).

Croatia's entry into the EU in 2013 is taken into account in the 2013-2017 FADN years. Table **??** provides farm group numbers for all *FADN* years related to the *V5* version of the model.

Table 19. Farm group number per *EU* country throughout the different *AROPAj* versions.

country (*FADN*)	*V1* (1997)	*V2* (2002)	*V3* (2004)	*V4* (2009)	*V5* (2007)	*V5* (2008)	*V5* (2009)	*V5* (2010)	*V5* (2011)	*V5** (2012)
bel	15	27	24	22	23	25	26	30	29	30
bgr				29	45	49	48	61	61	65
cyp			9	5	5	8	9	10	11	11
cze			23	22	42	44	45	44	47	47
dan	10	22	21	21	31	29	29	33	34	32
deu	99	144	138	172	180	195	215	229	215	218
ell	45	54	59	56	86	90	93	80	90	106
esp	99	155	150	134	190	178	182	185	182	200
est			11	13	16	18	17	22	22	23
fra	131	157	144	145	164	159	157	164	169	170
gbr	51	59	64	66	74	73	74	74	71	78
hun			45	49	54	53	57	53	54	48
irl	8	22	23	16	20	19	18	17	20	18
ita	161	278	207	154	323	249	233	239	238	256
ltu			22	25	30	28	30	30	30	31
lux	5	13	11	10	7	7	7	7	8	8
lva			19	26	31	31	24	25	27	26
mlt				6	9	11	11	11	10	11
ned	11	19	18	16	24	22	24	25	24	23
ost	23	38	39	41	43	43	41	41	43	45
pol			176	195	222	222	215	200	205	209
por	43	37	32	36	51	48	46	46	47	48
rou				65	38	55	102	178	171	180
suo	18	20	21	26	24	24	26	22	26	26
sve	15	29	27	30	25	28	30	30	30	30
svk			13	13	19	18	18	18	18	18
svn			11	21	22	21	25	27	28	36
EU	734	1074	1307	1414	1798	1747	1802	1901	1910	1993

* Including overseas regions

Table 20. Farm number represented per *EU* country throughout the different *AROPAj* versions (thousands).

country (*FADN*)	*V1* (1997)	*V2* (2002)	*V3* (2004)	*V4* (2009)	*V5* (2007)	*V5* (2008)	*V5* (2009)
bel	36	34	30	27	27	27	27
bgr				78	91	93	102
cyp			8	8	8	8	8
cze			12	13	13	13	13
dan	48	44	37	31	31	31	31
deu	256	205	200	164	163	163	164
ell	273	198	209	201	271	276	261
esp	302	330	275	281	343	276	281
est			7	7	7	7	7
fra	320	314	299	287	286	287	287
gbr	124	118	94	85	85	85	85
hun			67	65	62	62	65
irl	129	77	69	56	107	106	106
ita	597	374	354	328	382	376	378
ltu			30	39	39	39	39
lux	2	2	2	1	1	1	1
lva			20	22	22	22	22
mlt				1	1	1	1
ned	65	51	42	46	47	47	47
ost	79	74	73	63	63	63	63
pol			438	450	703	697	704
por	209	79	72	67	66	68	67
rou				245	815	834	835
suo	49	43	42	39	38	38	39
sve	38	39	31	26	26	26	26
svk			3	3	3	3	3
svn			33	38	36	38	38
EU	2527	1981	2446	2672	3738	3688	3700

Table 21. Utilized agricultural area represented per *EU* country throughout the different *AROPAj* versions (thousands of hectares).

country (*FADN*)	*V1* (1997)	*V2* (2002)	*V3* (2004)	*V4* (2009)	*V5* (2007)	*V5* (2008)	*V5* (2009)
bel	1355	1432	1339	1393	1333	1380	1411
bgr				2797	2752	2781	2985
cyp			137	110	107	125	112
cze			3430	2987	3277	3153	3222
dan	2349	2437	2397	2489	2486	2521	2590
deu	14736	14728	15414	15145	14932	15306	15411
ell	1418	1220	1651	2053	2357	2443	2423
esp	11816	11329	10939	12903	13566	13008	13253
est			669	2249	837	869	886
fra	22972	24042	24321	24778	25065	25208	25118
gbr	11360	11647	10337	10290	10327	10322	10606
hun			3477	3656	3724	3727	3737
irl	4014	3198	3087	3003	3992	4072	4195
ita	8295	7552	8125	7145	9109	7848	7982
ltu			1470	1813	1907	1889	1869
lux	104	130	124	125	125	125	127
lva			1000	1159	1142	1157	1178
mlt				2	2	2	2
ned	1867	1511	1234	1648	1716	1728	1718
ost	1882	1788	1853	1771	1780	1807	1791
pol			7706	9823	12594	13254	13374
por	2253	1990	1716	1971	2032	2267	2124
rou				4527	7917	10100	10182
suo	1765	1889	1978	2046	2024	2084	2110
sve	3228	2639	2560	2406	2482	2472	2496
svk			1761	1724	1925	1899	1854
svn			430	434	437	434	438
EU	89415	87532	107153	120448	129943	131980	133193

Table 22. Farm group number per *EU* country throughout the different years of the *AROPAj-V5* version.

country (*FADN*)	*V5* (2007)	*V5* (2008)	*V5* (2009)	*V5* (2010)	*V5* (2011)	*V5** (2012)	*V5** (2013)	*V5** (2014)	*V5** (2015)	*V5** (2016)	*V5** (2017)	*V5** (2018)
bel	23	25	26	30	29	30	31	33	30	25	29	29
bgr	45	49	48	61	61	65	67	70	68	69	67	73
cyp	5	8	9	10	11	11	10	8	8	9	8	8
cze	42	44	45	44	47	47	43	44	44	45	41	41
dan	31	29	29	33	34	32	31	39	37	41	40	40
deu	180	195	215	229	215	218	224	223	215	214	215	217
ell	86	90	93	80	90	106	116	98	103	99	98	89
esp	190	178	182	185	182	200	193	193	191	199	189	191
est	16	18	17	22	22	23	18	20	21	20	22	19
fra	164	159	157	164	169	170	179	172	175	167	165	166
gbr	74	73	74	74	71	78	73	78	74	78	82	81
hrv							31	37	35	33	36	36
hun	54	53	57	53	54	48	48	49	48	50	48	45
irl	20	19	18	17	20	18	17	17	15	18	13	13
ita	323	249	233	239	238	256	243	225	200	206	228	235
ltu	30	28	30	30	30	31	26	30	27	29	30	31
lux	7	7	7	7	8	8	7	8	7	10	8	10
lva	31	31	24	25	27	26	26	24	26	25	26	26
mlt	9	11	11	11	10	11	10	11	10	11	11	10
ned	24	22	24	25	24	23	26	26	26	28	27	26
ost	43	43	41	41	43	45	41	44	43	40	39	41
pol	222	222	215	200	205	209	210	172	168	174	174	211
por	51	48	46	46	47	48	45	42	43	40	35	40
rou	38	55	102	178	171	180	189	116	142	180	181	157
suo	24	24	26	22	26	26	22	21	23	24	24	19
sve	25	28	30	30	30	30	32	31	32	34	30	31
svk	19	18	18	18	18	18	17	21	16	16	19	16
svn	22	21	25	27	28	36	31	29	25	26	31	27
EU	1798	1747	1802	1901	1910	1993	2006	1881	1852	1910	1916	1928

* Including overseas regions

6 Parameter estimation

6.1 Variable costs

Key-points:

Insights about new cost estimates
Application of General Maximum Entropy
Individual estimates

6.1.1 Estimation of activity variable costs

One of the main problems related to the use of the *FADN* database in farm programming models designed for policy analysis is the aggregated nature of cost data. This means that costs for various types of inputs are reported at the farm level but their distribution among the farm various activities is not available to the analyst.

Least Squares (LS) was historically the first method used for such cost disaggregation (**?**). However, LS may possibly return negative or non-significant cost coefficients that cannot be considered as valid estimates. Moreover, LS is suitable for estimating regional cost coefficients since multiple regression with cross-sectional data can only yield estimates of the average cost of input use for each activity across the sample: denoting activities by $i = 1, 2, \ldots, I$ and farms by $n = 1, 2, \ldots, N$, the estimating equation for the average cost per activity for any given input is written as $Y_n = \sum_i X_{ni} b_i + u_n$, where Y_n is the aggregate cost paid by farm n for the said input, X_{ni} stands for the level of activity i in the n-th farm, b_i is the average cost of the input for activity i and u_n is the error term. However, the cost coefficients in the case of farm models should not correspond to average values but should instead reflect the differences in production costs among the same activities in different farms. Cost disaggregation for each farm in the sample therefore requires the estimation of different β_{ni} coefficients so that the budget equation $Y_n = \sum_i X_{ni} \beta_{ni}$ is satisfied for any given input (i.e. $\beta_{ni} \neq b_i$). This means that for databases like *FADN*, the estimation of activity- and farm-specific costs constitutes an ill-posed problem with negative degrees of freedom because the number of the coefficients to be estimated ($I \times N$) is greater than the number of observations (N).

Since the problem of possible negative coefficients and ill-posedness cannot be handled by traditional econometric techniques, the literature suggests the use of information-theoretic statistical approaches like entropy and Bayesian methods that make use of prior information based on agronomic data, expert opinion, or other sources (e.g. **??**). Of similar reasoning to the above is the use of goal programming, as implemented by **?** for the estimation of activity costs in *FADN* farms in Thessaly, Greece, using information on costs obtained by a farm survey.

Although the use of exogenous information provides an answer to the problem of activity cost disaggregation, there still remains the problem of a source of homogeneous data at the EU level because available information, especially when it relates to agronomic practices, is usually limited to only a finite number of crops and geographical locales. In this section we present a method that is based on a two-step procedure which combines regression analysis with the Generalized Maximum Entropy estimator (*GME*). Specifically, we remain within the entropy framework of for the disaggregation of costs, as suggested by the literature, but we draw prior information from

an initial regression exercise that aims at providing an approximate estimate of variable costs per activity at the regional or higher geographical scales.

Applying the *GME* estimator in the linear model $Y_n = \sum_i X_{ni}\beta_{ni}$ involves finding an entropy maximizing probability distribution over a vector of support values for every β_{ni} coefficient, so that each coefficient is re-defined as the expected value of its corresponding unknown probability distribution. The support vectors are specified in the feasible range of the unknown parameters and are defined a priori by the researcher. In this sense they constitute the most subjective aspect of the entropy estimation procedure. In fact, the selection of the support values gives rise to two issues. The first is the choice of the central value of the support interval and the second is the width of the interval, i.e., the width of the interval around the central value. Especially the latter can greatly affect the estimation procedure because the coefficients are not invariant to the chosen support intervals (**?**); narrow bounds can greatly increase the mean square error of the *GME* estimator (**?**), or even render the entropy maximization program infeasible.

The first step for estimating β_{ni} with *GME* is to construct an appropriate vector of support values. For this purpose, we estimate the equation $Y_n = \sum_i x_{ni}b_i + u_n$ in order to obtain the $\hat{b}_i$ coefficient and the residual e_n. The underlying hypothesis is that for a given cost type t, the coefficient β_{ni} for activity i in farm n should be located in an interval around the average cost $\hat{b}_i$ and the interval must also contain part of the residuals e_n. We can therefore redefine the unknown parameter as $\beta_{ni} = \alpha_{ni} + \upsilon_{ni}$, where α_{ni} is a parameter derived from the distribution of $\hat{b}_i$, whereas υ_{ni} represents an error term that can be defined with respect to e_n. Although α_{ni} is directly related to the distribution of β_{ni}, and thus its meaning is straightforward, the use of an error term merits justification. There are two reasons that require the definition of β_{ni} as the sum of two separate parameters, one of which is an "error" term. The first reason is the aggregation of explanatory variables (production activities) into general categories. For instance, some *FADN* activities are already aggregates of other individual activities which cannot be modelled separately for different technical or sampling reasons. Typical examples are the various age categories in livestock production, flowers, seed crops, fodder root crops and all subcategories indexed as "other..." (e.g., *K128: other cereals* and *K330: other protein crops*). The second reason that justifies the inclusion of error terms concerns the feasibility of the maximum entropy problem: as explained by **?**, the role of the disturbances is to "loosen" the model constraints so that the model returns an optimal solution. This implies that there should be a large enough error support interval in order to compensate for deviations between observed aggregate input costs and estimated cost allocations.

The goal is now to find support intervals for α_{ni} and υ_{ni}. The support interval for α_{ni} can be constructed using the "3σ rule", that is, three standard errors (`s.e.`) around the *LS* estimate: $\hat{b}_i - 3$ s.e. $(\hat{b}_i) \leq \alpha_{ni} \leq \hat{b}_i + 3$ s.e. $(\hat{b}_i)$. This method requires that (i) the $\hat{b}_i$ coefficient is positive and (ii) the resulting support interval does not contain any negative values. However, this may not always be possible with least square techniques because the analyst can neither impose non-negativity constraints on the estimation procedure nor ensure that the estimated coefficients are statistically significant, so that the interval strictly lies in $\mathbb{R}_+$.[6] A simple ad-hoc solution when $\hat{b}_i$ is not statistically significant would be to re-define the lower bound of the support interval as $\hat{b}_i -$ s.e. $(\hat{b}_i)$. However, if this still results in negative interval values, or if $\hat{b}_i$ is negative, then the lower bound should be set to zero.

[6] A higher variance will probably lead to a negative lower bound with the "3σ" rule.

On the other hand, the support interval for υ_{ni} can be defined with respect to the residuals of the initial regression, e_n. More precisely, e_n needs to be distributed among the various activities of the same farm. We will therefore assume that $-|e_n| \leq \upsilon_{ni} \leq |e_n|$. This interval is large enough to allow for an optimal solution to the entropy model and prevent any infeasibility problems.

For the formulation of the *GME* problem we will denote the unknown probability distribution corresponding to α_{ni} by $\mathbf{h}_{ni} = [h_{ni1}, h_{ni2}, \ldots, h_{niK}]$ and the distribution for every υ_{ni} term by $\mathbf{w}_{ni} = [w_{ni1}, w_{ni2}, \ldots, w_{niK}]$.

Similarly, the support vectors for α_{ni} will be denoted by $\mathbf{z}_{ni} = [z_{ni1}, z_{ni2}, \ldots, z_{niK}]$ and for υ_{ni} by $\mathbf{v}_{ni} = [v_{ni1}, v_{ni2}, \ldots, v_{niK}]$. Since entropy is a concave function of the non-negative probability vectors $\mathbf{h}_i$ and $\mathbf{w}_n$, estimating the allocation of any type of input cost across every activity i can be expressed in the form of the following optimization program which has a unique interior solution:

$$\begin{aligned}
\max_{h_{nik},\ w_{nik} \geq 0} \quad & H = -\sum_{n,i,k} h_{nik} \ln(h_{nik}) - \sum_{n,i,k} w_{nik} \ln(w_{nik}) \\
\text{s.t.} \quad & y_n = \sum_i \left[x_{ni} \left(\sum_k z_{nik} h_{nik} + \sum_k v_{nik} w_{nik} \right) \right] \\
& \sum_k h_{nik} = 1 \ , \quad \sum_k w_{nik} = 1 \\
& \forall n \in N, \quad \forall i \in I, \quad \forall k \in K
\end{aligned}$$

Assuming that h^*_{nik} and w^*_{nik} are the solutions to the maximum entropy problem, the unknown parameters are calculated as:

$$\alpha_{ni} = \sum_k z_{nik} h^*_{nik} \qquad \text{and} \qquad \upsilon_{ni} = \sum_k v_{nik} w^*_{nik}$$

Note that we propose an initial regression exercise that uses all observations (farms) in each *EU* country surveyed by *FADN*. An objection to this choice would be based on the expected heterogeneity in activity costs among the various regions due to different technologies which are adapted to the local physical conditions. Although this heterogeneity can possibly be captured with the use of dummy variables, the number of additional variables introduced in the model (No. of activities × No. of regions) would make this procedure too cumbersome. Another solution would be to perform regressions for every region, but at the cost of less observations for each activity-cost combination and the increased risk of not statistically significant coefficients.

6.1.2 Implementation details for cost estimating

6.1.2.1 Agricultural activities and cost data The agricultural and animal production activities in *AROPAj* constitute only a subset of the activities that appear in *FADN*. A detailed presentation of all activities considered in the cost estimation exercise, along with the corresponding *FADN* variable codes is given in Table **??**. Note that *FADN* includes several sub-categories for the various animal types, for example, cattle by age or by production orientation. Although *AROPAj* also uses various livestock categories in order to better represent feeding requirements in its constraint set, for the purpose of cost estimation we have grouped all animal activities.

A special note should be made for irrigated agricultural activities. More precisely, *FADN* includes data on irrigated agricultural land only for the main crop types (*FADN* variables `K1XX`),

whereas all `K3XX` variables are actually sub-categories of the `K1XX` ones. For example, cotton and fibre flax belong to the general category "*Other industrial crops*—`K135`". Although *FADN* reports total land for `K135` and the part of it which is irrigated (`K135AA` and `K135IRAA` respectively according to *FADN* nomenclature, with K135AA $\geq$ K135IRAA), it only reports total land for cotton and fibre flax. In order to estimate the irrigated areas for these two crops, we assumed that the irrigated shares of a general category is the same as the crops that comprise it. Hence, if K135IRAA/K135AA $= \lambda$ (with $\lambda \leq 1$), then λ is also the irrigated share for both cotton and fibre flax. Finally, for completeness we have created a "virtual" activity (`a26` and `r26`) that includes permanent crops, vines, orchards and nurseries which are not represented in *AROPAj* but nevertheless affect farm expenses.

Table 23. Activities considered and related *FADN* variable codes.

		Notation	
Activities	**FADN variable**	**Generic**	**Irrigated**
Agricultural Activities			
Common wheat and spelt	K120	$a1$	$r1$
Durum wheat	K121	$a2$	$r2$
Rye	K122	$a3$	
Barley	K123	$a4$	$r4$
Oats	K124	$a5$	
Grain maize	K126	$a6$	$r6$
Rice	K127	$a7$	$r7$
Other cereals	K125+K128	$a8$	$r8$
Dry pulses	K129	$a9$	
Potatoes	K130	$a10$	$r10$
Sugar beet	K131	$a11$	$r11$
Oilseed rape	K331*	$a12$	$r12$
Sunflower	K332*	$a13$	$r13$
Soya	K333*	$a14$	$r14$
Tobacco	K134	$a15$	$r15$
Cotton	K347*	$a16$	$r16$
Fibre flax	K373*	$a17$	$r17$
Fodder roots and brassicas	K144	$a18$	
Fodder maize	K326*	$a19$	$r19$
Other fodder plants	K327*	$a20$	$r20$
Other silage cereals	K328*	$a21$	$r21$
Fallow and set aside	K146	$a22$	
Temporary grass	K147	$a23$	
Permanent pasture	K150	$a24$	
Open ground field scale vegetables	K136	$a25$	$r25$
Sum of activities not in AROPAj	**	$a26$	$r26$
Animal Activities			
Cows	D23LU+...+D32LU	$a27$	
Sheep and goats	D38LU+...+D41LU	$a28$	
Pigs	D43LU+...+D46LU	$a29$	
Chicken and other poultry	D47LU+...+D49LU	$a30$	

*See text for more details on the calculation of irrigated areas.
**See text for more details on the activities not in *AROPAj*.

On the other hand, production costs reported in *FADN* include the value of all non-capital inputs used in crop and animal production, that is, the "*consumption*" of productive resources that correspond to farm output during the accounting year. The following are not part of production costs: : (1) Costs that do not correspond to the accounting year production; (2) changes in stocks of inputs, including costs accruing to growing crops; (3) cost of inputs used to increase, repair or replace the holding's fixed assets which are considered as "*Investment*"; (4) consumption of capital assets ("*Depreciation*") and/or acquisition of capital; (5) costs of production for private

consumption, e.g. inputs for kitchen gardens; (6) expenditure on cost items compensated during the accounting year, e.g. tractor repairs covered by an insurance policy the same year; (7) charges on land, quotas and other rights, and buildings, but not short-term renting and other charges; (8) interests paid.

According to *FADN* nomenclature, production costs can be divided into four general categories, namely animal-specific costs, crop-specific costs, labour and machinery costs, and farming overheads. As illustrated in Table **??**, we have aggregated specific *FADN* variables in order to produce 15 types of costs that need to be estimated per farm and activity.

Table 24. Production costs considered and related *FADN* variable codes.

Cost Notation	*FADN* variable	Description
Animal-specific		
$c1$	F64	Concentrated feedingstuffs for grazing livestock (farm)
	F65	Coarse fodder for grazing livestock (farm)
	F68	Feedingstuffs for grazing livestock (market)
$c2$	F66	Feedingstuffs for pigs (farm)
	F69	Feedingstuffs for pigs (market)
$c3$	F67	Feedingstuffs for poultry & other small animals (farm)
	F70	Feedingstuffs for poultry & other small animals (market)
$c4$	F71	Other specific livestock costs (market)
Labour and machinery		
$c5$	F59	Wages paid and social security costs
$c6$	F60	Contract work and machinery hire
$c7$	F61	Current upkeep of machinery and equipment
	F63	Car expenses
$c8$	F62	Motor fuels and lubricants
Specific crop costs		
$c9$	F72	Seeds and seedlings purchased
	F73	Seeds and seedlings produced and used on the farm
$c10$	F74	Fertilisers and soil improvers
$c11$	F75	Crop protection products
$c12$	F76	Other specific crop costs
	F77	Specific forestry costs
Farming overheads		
$c13$	F79	Electricity
$c14$	F81	Water
$c15$	F78	Upkeep of land improvements and buildings
	F80	Heating fuels
	F82	Insurance
	F83	Taxes and other dues
	F84	Other farming overheads

6.1.2.2 Regression analysis As explained previously, the first step of the cost disaggregation procedure involves the estimation of an average cost per production activity ($\hat{b}_i$) by means of regression analysis. For this purpose, a different model was built for each cost category (15 in total), depending on the activities which were most likely to involve a certain type of cost. For example, labour and machinery costs should correspond to all activities, whereas, animal specific costs were regressed only against the corresponding animal activities and crop-specific costs against agricultural activities. For $c8$ (motor fuels and lubricants), all agricultural areas were separated

into irrigated and non-irrigated ones. The latter (which will henceforth denoted by 'd') were calculated for each farm by subtracting 'r' areas from their respective 'a' values: $dX = aX - rX$. The reason for this distinction is that irrigated activities may exhibit higher costs than non irrigated ones. Therefore in these regression models we used both r and d entries as independent variables, instead of an aggregate 'a' activity. The same rationale was employed for the $c13$ (electricity) and $c14$ (water) costs, since they were regressed against irrigated activities and animals only. For the remaining models we used the total agricultural areas as independent variables (a entries), since we assume that these costs per hectare should not differ between irrigated and non irrigated activities. Finally, agricultural activities that do not involve marketable outputs were not used in certain models. For example, activity $a22$ (fallow and set aside) was de-associated from cost types $c9$ (seeds and seedlings), $c10$ (fertilizers and soil improvers), and $c11$ (crop protection products). Table **??** presents the explanatory variables (activities) used in the estimation of the different cost types.

Table 25. Production costs considered and related *FADN* variable codes.

Cost types	Agricultural activities	Animal activities
$c1$		$a27$, $a28$
$c2$		$a29$
$c3$		$a30$
$c4$		$a27$, $a28$, $a29$, $a30$
$c5$	$a1$, ..., $a26$	$a27$, $a28$, $a29$, $a30$
$c6$	$a1$, ..., $a26$	$a27$, $a28$, $a29$, $a30$
$c7$	$a1$, ..., $a26$	$a27$, $a28$, $a29$, $a30$
$c8$	$d1$, ..., $d26$, $r1$, ..., $r26$	$a27$, $a28$, $a29$, $a30$
$c9$	$a1$, ..., $a21$, $a23$, ,$a25$, $a26$	
$c10$	$a1$, ..., $a21$, $a23$, ..., $a26$	
$c11$	$a1$, ..., $a21$, $a25$, $a26$	
$c12$	$a1$, ..., $a21$, $a25$, $a26$	
$c13$	$r1$, ..., $r26$	$a27$, $a28$, $a29$, $a30$
$c14$	$r1$, ..., $r26$	$a27$, $a28$, $a29$, $a30$
$c15$	$a1$, ..., $a26$	$a27$, $a28$, $a29$, $a30$

Since the goal of the initial regressions is to aid in the construction of the support intervals for the *GME* model and not to accurately estimate the average cost coefficients, for simplicity's sake at this stage, we chose to perform equation by equation *OLS* estimation with no constant term. The reason is that a constant term would also have to be distributed to different activities, like the residuals, in each farm in order for the budget constraint to hold. Future work can use *WLS* estimation in order to also account for the representation weights attributed to each farm in the *FADN* sample.

6.1.2.3 The maximum entropy program The outputs from each model (cost type) were the $\hat{b}_i$ coefficients, the standard errors and the residuals for each farm–observation. These results were used as data inputs for the second step of the cost disaggregation procedure that involves the use of a *GME* model for each farm included in the *FADN* sample. One fundamental differentiation from the initial regressions, was that for the entropy model, we explicitly distinguished between irrigated and non irrigated activities for all cost types. Since the explanatory variables for most regression models were total crop areas ('a'), some data manipulation was required in order to render the regression results consistent with the *GME* model. Hence we assumed that the $\hat{b}_i$ coefficients and

the standard errors for all cost types were the same for both irrigated and non irrigated crops in the same farm, which finally lead to the same support intervals for the two crop variants. Additionally, as explained before, negative coefficients were set to zero, which meant that the support interval for β_{ni} also had a lower bound of zero.

For the maximum entropy program, we introduced additional constraints for $c8$ (motor fuels and lubricants), $c13$ (electricity) and $c14$ (water) in order to reflect possible differences between the two crop variants if they are both recorded in the same farm. More precisely, $c13$ and $c14$ were not considered at all for non irrigated crops (as is also shown in Table **??**), whereas $c8$ for irrigated crops were set to be equal or greater than non irrigated ones. Especially for $c8$ (motor fuels and lubricants), we assumed that any additional costs were the result of using a petrol-based irrigation system. Thus the existence of electricity expenses in the same farm meant that $c8$ costs were the equal between irrigated and non irrigated crops. Finally, farms with only one recorded activity were excluded from the *GME* model because the repartition of costs is straightforward. The model was solved within a "loop" command in *GAMS* which, since entropy was maximized independently for each farm, lead to a number of sub-models equal to the number of remaining farms. A matrix of farm-activity correspondences was created for each remaining farm based on *FADN* data. Hence each model maximized entropy for farms with more than one activity and for those costs that are actually recorded. This constraint was deemed necessary for feasibility reasons when solving the model with the *GAMS* programming language.

Since the main decision units in *AROPAj* are farm-types and not individual farms, the results of the entropy model, i.e., the input costs per hectare for each farm, were transformed into average values for every farm-type. Hence, we have implicitly assumed that all costs are the same for the irrigated and non irrigated variant of the same activity and are numerically equal to the average of costs in the examined farm-type. An exception was made for fuel costs ($c8$) which were distinguished between irrigated ($c8_{ir}$) and non irrigated activities ($c8_{nir}$) because irrigation with no electricity costs implies the use of oil-based water pumps. Therefore, when non zero electricity ($c13$) expenses were recorded for the irrigated activity then $c8$ was averaged out for both the irrigated and non irrigated activity variants so that $c8_{ir} = c8_{nir}$. In contrast, for $c13 = 0$, we required that $c8_{ir} > c8_{nir}$. Since the two activity variants at the farm-type level come from different individual farms, a numerical problem that arises is that the previous inequality may not always hold. In such case, and when $c13 = 0$, the estimated $c8_{ir}$ was assumed to be equal to $c8_{nir}$, plus their difference:

$$\hat{c8}_{ir} = c8_{nir} + (c8_{nir} - c8_{ir}) = 2c8_{nir} - c8_{ir}$$

6.2 Parameters estimating for farm groups

Key-point:

Estimating farm groups' parameters

The majority of *AROPAj* parameters are provided by estimations from *FADN* data. The previous section describes the calculation process required to deal with the insufficiently detailed information per crop. After that, it is easy to estimate costs at the farm group level thanks to the provision of (estimated) individual costs for the entire *FADN* sample. We did that separately for

fertilizer costs, other variables costs for all *AROPAj* versions, and for irrigation costs for the *V5* version.

Other parameters are provided with two methods. Some are calculated by the means of simple linear regression or estimated average. This is done through *R*-code programs which can be automatically managed per series of parameter categories for all countries. Parameter categories of this part refer to:

- crop yields
- areas (requested by the *AROPAj* calibrating process)
- herds
- on-farm re-used crop products (useful for improving the *AROPAj* calibrating process)
- variable costs (in case of non-operating *GME* estimator)
- farm numbers per crop or per crop category
- ratios between crops or fodders, and categories of crops or fodders
- milk productions and quotas, sugar beet quotas

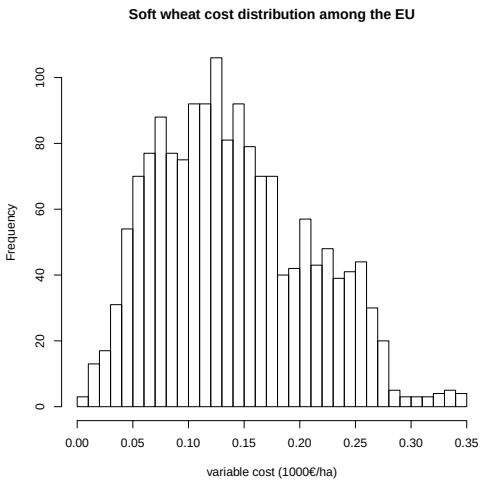


Figure 8. Fertilizer cost (1000€/ha)

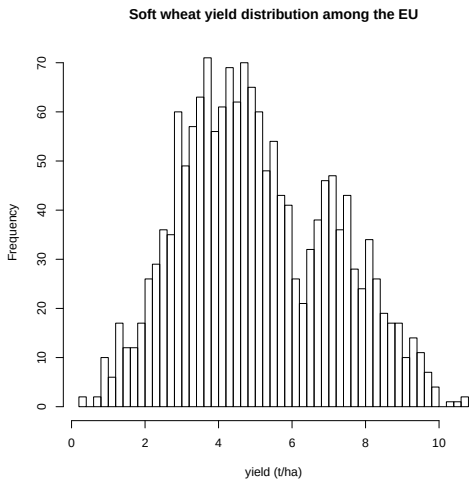


Figure 9. Yield (t/ha)

Distribution of fertilizer costs (left) and yields (right), for soft wheat, among the EU AROPAj farm groups.

From expert sources, we gather other important information such as:

- *IPCC* guidelines sourced recommendation
- energy and protein compounds of concentrates and on-farm sourced feed-stuff
- animal requirements in energy, proteins and dry matter, and limits or thresholds regarding animal feeding
- by default characteristics of fertilizers
- by default characteristics of manure
- by default value of ratios within the bovine herd
- *CAP* tool values linked to (more or less) stylized *CAP* tool implementation

Finally, in some cases, parameter values can be provided by the *AROPAj* model itself, when outputs become parameters. This is the case for the following parameters:

- baseline *GHG* emissions (to be quotas)
- baseline cereals, seed and protein crops (to be *CAP* supported areas)
- shadow prices (to replace market prices in the case of endogenous computation of yields)

Figures **??** and **??** illustrate the parameter heterogeneity in case of fertilizer costs and yields related to soft wheat, present in 1642 among the 1815 farm groups related to the *V5 version* and the *FADN*-2009. The figures show a bi-modal distribution.

7 Calibration of the *AROPAj* model

Key-points:

Finding a *LP* solution at the farm group level
Improving all time-reference *LP*s through parameter re-estimates
Applying the algorithm, and running time

The first problem we face after prior estimation of parameters is to obtain a *LP* solution. To do so, we check parameters' consistency with regards to a series of potential problems due to missing information. For instance, at the farm group level, the *FADN* may provide milk quota and milk production but no milk price ; it may provide crop area but no crop production neither crop price. This poses a problem even when a *LP* solution exists, and even more when there is no *LP* solution. A series of "expert" checking points are processed in the earlier steps of the *AROPAj* building. They are based on fulfilling farm group data sets with nested calculations: given a parameter (e.g. a yield, a price), in case of missing farm group information, we use regional average values; in case of missing region information, we use country-level average values; in case of missing country information, we retain the default value. In practice, default values were used in very few cases, related to bovine activities (uncomplete livestock categories) in small countries (Cyprus).

The second problem arises when we compare computed optimal activities and actual activities, once we have a *LP* solution derived from prior estimates of parameters. The *FADN* provides a wide range of parameters, some of them being related to *LP* command variables, and, at the same time, the *FADN* has to be supplemented by exogenous technical information (e.g. energy and protein contents of feed intakes, feed requirements, fertilizer characteristics, livestock life cycle, agronomic cropping limits in terms of land allocation).

The calibration process consists of re-estimating a subset composed of previously estimated parameters (crop ratios) and of non-*FADN* sourced parameters (feed requirements, livestock life cycle). In practice, more than one hundred parameters per farm group are re-estimated in this way. The calibration criterion is a sum of weighted squared deviations, each of them being the difference between *FADN* observation and the associated optimal *LP* variable. The set of variables is composed of crop and fodder areas (hectares), livestock categories (head units) and feed intakes (metric tons). Table **??** lists the variables which contribute to the calibration criterion and fully describes the criterion calculation.

Figure **??** illustrates the geometric interpretation of the calibration phase, which can be viewed as a rotational and translational movement of constraints in the primal variable space.

The calibration algorithm is based on sequential calculations, combining Monte-Carlo and gradient methods. In practice, randomization of calibration parameters alternating with "local" gradient based improvement of the criterion is generated from a large number of *LP* runs for each of *AROPAj* farm groups. All operating *AROPAj* versions are based on more than one thousand farm groups. For one farm group, the number of runs ranges from 1000 to 1500 (*AROPAj* versions *V2* and *V3*) or to 2000 (versions *V4* and *V5*). The calibration process is stopped either when the run number limit is reached (i.e. pre-defined number for Monte-Carlo and gradient phases), or when additional change of parameters' value does not improve the criterion during a phase.

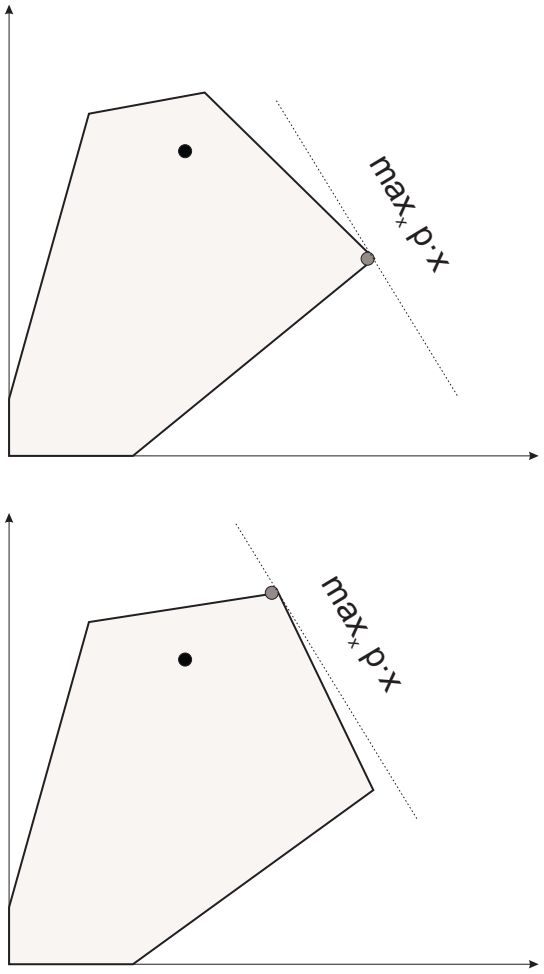


Figure 10. Geometric overview of the calibrating process on a two-dimension subspace: the target fixed point is marked by the black circle, the LP point maximizing the gross margin $p \cdot x$ is marked by the grey circle, the feasible solution has to belong to the colored area bounded by the LP constraints. Calibrating consists on rotational and translational movements aiming to bring the two points closer.

Table 26. Detailed expression of the calibrating criterion to be maximized c (δ refers to any additional deviation, w is 1 by default); each deviation enters the criterion such as $c \rightarrow c - 1/w\ \delta^2$.

variable	additional weight $1/w$	(δ unit)
each of 8 cereals		(ha)
cereals together		(ha)
each of 3 oilseeds		(ha)
each of 2 proteins crops		(ha)
sugarbeet		(ha)
potatoes		(ha)
each of other crops		(ha)
fallow		(ha)
CAP set aside		(ha)
each of fodders		(ha)
grasslands		(ha)
fallow & set-aside		(ha)
bovine less than 1 year	16	(Lh)
one-year bovine male	4	(Lh)
one-year bovine female	4	(Lh)
male more than 2 years		(Lh)
heifers and cows		(Lh)
other animals		(Lh)
cow life cycle -5	4	(1)
each of feed categories	25	(t)

Results provided by the versions *V2*, *V3* and *V4* based on linear cost estimates are illustrated on Figures **??**, **??**, and **??**.

The calculations were carried out on a high-performance workstation for the *FADN* years 2007 up to 2012, initially. Results are illustrated on the figures **??**, **??**, **??**, **??**, **??**, **??**. The operations were renewed for the *FADN* years 2013-2018, after the *FADN* data were modified by DG-AGRI of the EC. Corresponding illustrations are provided in figures **??**, **??**, **??**, **??**, **??**.

Calibration may be considered as acceptable when the criterion $log(-c)$ is lower than 0, it is correct when $log(-c)$ is lower than -2.3 and very good when $log(-c)$ is lower than -4.6. We have to insist on the difficulty of comparing the calibrated farm group distribution involving different *AROPAj* versions: these versions usually refer to different stages of the European Union building, and recently added countries usually strongly differ in farm types. Typically, East Germany and other countries of Central and Eastern Europe had very large farming structures which did not exist in the founding *EU* countries. Finally, the calibration criterion was refined *after* the *V2* version.

The computation run time used to be a real difficulty. In the early 2000s, several weeks were required to calibrate from 1000 (*V2*) to 1400 (*V4)* farm groups.Now, when new iterations are added, the 1800 farm groups can be calibrated in less than a full day.

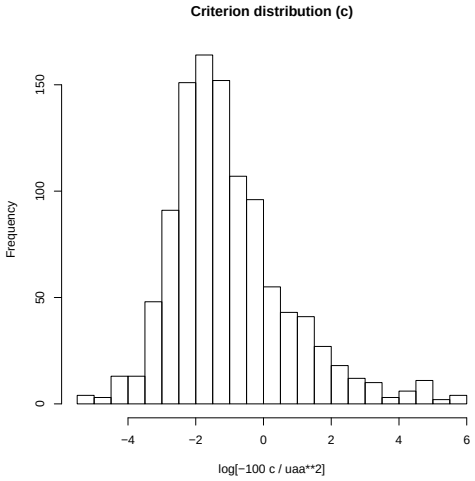


Figure 11. "V2" (*FADN-2001*)

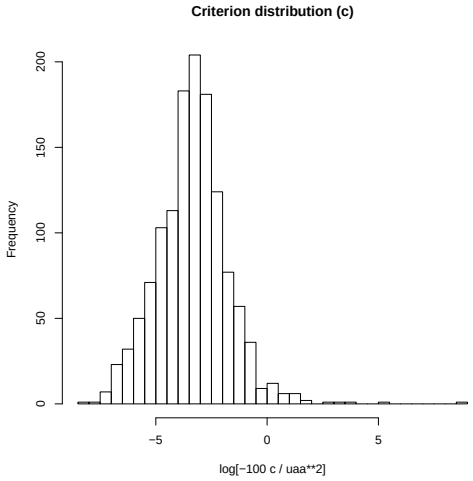


Figure 12. "V3" (*FADN-2004*)

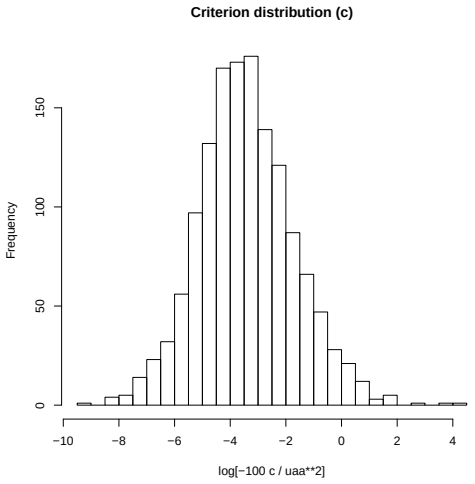


Figure 13. "V4" (*FADN-2009*)

Distribution set related to the calibrating criterion among farm groups for different FADN years - V2, V3, V4 versions.

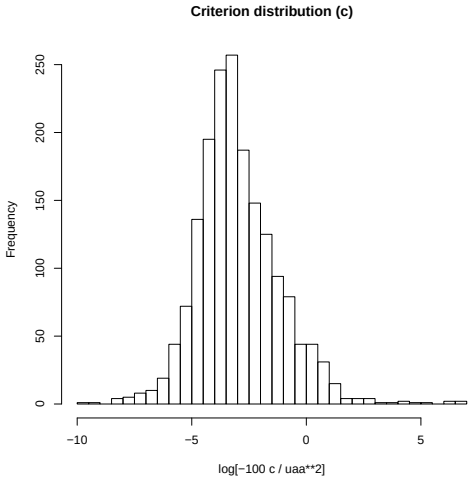


Figure 14. "V5_2007"

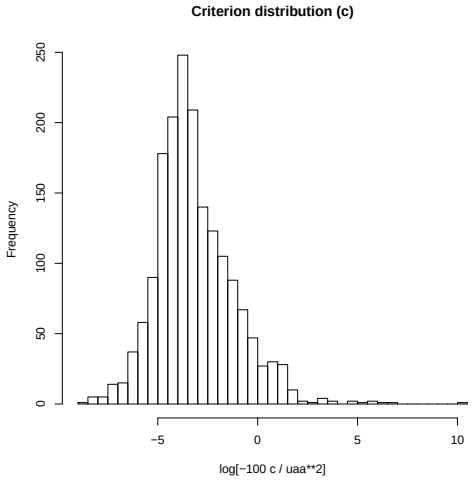


Figure 15. "V5_2008"

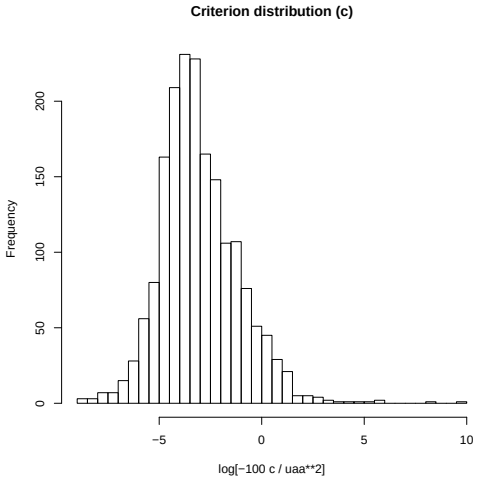


Figure 16. "V5_2009"

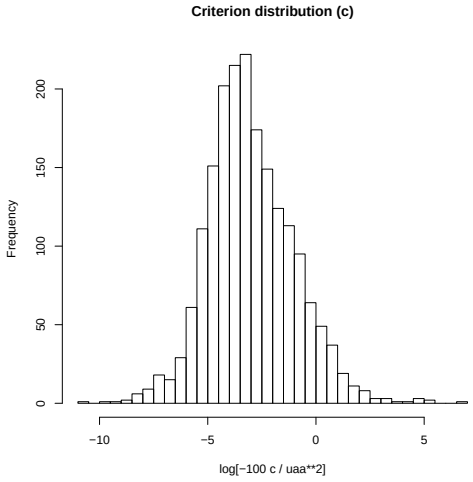


Figure 17. "V5_2010"

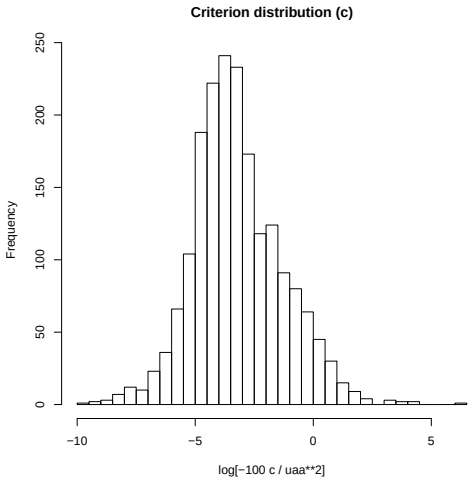


Figure 18. "V5_2011"

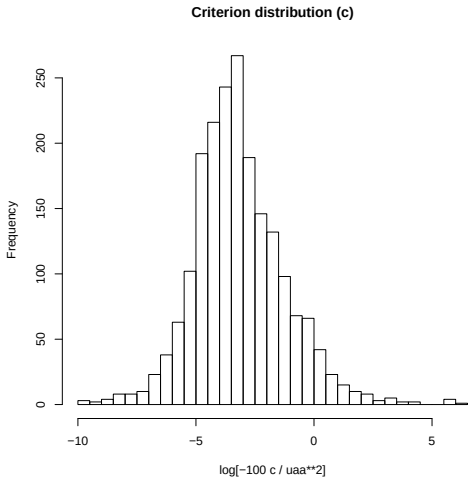


Figure 19. "V5_2012"

Distribution set related to the calibrating criterion among EU-27 farm groups for 6 different FADN years (2007 up to 2012) - AROPAj version related to GME cost estimates.

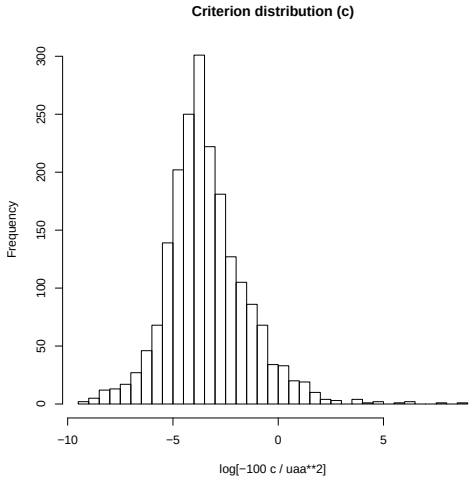


Figure 20. "V5_2013"

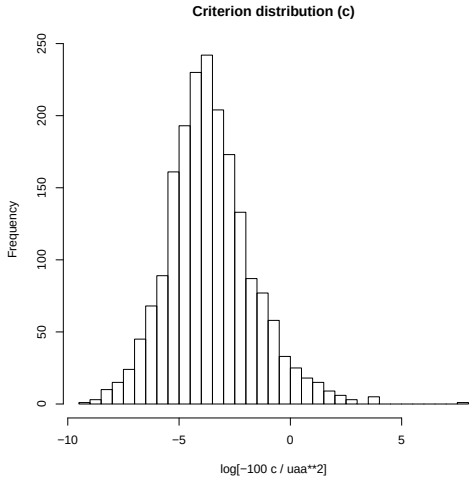


Figure 21. "V5_2014"

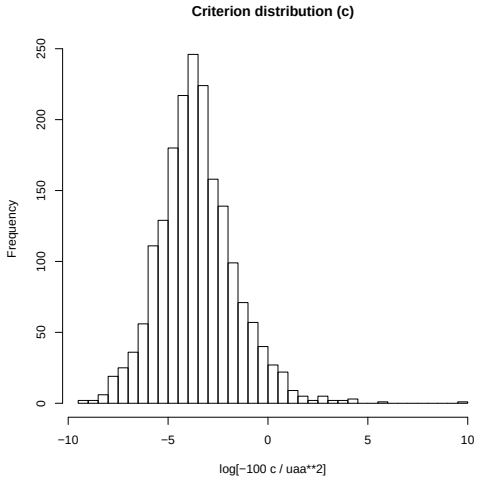


Figure 22. "V5_2015"

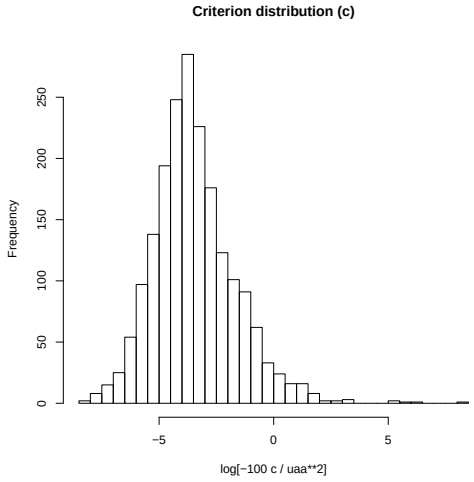


Figure 23. "V5_2016"

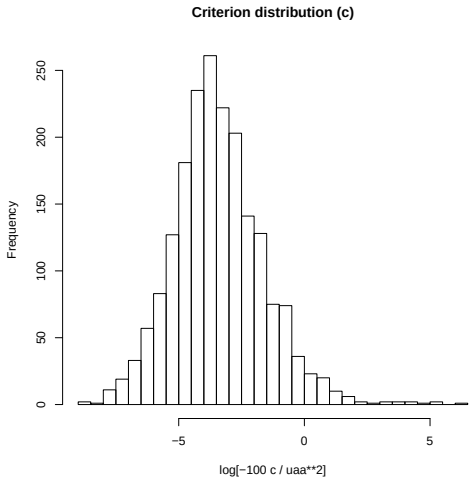


Figure 24. "V5_2017"

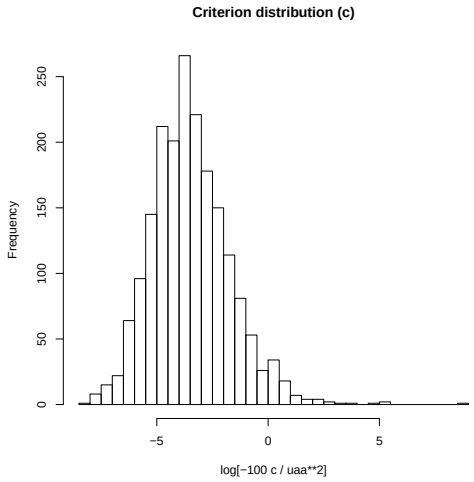


Figure 25. "V5_2018"

Distribution set related to the calibrating criterion among EU-28 farm groups for 6 different FADN years (2013 up to 2018) - AROPAj version related to GME cost estimates.

8 Spatialization of *AROPAj* outputs

Key-points:

Obtaining spatialized probabilities for crop location
Calling spatial econometrics and the Multinomial logit approach
Generalized cross entropy aiming at refining the probabilities

The proposed spatial disaggregation approach (**?**) shows how land use data and biophysical data, together available at a fine resolution level, can be combined to disaggregate regional data and provide more complete information. The proposed method is based on a two-step process. First, the estimation of a disaggregated econometric model (multinomial logit -*MNL*) of land use with Land Use/Cover Area frame Statistical Survey -*LUCAS* data in combination with data on soil, climate, altitude and land cover (CORINE Land Cover -*CLC*). Second, using the estimates from the first stage as priors and refining them through a cross entropy method, we disaggregate the *FADN* data at a fine level (100m pixel resolution).

8.1 Disaggregation approach in spatial econometrics devoted to crop location

The proposed disaggregation approach is a two-step procedure. In the first step, we estimate a land use model. In the second step, using estimations from the first step as "priors", we disaggregate the aggregated regional data available at the regional administrative level with the Generalized Cross-Entropy method (*GCE*).

8.1.1 The land use model

The *MNL* model estimates the probability outcome associated with each category of land use depending on a set of explanatory variables. The probability of observing the land use c at parcel l can be expressed as:

$$p_{lc} = \frac{exp(\beta_c^{'} x_{lc})}{\sum_{c=1}^{C} exp(\beta_c^{'} x_{lc})} \tag{2}$$

where C is the number of possible land use categories, $\beta_c^{'}$ is a vector of parameters to be estimated for land use c, and x_{lc} are explanatory variables associated with land use c and parcel l.

The log-likelihood function of a sample of L parcels is given by:

$$\ln(L(\beta)) = \sum_{l=1}^{L} \sum_{c=1}^{C} ln p_{lc}^{y_{lc}} \tag{3}$$

where y_{lc} is a dummy variable such that $y_{lc} = 1$ if land use c is observed at location l and $y_{lc} = 0$ otherwise.

8.1.2 The disaggregation model: the Generalized Cross Entropy approach

The objective here is to utilize land use estimation results to find a disaggregated spatial allocation of regional agricultural land use data. We used the Generalized Cross Entropy (*GCE*)

approach (**?**) to carry out the disaggregation procedure. The entropy concept in information theory was originally proposed by **?** as a measure of uncertainty or missing information. Suppose there are C possible outcomes for some event and the probability distribution $(p_1, p_2, ...p_C)$ can be used to explain outcomes. Shannon's information entropy is defined as:

$$H(p_1, p_2, ...p_C) = -\sum_{c=1}^{C} p_c ln(p_c). \tag{4}$$

To estimate the unknown probabilities p_c, **?** proposed maximizing entropy and subject to sample-moment information and then adding up constraints on the probabilities. This approach is based on the idea that the frequency that maximizes entropy is a reasonable estimate of the true distribution when we lack other information. If we have any additional information from previous experiments or observations, we can use those *priors* to alter this estimate.

? introduced the notion of Cross Entropy[7] (CE). Unlike the ME approach, where we maximize uncertainty implied by the probabilities, in the CE framework we minimize the CE, which is a measure of discrepancy between the posterior probabilities p and their priors q. This yields the minimization problem:

$$Min_p CE(p,q) = \sum_{c=1}^{C} p_c \ ln(p_c/q_c) \tag{5}$$

subject to all relevant adding up constraints on probabilities, where $p = p_1, p_2, ...p_C$ and $q = q_1, q_2, ...q_C$.

The GCE approach is applied here to estimate agricultural land use choices at a disaggregated level taking into account the observed aggregated regional level data. The estimation of the land use model by the Multinomial logit provided us with some prior knowledge about the land share of each crop j in each pixel i. This corresponds to the prior probabilities $\hat{\pi}_{ij}$ that will enable us to estimate probabilities p_{ij} to observe crop j in pixel i by solving the non-linear optimisation program:

$$\min_p \ CE(p_{ij}, \varepsilon) = \sum_{i=1}^{I} \sum_{j=1}^{C} p_{ij} \ ln(p_{ij}/\hat{\pi}_{ij}) + \sum_{j=1}^{C} \sum_{n=1}^{N} \varepsilon_{jn} ln(\varepsilon_{jn}) \tag{6}$$

subject to:

$$\sum_{i=1}^{I} p_{ij} \times s_i + \varepsilon_j = S_j^R, \forall j = 1, ..., C. \tag{7}$$

$$\sum_{j=1}^{C} p_{ij} = 1, \forall i = 1, ..., I \ and \ p_{ij} \in [0,1]. \tag{8}$$

$$\sum_{n=1}^{N} \varepsilon_{jn} = 1, \forall j = 1, ..., C \ and \ \varepsilon_{jn} \in [0,1]. \tag{9}$$

where s_i is the area of pixel i and S_j^R is the area allocated to crop j at the regional level. Following the GCE framework, the error terms ε_j associated with the constraints (**??**) are reparameterized in terms of probabilities such that $\varepsilon_j = \sum_{n=1}^{N} \zeta_n \varepsilon_{jn}$ where $\zeta_1, \ldots, \zeta_N$ with $N \geqslant 2$ is the support set associated with probabilities $\varepsilon_{j1}, \ldots, \varepsilon_{jN}$.

Equation (**??**) is the data compatibility constraint, which ensures that the predicted land allo-

[7] The Cross entropy, or Relative Entropy, is also known as the Kullback-Leibler distance or divergence. This distance is not, in general, symmetric

cated to a specific crop j across all pixels is equal to the observed area allocated at the regional level to crop j. Equation (**??**) is the adding-up constraint that ensures that the area allocated to all crops in a given pixel i is exactly equal to the area of pixel i. Equation (**??**) ensures that $\varepsilon_{j1},\ldots,\varepsilon_{jN}$ is a probability distribution. The objective of the *GCE* problem (**??**)-(**??**) is to recover the unknown probability distributions p_{ij} and ε_{jn}. This is a standard *GCE* problem ; due to its convexity, there will be unique solution.

8.2 Method and process implementation

Key-points:

Obtaining spatialized contribution of farm groups to land based activities

Algorithm implementation

8.2.1 Crop location estimates at a fine resolution level

Data utilized for the spatialization block refers to raw physical data, as defined in section **??**.

For the spatialization of *V5* of *AROPAj*, we use *R*, *GDAL*, *POSTGIS* and *GAMS* scripts and procedures.

Data treatment

The spatialization procedure combines a large variety of information. The starting point is the *LUCAS* database which contains points of observations on crops (land uses) and land cover. For each *LUCAS* point data on soils, climate, relief and land cover (*CLC*) is extracted (see Figure **??**). In the same sense, *CLC* raster data is intersected with the other layers. In order to do so, rasters devoted to agricultural activities treated in the *AROPAj* model are extracted and polygonized through *GDAL* programs. Afterwards, the obtained polygons are intersected with the polygons containing the other raw information.

Aggregation of crops into crop classes

Crops observed in the *FADN* data are grouped in different classes depending on their representativeness in the *FADN* region. When a given crop surpasses a certain threshold (for instance, 5% of agricultural area in the *FADN* sample), it is modeled independently. Otherwise, it is grouped with other crops. The method is summarized in Table **??**.

The same classification of the crops is applied to the *LUCAS* points before the estimation of the *MNL* model.

MNL and estimation of priors

In order to establish the *MNL* model formula a dedicated *R* script is used (package *VGAM*). The algorithm defines a list of variables (that actually vary in the region) and their corresponding interaction terms. Three loops over this list are then executed. In each loop, each variable (or interaction terms) is added to the *MNL* model formula. The score of the model is defined as the percent of correctly predicted observations. If the score of the model with the additional variable

Table 27. Method of grouping crops depending on area ratio limits.

Crop	Crop >5% (Cat.1)	Crop <5% (Cat.2)	Cat.2 < 5% (Cat.3)
Cereals			
Wheat (common)	Wheat (common)	Cereals	Cereals
Wheat (durum)	Wheat (durum)	Cereals 2	Cereals
Barley	Barley	Cereals 2	Cereals
Oats	Cereals 2	Cereals 2	Cereals
Rye	Cereals 1	Cereals 1	Cereals
Other cereals	Cereals 1	Cereals 1	Cereals
Maize	Maize	Cereals 2	Cereals
Rice	Rice	Cereals	Cereals
Industrial crops			
Dry pulses	LEG 1	IND	IND
Potatoes	PDT	IND	IND
Sugar beet	BTS	IND	IND
Tobacco	TAB	IND	IND
Cotton	COT	IND	IND
Flax fiber	LIN	IND	IND
Other industrial crops	AIN	IND	IND
Oilseed rape	CLZ	PR1	PRO
Sunflower	TOR	PR1	PRO
Soya	SOJ	PRO	PRO
Fodder crops	FOU	PPR	PPR
Fallows and set aside	FRIC	PPR	PPR
Temporary grass	FRIC	PPR	PPR
Other arable crops	AUC	AUC	AUC
Permanent pasture	PPR	PPR	PPR

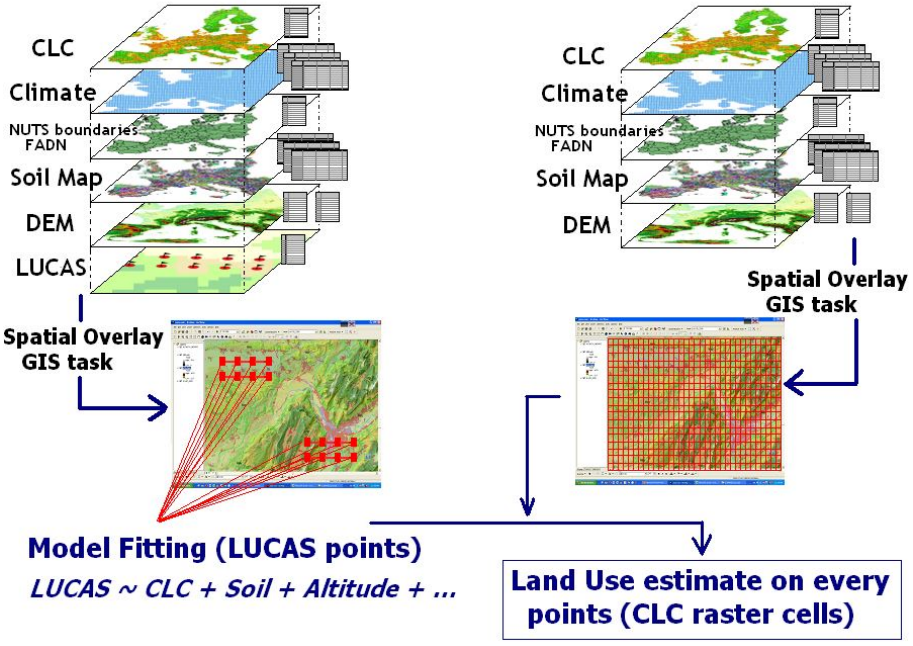


Figure 26. Intersecting and overlapping of information layers requested by *AROPAj* spatialization procedures. *Source:* **?**.

is at least as good as the best one already obtained, then the variable is retained in the formula. When all three loops are completed the formula with highest score and lowest number of variables is selected. Based on this formula, predictions of the probability of presence of the different land uses are estimated and exported as priors for the following *GCE* procedure.

GCE

In this step the predicted crop areas in the *MNL* phase are compared with the crop areas observed in the *FADN* database (as ratios) and the priors are refined (see Section **??**)).

8.2.2 Estimation of probabilities for farm types

In *V5* of *AROPAj* the irrigation plays a central part in the model. In order to account better for it in the spatialization we use additional information from the *LUCAS* database, namelly the presence of irrigation for agricultural needs at a given point. We first definy a 10 km x 10 km regular grid. For each cell of the grid we verify the presence of an irrigating point in the *LUCAS* data. If there are such points we assume that the cell has a greater potential to accomodate agriculutural activities subject to irrigation. The polygons of the grid are then intersected with those resulting from the polygonization of the *CLC* database and its intersection with the other data layers.

We describe below how to distribute farm group contribution to agricultural activities accross all pixels over a geographical area.

Prior information

We consider that the *FADN* region is the basic aggregation unit, and all requested information are region-based.

At the region scale, we have to distinguish:

- the geographical area
- the agricultural area estimated through general spatial statistics and related to *AROPAj* activities
- the *AROPAj* area (based on *FADN* estimates)

We denote by:

- i : the pixel
- j : the crop
- k : the farm group
- a : the elevation class (let us recall that the value of 1 refers to $[0, 300m[$, the value of 2 to $[300m, 600m[$, the value of 3 to altitudes greater than or equal to $600m$
- *irr* vs *nir* respectively for irrigated vs non irrigated areas (when available)

We assume to be informed on:

- the location probability to find the agricultural activity j on the pixel i, denoted by P_{ij}, $\forall i, \forall j \ : P_{ij} > 0$, $\forall i \ : \sum_j P_{ij} = 1$;
  P_{ij} is provided by *MNL* and cross entropy estimating process
- the *AROPAj* area devoted to the j crop by the k farm group, denoted by S_{jk} (prior estimate provided by *FADN*), and the k's utilised agricultural area $L_k = \sum_j S_{jk}$
- possibly informed share of the S_{jk} area between irrigated (S_{jk}^{irr}) and non irrigated (S_{jk}^{nir}) areas ($S_{jk}^{irr} + S_{jk}^{nir} = S_{jk}$)
- the share of k' utilized agricultural area (UAA) of a elevation class U_{ak} (in accordance with $\sum_a U_{ak} = 1$);
  We consider that any i pixel is elevation class homogenous ("a"⇒"i")

At the regional level, we assume to know:

- the total amount of irrigated areas relating to *AROPAj* activities (denoted by $\mathbb{I}$)
- the total amount of pixels of elevetion a covered by *AROPAj* activities (denoted by $\mathbb{A}(a)$)
- the total amount of areas covered by *AROPAj* activities (denoted by M)

The proposed estimate of the contribution of the k farm group to agricultural activities (among *AROPAj* ones) follows:

$$T_{ik} = \sum_j P_{ij} \left(\alpha \frac{S_{jk}}{\sum_{k'} S_{jk'}} + \beta \frac{S_{jk}^{irr}}{\sum_{k'} S_{jk'}^{irr}} \mathbb{1}_{\mathbb{I}}(i) \right) + \sum_a \gamma_a U_{ak} \mathbb{1}_{\mathbb{A}(a)}(i) \tag{10}$$

using the indicator functions $\mathbb{1}_{\mathbb{A}(a)}(i) = 1$ if pixel i elevation is a, else the value is 0, and $\mathbb{1}_{\mathbb{I}}(i) = 1$ if pixel i is irrigated, else 0.

Some *a priori* constraints have to be checked:

- T_{ik} has to be a probability set such that $\forall i \ : \sum_k T_{ik} = 1$, and consequently $\alpha + \beta + \sum_a \gamma_a = 1$
- When the region is homogenous in elevation, the γ_a parameters have to be equal to 0
- When there is no information upon irrigation, the β parameter has to be equal to 0

At this stage, there is no reason to consider any uplet $\{\alpha, \beta, \gamma\}, \alpha + \beta + \gamma = 1$, as a mean to provide the best estimate of probabilities, in the sense the linear form of term combination is questionable (linearity means that the parameter uplet does not depend on the i pixel). But we have a real probability set, for which we just need to set four of the five parameters α, β, γ_a, $a \in \{1,2,3\}$. Additionnal information may be useful to do that.

Let us denote the set of irrigated pixels by $\mathbb{I}$ (amount $\|\mathbb{I}\|$ in hectares), and the set of a elevation by $\mathbb{A}(a)$ (amount $\|\mathbb{A}(a)\|$ in hectares). First, irrigation is one of the characteristics of the i pixel (through the *LUCAS* database). Secondly, the i elevation is assumed to be known (through the *DEM* database). Three equations describe these informations and relate them to *AROPAj*-sourced information:

$$\sum_i 1 = \quad M \quad = \frac{1}{\rho} \sum_{jk} S_{jk} \tag{11}$$

$$\sum_i \mathbb{1}_{\mathbb{I}}(i) = \quad \|\mathbb{I}\| \quad = \frac{1+\delta_{irr}}{\rho} \sum_{jk} S_{jk}^{irr} \tag{12}$$

$$\forall a : \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) = \quad \|\mathbb{A}(a)\| \quad = \frac{1+\delta_a}{\rho} \sum_k U_{ak} \sum_j S_{jk} \tag{13}$$

where ρ reflects the *AROPAj* representativity in area of the *AROPAj*-activities in the region and the δ refer to distortions due to the fact that *AROPAj* may differ in farming system representativity accounting for irrigation or for elevation. The ρ parameter is determined by the known amount of *CLC* pixels dedicated to *AROPAj*-like activities.

Combining the definition of the T_{ik} and the previous relations and respectively accounting for irrigation calculation (whatever is the elevation) and elevation calculation (whatever is the irrigation) in terms of area balance, we obtain:

$$\begin{aligned} \sum_i \mathbb{1}_{\mathbb{I}}(i) &= (\alpha+\beta) \sum_i \mathbb{1}_{\mathbb{I}}(i) + \sum_a \gamma_a \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) \mathbb{1}_{\mathbb{I}}(i) \\ \forall a : \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) &= (\alpha+\gamma_a) \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) + \beta \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) \mathbb{1}_{\mathbb{I}}(i) \end{aligned}$$

leading us to the equations:

$$(1-\alpha-\beta)\,\|\mathbb{I}\| = \sum_a \gamma_a \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) \mathbb{1}_{\mathbb{I}}(i) \tag{14}$$

$$\forall a : (1-\alpha-\gamma_a)\,\|\mathbb{A}(a)\| = \beta \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) \mathbb{1}_{\mathbb{I}}(i) \tag{15}$$

We can be comfortable with these formulas in particular cases such as when there is no irrigation in the region, or when only one elevation class exists in the region. We aim at providing an acceptable solution in $\{\alpha, \beta, \gamma\}$ in any case, taking account of amounts of areas intersecting irrigation and elevation class (we denote $\|\mathbb{I}\mathbb{A}(a)\| = \sum_i \mathbb{1}_{\mathbb{A}(a)}(i) \mathbb{1}_{\mathbb{I}}(i)$ possibly estimated through overlapping of *LUCAS* and *DEM* databases).

No irrigation When $\forall i\ \mathbb{1}_{\mathbb{I}}(i) = 0$ (implying $\mathbb{I} = 0$), necessarily $\beta = 0$ (thanks to **??** and $\alpha + \beta + \gamma = 1$). There remains to set γ_a values. We suggest to retain positive values increasing whith absolute values of distorsion δ_a such as $\gamma_a = 1 - \frac{M - \mathbb{A}(a)}{M} e^{-|\delta_a|}$. Let us note that the subcase in which the region is homogeneous in elevation leads to $\gamma_a = 0$, thanks to Equations **??** and **??**. This is compatible with the case below.

Homogeneous elevation class When the region is homogeneous in elevation class, there is nothing to loose to set $\gamma = 0$. In case of irrigation areas positively estimated by the *FADN-AROPAj* source, we suggest to retain a value reflecting the distorsion in irrigation between *FADN* and *DEM* sources for estimation, i.e. $\beta = 1 - \frac{M - \mathbb{I}}{M} e^{-|\delta_{irr}|}$. Let us note that when the part of irrigated area is close to 0 or to 1, the β value is close to 0. This is a consequence of Equation **??** which implies that δ_{irr} tends to 0 when $\mathbb{I}$ is close to M).

Mixed cases with areas splitted in irrigation and in elevation Let us consider the case where irrigation information exists in *FADN* (correlatively *AROPAj* S^{irr}_{jk} are estimated). Let us consider the system defined by Equation **??**, Equations **??** (as many as a elevation classes) and $\alpha + \beta + \sum_a \gamma_a = 1$ which is linear in the weights parameters α, β, and γ_a. Let us note that $\sum_a \|\mathbb{IA}(a)\| = \|\mathbb{I}\|$, due to the fact that the a areas do not intersect. With some re-arrangements of the system, we get that $\beta \|\mathbb{I}\| = \beta \sum_a (\|\mathbb{IA}(a)\|)^2 / \|\mathbb{A}(a)\|$. There is generally no solution except $\beta = 0$, which means that weights of our probability set (Equation **??**) is based on redundant information (irrigated areas on the one hand and irrigated areas intersecting elevation classses on the other hand). Moreover, that leads to keep equal the γ values, i.e. $\forall a : \gamma_a = \gamma$. We suggest to overcome this parameter under-determination by retaining the parameters as follows:

$$\gamma = 1 - \alpha = 1 - \frac{M - \mathbb{I}}{M} e^{-|\delta_{irr}|}$$

In other words, we rank first irrigation before elevation, which is nevertheless taken into account through the part of k farm group area of elevation a (U_{ak}). The probability set becomes:

$$T_{ik} = \alpha \sum_j P_{ij} \frac{S_{jk}}{\sum_{k'} S_{jk'}} + \gamma \sum_a \frac{U_{ak}}{\sum_{k'} U_{ak'}} \mathbb{1}_{\mathbb{A}(a)}(i) \tag{16}$$

The distorsion term δ_{irr} comes from Equations **??** and **??**: $\delta_{irr} = 1 - \frac{M}{\mathbb{I}} \frac{\sum^{irr}_{jk}}{\sum_{jk}}$.

8.3 Application

User-friendly application is anchored to the general AROPAj-menu ("miraj", see Appendix **??**), aiming first at managing the compilation of databases, precisely regarding:

1. FADN region list
2. CLC for Europe raster
3. Soils
4. Elevation
5. Slope
6. year-dependent climate

7. LUCAS
8. intersecting CLC and countries involved in the application
9. FADN-based AROPAj-dependent information related to farm group's area per crop group

The next steps of the application follow the algorithm described in the previous section, performing the calculations of crop location's priors and posteriors, and the geo-referenced probability sets of farm groups.

We implemented the calculation process for the *V5 AROPAj* version applied to most of EU-27 FADN regions and for the 6 years 2007-2012. This is illustrated for France and the 2009 *FADN* on Figures **??** providing the location probabilities of two of the farm groups in the Bourgogne region.

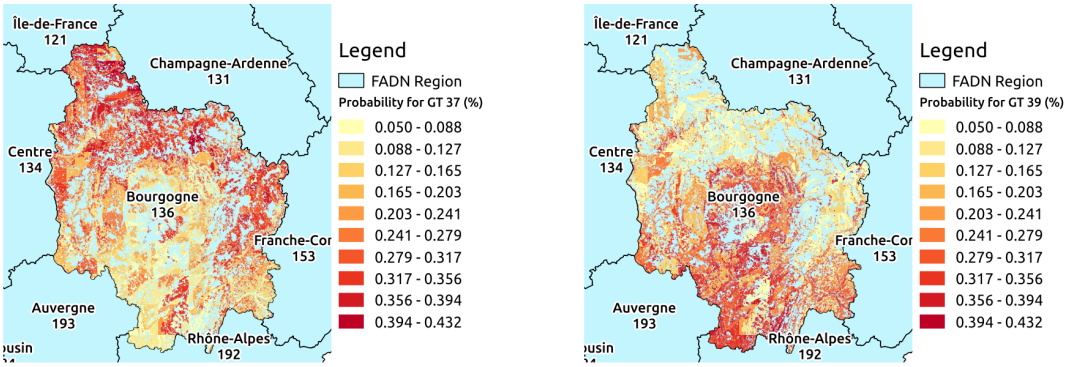


Figure 27. Location probabilities of farm groups "GT37" and "GT39" in the Bourgogne region (2009-*V5 AROPAj* version).

9 Models linkage

The interest of linking the *AROPAj* economic model with biophysical models relies on the possibility of accounting for the impact of technical as well as environmental and physical variables (i.e. soil and climate) on the agricultural production. Moreover, linking *AROPAj* to biophysical models is important for understanding the environmental externalities. Regarding crop production, the general rationale is to introduce in the model a relation between crop production inputs (e.g. fertilizer applications) and crop production outputs (e. g. yields). The possible connecting points between the economic and the biophysical models are the farmer practices found in the accountings, such as irrigation, application of pesticides, and fertilizers. Among those practices, we chose nitrogen (N) fertilization because of (i) its key role in environmental policies in relation to non-point source pollution and greenhouse gas emissions, (ii) its traceability in accounts through fertilizers costs, and (iii) the wealth of usable information for our model on N effects on yield.

9.1 STICS

Key-points:

Introducing crop models and replacing constant yields and inputs by nitrogen to yield functions
The integrated *ArTiX* database and applications

Economic modeling benefits from accessing biophysical information on the systems they represent. In the case of agricultural production, crop models are often associated with economic models, with the information being presented in a simplified form to fit the economic model's constraints. An important variable is the yield of the crop. This depends on environmental factors, for instance climate, as well as farmers' decisions (fertilizer and water use, technical management etc...). In the case of *AROPAj*, the generic crop model *STICS* (**??**) was used to generate data linking the yield of major crops with the technical choices made at the farm level.

The genericity of the *STICS* crop model was one of the main reason for choosing it among all the existing crop models. The potential crop model linked to *AROPAj* needed to simulate yield as well as environmental variables (such as nitrate fluxes) for all of the main crops simulated by the economic model across the *EU* territory. The *STICS* model simulates crop growth, development, and environmental impacts for the main European crops, in the various pedo-climatic conditions of the *EU*. Indeed, the main processes simulated by *STICS* are crop growth and development as well as water and nitrogen soil balances. A of this information is dynamically driven by data such as daily weather data and information regarding soil characteristics and management. The model is organized in several modules: three modules are dedicated to the simulation of plant aerial part (*LAI*, biomass and yield), three to the soil simulation (water and nitrogen balances, and root growth), and the last one additional module dedicated to the management of interactions between cropping techniques and soil–plant system (i.e. water or fertilizer supplies, or climatic environment of the plant). *STICS* is a deterministic model that simulates the soil–plant–weather system in supposedly homogenous physical and technical conditions, which corresponds to each crop of each European farm type of the *AROPAj* model. Regarding the nitrogen sources, the *STICS* model was also able to account for both chemical fertilizers and on-farm produced manures.

9.1.1 N to yield functions

We decided to provide *AROPAj* with response functions $y(N)$ that describes crop yield level y as a function of nitrogen input N. With such a response function, the production process is summarized in a simple function that focuses on the main limiting production factor (N) while also implicitly considering the other ones. No general consensus appears in the literature on the ideal shape of such a response curve (**?** cited in **?**). However, the use of concave and strictly increasing curve is consistent from both an economic perspective (requirements for the estimation of a mathematical optimum), and an agronomic perspective (consistent with expected shape, and defined by easily interpretable parameters). Hence, we selected an exponential function, expressed as the following:

$$y(N) = B - (B - A)e^{-\tau N}$$

where y is crop yield, N is the nitrogen fertilizer amount, A, B, and τ are, respectively, the minimal yield, maximal yield and sensitivity to N input.

To build up the nitrogen response curve, a two-step procedure was set up and repeated for each crop of each farm type. In the first step, *STICS* was run for each crop of each farm type using inputs derived from available databases and expert knowledge (see sections 3.2.2, 3.2.4 and 3.3.1). Each case was thus first characterized by weather conditions and by several combinations of soil and agronomic choices. To avoid too many combinations, the most representative soils and techniques were selected for each case studied: namely, 5 soil types, 3 sowing dates or earliness groups, 2 preceding crops and, given the crop, rainfall or irrigated option were selected for one case studied. This procedure yielded a set of 30 or 60 likely N-response curves.

In the second step, economic data (i.e. yields and price ratios extracted by *AROPAj* from the *FADN*, and from price databases) were used to select the relevant response curve so that a single N-response curve emerges with the required biophysical and economic properties. To do so, we first selected the curves for which the reference yield of the crop, given by the *FADN* database, was reached. Second, we selected the curve for which the slope of the tangent, at the point of the curve corresponding to the above mentioned reference yield value, was the closest to the price ratio (price of nitrogen/price of crop).

In the process, special attention was paid to livestock farms. For these farms, effluents from livestock are spread in addition to (or instead of) chemical nitrogen fertilizer. In this case, the *STICS* model needs additional data to simulate the appropriate N-response curves: first, a curve for the combined organic sources of nitrogen and second a curve of yield against chemical nitrogen only. Indeed, this additional organic nitrogen affects the crop yield but does not have to be represented on the X-axis of a curve showing yield response to chemical nitrogen fertilizer. Consequently, potential organic nitrogen available in each livestock farm type was estimated from the *FADN*. Used as *STICS* inputs (complementary N source apart from fertilizer N), it enabled one to choose the optimal *STICS* input set in step 2. Another run of the *STICS* model, using the selected set of inputs (soil, agronomic choices) then generated the response curve to chemical fertilizer nitrogen itself. The theoretical framework used for the inclusion of different nitrogen sources in the N-response equation was based on an equivalence rule allowing us to compare any N source to a reference chemical nitrogen fertilizer. Once the exponential shape had been selected for the N-response curve, the minimal and maximal yields were assumed to be independent of the N source,

and N sources were then distinguished by the "rate of increase" of the function. Practically, the N-response curve was expressed using the following form:

$$y(\{N_i\}) = B - (B - A)e^{-\sum_i \tau_i N_i}$$

where y is the yield for a given crop on a given farm type, B and A represent the maximal and minimal yield of that crop, respectively, and τ_i represents the rate of increase of the yield response function to a nitrogen source i (e.g. manure, slurry etc.) the quantity of which is N_i.

9.1.2 N to emission functions

In addition to the amount of fertilizers, the *STICS* crop model provides us with the leaching associated to each type of soil and crop together with the climatic conditions and input application date. We derive the leaching functions from *STICS* outputs when nitrogen input varies within a reasonable range. Linear regression applied to *STICS* outputs allows us to relate N-losses and fertilizer amounts (through affine functions). Consequently, the emission function for pollutant i, related to crop j and farm group k, depends on the N-input through relation (**??**); α and β are estimated for a given plant on a given farm group in the form:

$$[e_i]_{j,k}(N) = [\alpha_i]_{j,k}\ N + [\beta_i]_{j,k}\ ;\ \ i = \{NO_3, N_2O, NH_3\}. \tag{17}$$

It should be noted that we focus only on the N-losses at the soil-root level (i.e. NO_3) and on two gas emissions (N_2O and NH_3). The time lag existing between the N-fertilizer application and the actual pollution is not taken into account.

The spatialization method provides the contribution of each farm group on each cell of the grid to the regional ($FADN$) agricultural activities and enables $AROPAj$ outputs to be distributed over France. Once the $AROPAj$ farm groups have been distributed, we aggregate the results in terms of emissions among other model outputs.

For instance, Table (**??**) represents the emissions for NH_3, N_20 and NH_3 for the main crops analyzed in $AROPAj$ in France

Crop	α_{N_2O}	β_{N_2O}	α_{NH_3}	β_{NH_3}	α_{NO_3}	β_{NO_3}
hard wheat	0.0232	0.7106	0.0537	1.6084	0.1081	36.3959
soft wheat	0.0207	0.7923	0.0477	1.8510	0.0366	46.3375
sugar beet	0.0216	0.0526	0.1109	0.3511	0.1412	18.3436
rapeseed	0.0183	0.6038	0.0691	2.1157	0.1854	41.5223
corn	0.0214	0.0779	0.0810	0.2499	0.1998	28.3517
winter barley	0.0075	0.2624	0.0163	0.5724	0.0550	26.4918
potatoes	0.0202	0.0268	0.0469	0.0321	0.2552	2.6380
sunflower	0.0209	0.0105	0.0440	0.0560	0.1331	16.2587

Table 28. N-loss i-emission function for crop j, $i = \{NO_3, N_2O, NH_3\}$: $[e_i]_j(N) = [\alpha_i]_j\ N + [\beta_i]_j$; β denotes natural nitrogen loss in kgN/ha/year.
source: **?**

9.1.3 *ArTiX*

We couple the *AROPAj* model to the *STICS* crop model in order to estimate input-output relationships (see above sections) for particular crops and the economic agents considered in the *AROPAj* model. The combination of an economic agent and a crop is hereafter referred to as a **crop-agent case**. For a specific crop-agent case, we estimate the input-output relationships from several simulations of the *STICS* crop model, while using various sources of data defining its physical (soil, altitude, slope, climate) and technical (such as sowing date, precocity group of the cultivar used, use of irrigation or not, fertilization calendar, preceding crop etc.) environments. This process thus requires the manipulation of various types of geo-referenced data as inputs to the *STICS* crop model, as well as simulation results. In order to automatize such a process for the numerous crop-agent cases and scenarios, we developed an *informatic framework* named **ArTiX** (described in this section).

ArTiX designates both a *database* containing various types of data, and a *software* interfacing the database with the *STICS FORTRAN* code and the user:

The *ArTiX* database. It is a *Postgresql* database, containing data in the form of tables and relations. This data is geo-referenced with either an *FADN* region identifier or an identifier of another spatial distribution (vector or raster form), and describes:

- All the parameters required by the *STICS* model to describe its physical and technical environment (e.g., descriptors of soil texture, daily weather, etc.),
- All the *STICS* output relevant for futher processing (e.g., crop yield, fertilizer application doses and dates, dates of phenological stages),
- Additional information concerning crop-agent cases (e.g., crop output selling price, etc.),

The *ArTiX* software. It is a software coded in Java, that allows the following functionalities:

- To perform simulations for a particular crop-agent case specified by the user. This includes automatic query of input information required by the *STICS* crop model, automatic generation of *STICS* input files, launching of *STICS* simulations and management of errors, and reading of output files from which results of interest are read and stored in the database.
- To allow the user to define a particular setup for the simulations to be done (e.g., scenarios for climate data, or crop management practices).
- To allow the user to query data of any type, contained in the database, via *psql* queries.

The overall concept is summarized in **??**, illustrating the database, the java code, and the user interface.

9.2 *Carbo-PRO*

Key-points:

Improving the insertion of organic matter sourced nitrogen

Applying the method to urban organic residue (*UOR*)

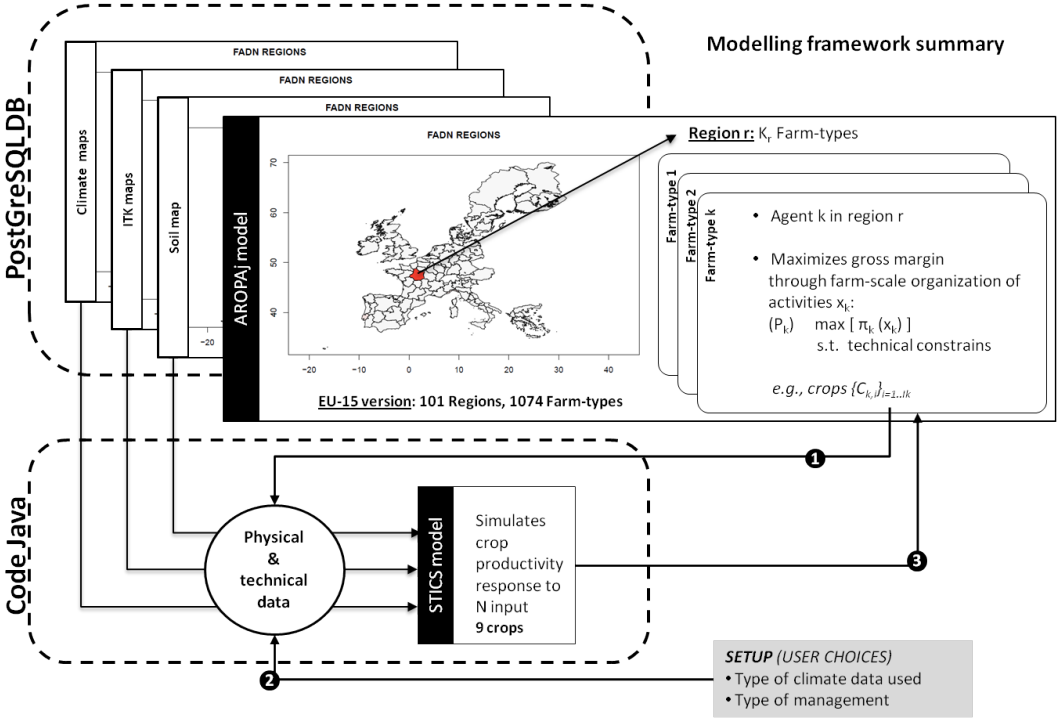


Figure 28. Schematic overview of the *ArTiX* database and Java interface between the *AROPAj* and *STICS* models, and the user.

We develop a method for introducing organic N sources in the *AROPAj* model, thereby completing the *LP* N-block presented in Section **??**. Manure is an important endogenous *AROPAj* activity related to livestock. Urban wastes are exogenous N sources. Agronomic long term effects have been introduced. This consideration can be attributed to the use of the *Carbo-PRO* model.

The method was applied to the French Ile-de-France context. It allowed to provide the value of the nitrogen sourced by *UOR* when considering the intensive effect of N-spreading at the plot level, the extensive effect of land cover change among crops at the farm-type level and the N allocation between farming system at the regional level (**?**).

9.2.1 Organic nitrogen sources

The nitrogen block includes new forms of nitrogen input : urban organic residues (*UOR*). At present, parameters are available for two types of *UOR* : green waste compost (*GWC*) and urban sewage sludge (*USS*)[8]. At the plot scale, they are considered as any other source of organic matter (*OM*), *i.e.* livestock manure.

Integrating long term effects of organic matter spreading in the yield-to-nitrogen response function

An enhanced form of the response function has been developped in order to integrate the long-term effect of organic nitrogen spreading. Organic residue spreading has a short-term fertilization effect (content in nitrogen, phosphorus and potassium, impact on biological and pH properties of soil). It also has a long-term effect, increasing soil *OM* and so supporting soil properties (water

[8]This work emanated from the PRO-EXTERN project from the French Environment and Energy Management Agency (*ADEME*)

retention, cation exchange capacity, nitrogen-to-plant availability). This leads to some extra nitrogen available to the plant, provided by the slow mineralization process of additional soil humus. *OM* spreading could then be considered as a double source of nitrogen for crops via short and long term effects.

The long term effect is assumed to modify the A parameter of the yield function, B potentially remaining unchanged (see the yield function and the parameters to which it refers on page **??**). The difference between the function that integrates a long term effect ($Y^{LT}(N^m,N^o) = B-(B-A^{LT})e^{-\tau(N^m+\alpha^{ST}N^o)}$) and the initial short term function ($Y^{ST}(N^m,N^o) = B-(B-A^{ST})e^{-\tau(N^m+\alpha^{ST}N^o)}$), leads to the following:

$$Y^{LT}(N^m,N^o) - Y^{ST}(N^m,N^o) = (B-A^{ST})\left(1-\frac{B-A^{LT}}{B-A^{ST}}\right)e^{-\tau(N^m+\alpha^{ST}N^o)}$$

in which the term $\left(1-\frac{B-A^{LT}}{B-A^{ST}}\right)$ remains within the interval $]0,1]$ and can be expressed as $e^{-\tau\alpha^{LT}N^o}$ (with α_l^{LT} being positive).

The exponential form of the Mitscherlich-Baule function suggests that the long term yield function can be thought of as a leftward translation of the short-term one. This effect can be translated as a decrease of the nitrogen needed to reach a given yield level, all other things being equal. Given the amount of organic nitrogen application, the mineral nitrogen gap between short term and long term leading to a given yield is constantly equal to $\Delta N = -\frac{1}{\tau}\log\left(\frac{B-A}{B-A^{LT}}\right)$, as represented in Figure **??**.

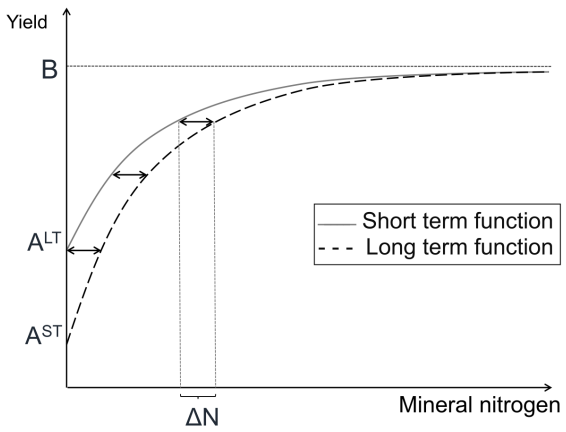


Figure 29. Long term effect of *OM* spreading on the N-to-yield response function.

The nitrogen-to-yield function, wich is generalized to L sources of organic nitrogen and which integrates both short and long term effects of organic matter spreading, is amended and as expressed by the following equation (with remarkably unchanged A and B values):

$$Y^{LT}(N^m,N_1^o,N_2^o,...,N_L^o,) = B-(B-A^{ST})e^{-\tau(N^m+\Sigma_{l=1}^{L}(\alpha_l^{ST}N_l^o+\alpha_l^{LT}N_l^o))} \tag{18}$$

The α_l^{LT} coefficient, related to the long term effects of organic matter spreading, is obtained via the *Carbo-PRO* tool (see www.CARBO-PRO.fr and section **??**).

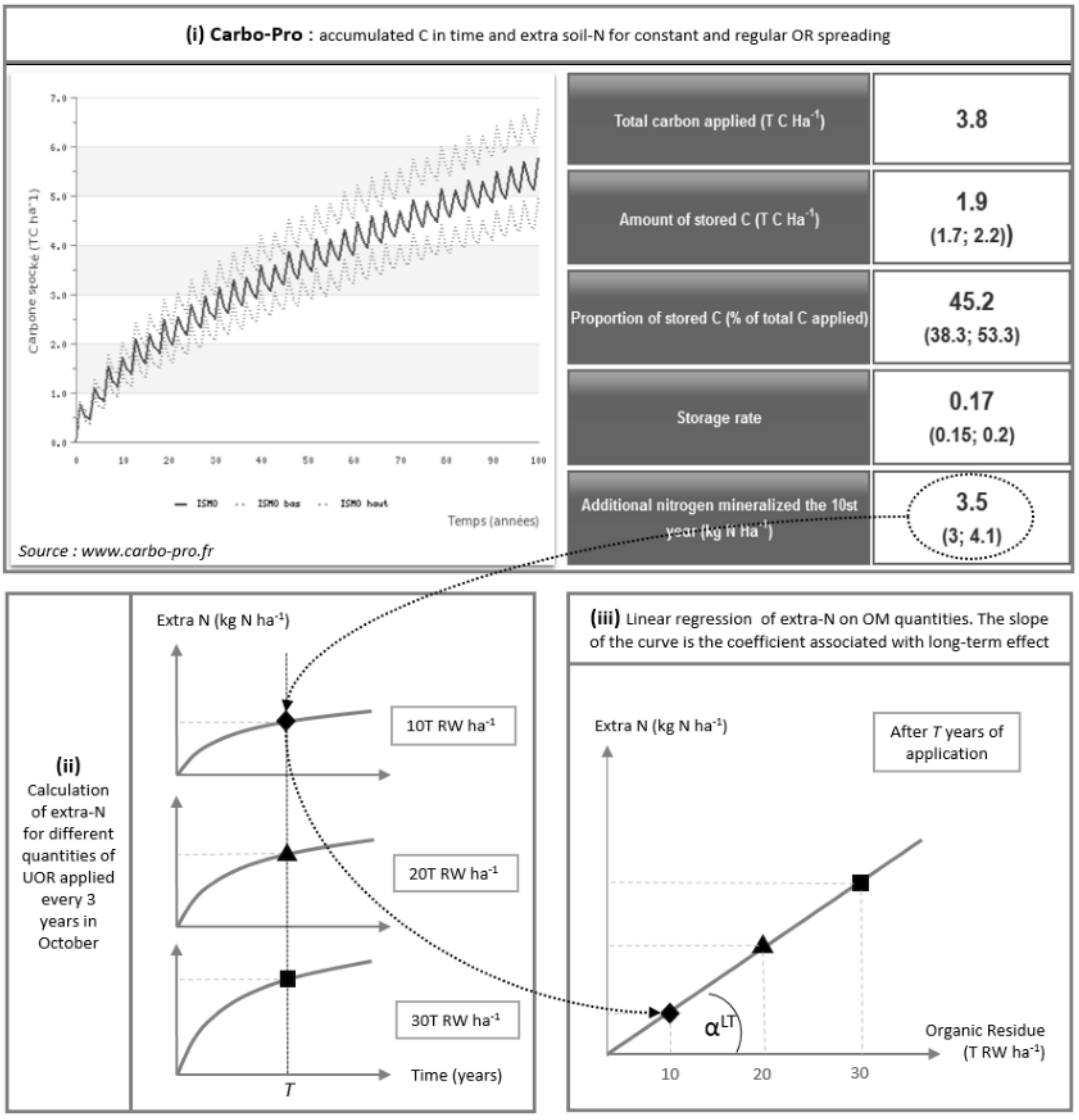


Figure 30. Illustration of the procedure to get coefficients associated with long-term impact of organic residue from *Carbo-PRO*.
Source: **?**

Modification of the optimization problem for UOR integration as organic nitrogen sources.

At the regional scale, a constraint relative to the UOR availability is added. Let us focus on the constraints associated with nitrogen supplies (N^m_{jk} and N^o_{jkl}) and extract:

- positivity constraints;
- the per area unit limit of organic matter spreading ($\bar{Q}$ not depending on the farmer k neither on the crop j, in accordance with real regulation);
- the upper bound of available organic matter M_{lk} referring to the l organic source. This constraint only concerns UOR availability, since livestock manure quantities are endogenous to the model (related to livestock management) and are not included in organic matter exchanges between farm types. This constraint allows for the transfer of UOR between farm-types. Optimal sharing is achieved when all farm-types' individual shadow prices are equal. If l refers to UOR, the regional amount of related organic matter M_{lk} is an external parameter. If l refers to animal-sourced organic matter, the M_{lk} amount refers to endogenous (bounded) animal activities (livestock and manure management).

Optimal UOR availability sharing is obtained via the following aggregated regional program:

$$\left|\begin{array}{lrl}
 & \max\limits_{x_k,a_k,N^m_{jk},N^o_{jkl}} \sum\limits_{k=1}^{K} \pi_k(x_k,a_k,N^m_{jk},N^o_{jkl}) & \\
 & \forall k \;\; \forall j: \;\; F_k(x_k,a_k,N^m_{jk},N^o_{jkl}) \leq O & (\psi_k) \\
 & \forall k \;\; \forall j: \;\; N^m_{jk} \geq O & (\beta^m_{jk}) \\
\text{s.t.} & \forall k \;\; \forall l \;\; \forall j: \;\; N^o_{jkl} \geq O & (\beta^o_{jkl}) \\
 & \forall k \; \forall j: \;\; \sum\limits_{l=1}^{L} N^o_{jkl} \leq \bar{Q} & (\lambda_{jk}) \\
 & \forall l: \;\; \sum\limits_{k=1}^{K}\sum\limits_{j=1}^{J} N^o_{jkl}\; a_{jk} \leq \sum\limits_{k=1}^{K} M_{lk} & (\xi_l)
\end{array}\right. \qquad (\mathscr{R})$$

Implementation of the long-term effect in $AROPAj$ simulations.

Thanks to the *Carbo-PRO* tool (see www.CARBO-PRO.fr and section **??**), we obtain a long term effect parameter (α^{LT}_l) for each organic matter source in *AROPAj*. Integrating long term extra-N as a reduction of the nitrogen need of plants makes it difficult to keep the LP structure of the model. Indeed, non linearity of organic matter $N_{o,l,j,k}$ enters the optimization program when long term impact of repeated organic matter application is taken into account. Linearity is ensured when optimal $N_{o,l,j,k}$ tends to converge toward the value for which long term effect is estimated. We use the iterative computational process below, where the short term $N_{o,l,j,k}$ computed at a given iteration is used to estimate long term effect for the following iteration:

1. The following two-step subprocess is implemented:
   (a) Market prices replace shadow prices to solve the plot optimization programs (see section **??**). Null-values are considered for additional long term impact of organic nitrogen.
   (b) The (**??**) regional optimization program is solved, doing so for yields and nitrogen applications provided by the step (a).

2. Shadow prices required for pre-solving the plot optimization program are extracted from the previous calculus' step At the same time, quantities of organic matter spread are extracted, and average quantities of organic nitrogen are computed against land endowment.
3. The three-step subprocess is implemented:
   (a) Shadow prices and long term organic nitrogen amount provided by the previous step are injected in the plot optimization program.
   (b) The (**??**) regional optimization program is solved, doing so for yields and nitrogen applications provided by the step (a).
   (c) Go to step (2.) until a predefined number of iterations is reached or a predefined convergence threshold is attained.

9.2.2 Carbo-PRO

The *Carbo-PRO* tool, designed by *INRA* (Department of "Environnement et Grandes Cultures") and the *Veolia* company, is used to get coefficients associated with the long term effects of organic matter spreading (α_l^{LT}). More details on the long term effect are available in the section focusing on nitrogen. *Carbo-PRO* aims at computing the carbon stock in soil that is created by repeated organic residue application. It is based on the Henin-Dupuis model of soil carbon dynamics (**?**) and on the *RothC* model that simulates the evolution of organic matter (**?**). It also provides estimations of soil stability improvement, water retention, cation exchange capacity and quantity of nitrogen stocked in soil after many years of mineralization process.

We only use the nitrogen module of *Carbo-PRO*. The module is concerned with the process of nitrogen mineralization related to organic matter degradation that renders organic nitrogen available for plant. This extra soil-N is estimated as the product of soil nitrogen content and the $K2$ isohumic rate, or rate of mineralization of the soil-organic matter. More details on *Carbo-PRO* operation are available on www.CARBO-PRO.fr.

To get coefficients associated with long term effects of organic matter spreading, a spreading scenario has to be entered in *Carbo-PRO* with information regarding soil type, climate and spreading specifications (application rate and frequency, humic potential indicators, dry matter and organic matter content). The procedure is presented in Figure **??**: For a given quantity of organic matter, *Carbo-PRO* provides the soil extra-N quantity. This extra-N at the end of the spreading period is obtained for different quantities of organic matter spread. A relation between extra-N and OM quantity at a given time can then be traced. As the relation is linear, the α_l^{LT} coefficient corresponds to the derivative of this function, related to the number of years of spreading.

The procedure has to be repeted for each source of organic nitrogen. We obtain one parameter for every type of organic matter. In the core of *AROPAj*, the extra-N is integrated as a reduction of the nitrogen need of plants (Needs - "long-term extra-N" $\leq$ Sum of $N_{m/o}$-inputs), requiring the iterative process described in section **??**.

9.3 *PaSim*

Key-points:

Overcoming of the lack of *STICS* sourced response functions for pastures and fodders

Rough replacement of fodder yields by *PaSim* in different climate change scenarios

The pasture simulation model (*PaSim*, (**?**)) is a multi-year, process-based biogeochemical grassland model, which is based on the Hurley Pasture Model ((**?**)). It simulates fluxes of carbon, nitrogen, water and energy at the soil-plant-animal-atmosphere interface for managed grasslands at the plot scale. Simulations may run over one or several years.

PaSim simulates grassland processes on a time step of 1/50th of a day. It provides detailed sub-daily dynamics and ensures energy budget stability. A fixed legume fraction in the grasslands is assumed in the *PaSim* model, in which vegetation is represented by mixed sward, when grasses and legumes co-exist. *PaSim* was developed to simulate the meat and milk production of cattle in response to different climate and management conditions. Animal processes are limited to the plot scale when animals are grazing at pasture (not in the barn). *PaSim* has been used to simulate greenhouse gas fluxes over grasslands and to examine the role of grasslands in the European greenhouse gas budget ((**???**)).

The method for the coupling of the *AROPAj* economic model and the *STICS* crop model (**?**), developed by **?**, is best adapted to major agricultural crops (cereals, oilseeds, etc.). Hence, it is not still appropriate for forage and grass plants, due to the lack of information in terms of productivity and prices. In order to obtain a reliable measurement of grass and fodder yields, we have coupled the *AROPAj* model and the *PaSim* grassland system simulation model. This allows us to better measure the grass yield and the interactions between climate, crop system and more generally the production intended for animal feed (see **?**).

To this end, the model was run as explained in Section**??**.

The level of nitrogen input at the European scale for the *A2H2* far future scenario is given in figures in Table **??** [9]. As shown, the amount of nitrogen applied varies from 0 to 0.10 metric ton of N per hectare (tN/ha) in most European Union countries, except for Germany, Spain, Ireland, the United Kingdom and central-eastern regions of France, with a mean variation in nitrogen supply from 0.10 to 0.30 tN/ha.

The *PaSim* model provides the grass yield (in dry matter) for each pixel for the entire Europe at three levels of nitrogen intensity, which is different from the agricultural yield in the *AROPAj* model where yields are estimated in raw material. This generates a group of points, in which a N-yield response curve was statistically adjusted for each crop and *EU* region. These dose-response N-yield functions provided by the *PaSim* model replace the previously calibrated values in the *AROPAj* model for fodder beet, fodder maize, permanent grassland, temporary meadows and other grasslands (figures in Table **??**).

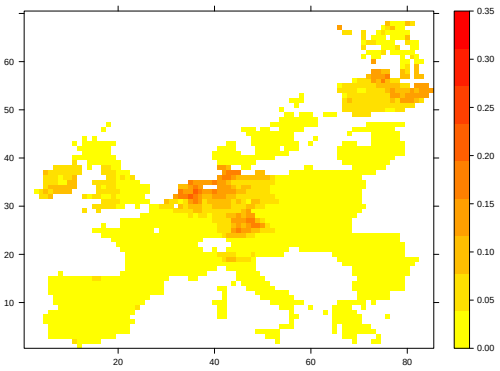
9.4 *MODCOU*

Key-points:

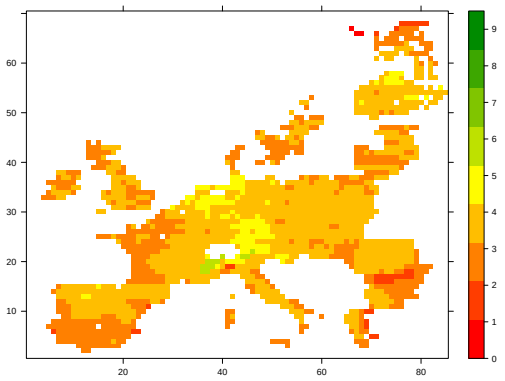
How to couple *AROPAj* with a model dealing with subsoil losses of nitrogen at a fine spatial resolution level

Modeling experiments with the hydro-geological model *MODCOU*

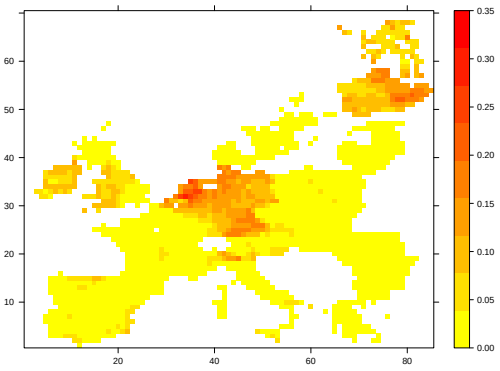
[9] The *A2H2* and *B1H2* scenarios are among those proposed by the *IPCC* to explore far future impacts of climate change (for more explanation, see for instance: http://www.ipcc.ch/ipccreports/sres/emission/index.php?idp=98).



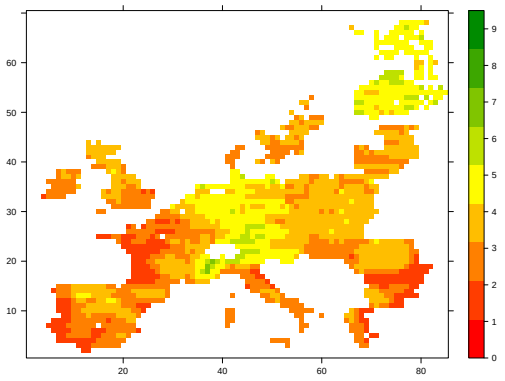
(a) Nitrogen input – Low intensity



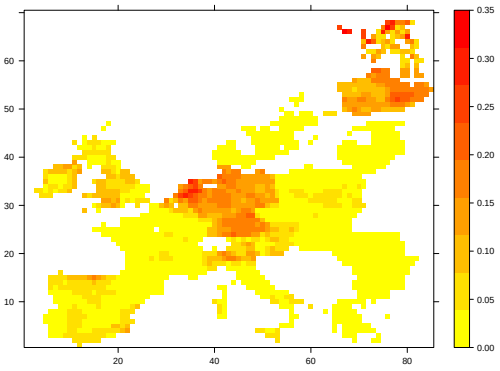
(b) Grass yield – Low nitrogen intensity



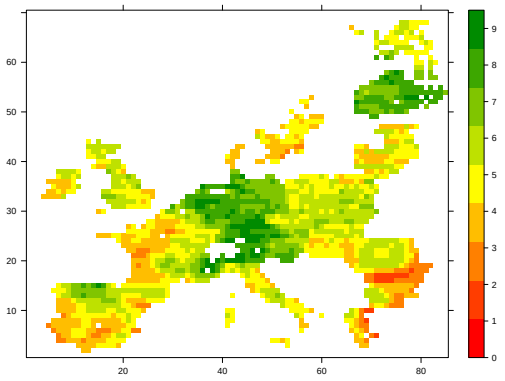
(c) Nitrogen input – Medium intensity



(d) Grass yield – Medium nitrogen intensity

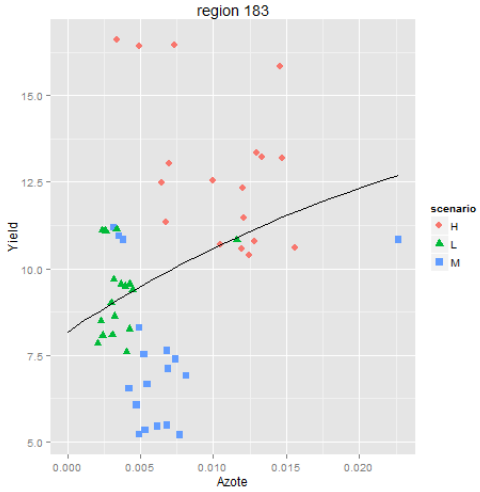


(e) Nitrogen input – High intensity

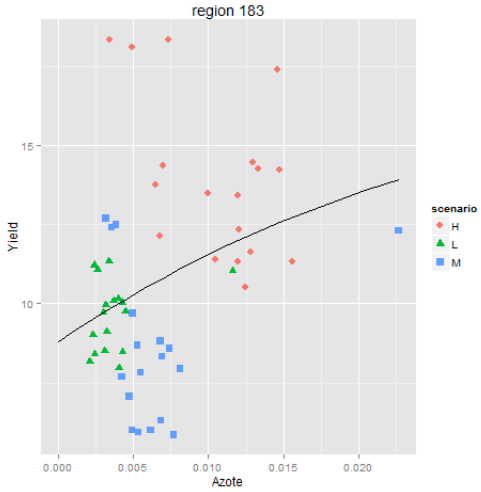


(f) Grass yield – High nitrogen intensity

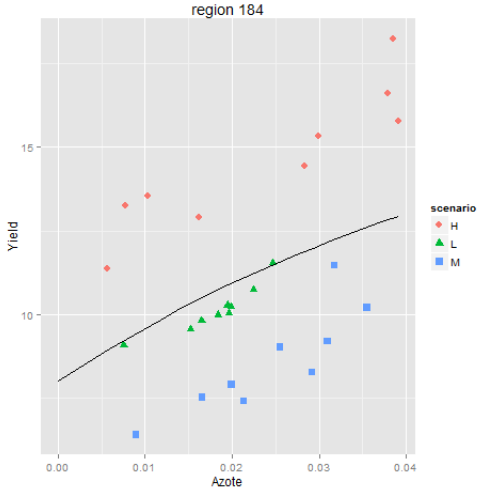
Table 29. Results of the sensitivity analysis of fodder maize yields to nitrogen input in three French regions under two climate scenarios *A2H2* (left) and *B1H2* (right) over the far future 2070-2100 period.



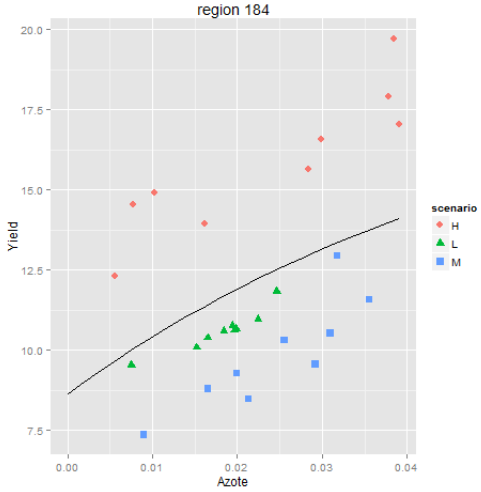
(a) *A2H2* Midi-Pyrénées



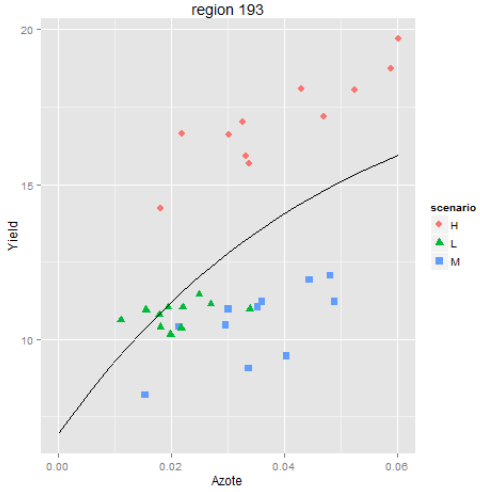
(b) *B1H2* Midi-Pyrénées



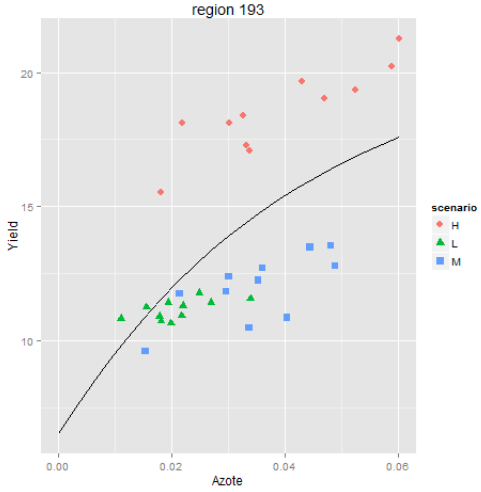
(c) *A2H2* Limousin



(d) *B1H2* Limousin



(e) *A2H2* Auvergne



(f) *B1H2* Auvergne

Table 30. The results of the sensitivity analysis of fodder maize yields to nitrogen input in three French regions (Midi-Pyrénées, Limousin, Auvergne) under two climate scenarios *A2H2* (left) and *B1H2* (right) over the far future 2070-2100 period.

In this section, we present a method of linking *AROPAj* with a hydrological model, *MODCOU*, in order to estimate the evolution of a pollutant concentration in aquifers when an environmental policy modifies the amount of fertilizer used by farmers.

The *MODCOU* application focuses on the Seine basin (95,560 km^2) as well as three major aquifers with different geological characteristics: the Oligocene (sands and limestone), the Eocene (sands and limestone), and chalk (Cretaceous chalk), with a spatial resolution varying from 1 to 8 km (**???**).

The use of the hydrological model *MODCOU* traces the nitrate transfer from the root zone toward the groundwaters of the Seine river basin. The hydrological model *MODCOU* (**???**) computes the daily water balance using climatic data such as rainfall and potential evapo-transpiration (*PET*), the water flow to and in the river network, and also the flow to, in, and between aquifer layers and the interactions between rivers and each layer of the free part of the aquifer. It is also able to compute the convective transfer of solute to, in, and between those layers.

Regarding the *MODCOU* model, the initial condition (in 2002) can affect the simulation over several decades due to transfer time through the unsaturated zone (*UZ*) and aquifer nitrogen accumulation. Thus, special attention was paid to obtaining realistic initial conditions for the *AROPAj-MODCOU* simulations. Initialization is based on the *MODCOU-STICS* simulations (as described in **?** and in **?**), and is aimed at correctly reproducing the nitrate stock in the *UZ*, among other things, at the starting date of simulations (2002). In such simulations, *STICS* provides estimates of the nitrate flux entering in the *UZ* over the period 1971-2002 and reaching the aquifer from 2003. Moreover, *MODCOU-STICS* provides the NO_3 reference flux as the mean value of 2002-2004 fluxes (see Figure **??**).

Coupling of *AROPAj* and *MODCOU* comes into play from the simulation starting date (2002). In this respect, we use the *AROPAj* spatialization process which geographically distributes *AROPAj* output at a fine level of resolution (see Section **??**). The spatialization method provides the contribution of each farm group on each cell of the grid to the regional (*FADN*) agricultural activity. This enables any *AROPAj* output to be distributed over the geographical area, i.e., the Seine river basin. Soil-root nitrate losses related to fertilizer use are then distributed on a grid compatible with the spatial resolution of the hydrological model. Following the temporal tax path designed by the theoretical economic model exposed in **?**, spatialized simulated NO_3 fluxes provided by the *AROPAj* model feed *MODCOU* by topsoil nitrate changes according to the formula:

$$\begin{aligned} NO_{3input}(year) &= Mean[NO_{3MODCOU-STICS}(2001-2003)] \\ &* \quad [1+\frac{NO_{3AROPAj}(year)-NO_{3AROPAj}(2002)}{NO_{3AROPAj}(2002)}] \end{aligned} \tag{19}$$

For each year, the signal transmitted to *MODCOU* is based on the average emissions provided by the coupled physical models (*MODCOU-STICS*) for the period [2001,2003], and is augmented with the annual variation provided by *AROPAj* and then related to the temporal tax path.

Finally a "scenario" is defined as a set of time-ordered annual *AROPAj* inputs from date 0 up to a given horizon T. Any of the t-annual *AROPAj* runs provides an NO_3 lixiviated flux which is transferred to the groundwater by the dynamic hydrological model *MODCOU*. A scenario gathers different incentive instruments that can be implementable into the *AROPAj* model. It could be an

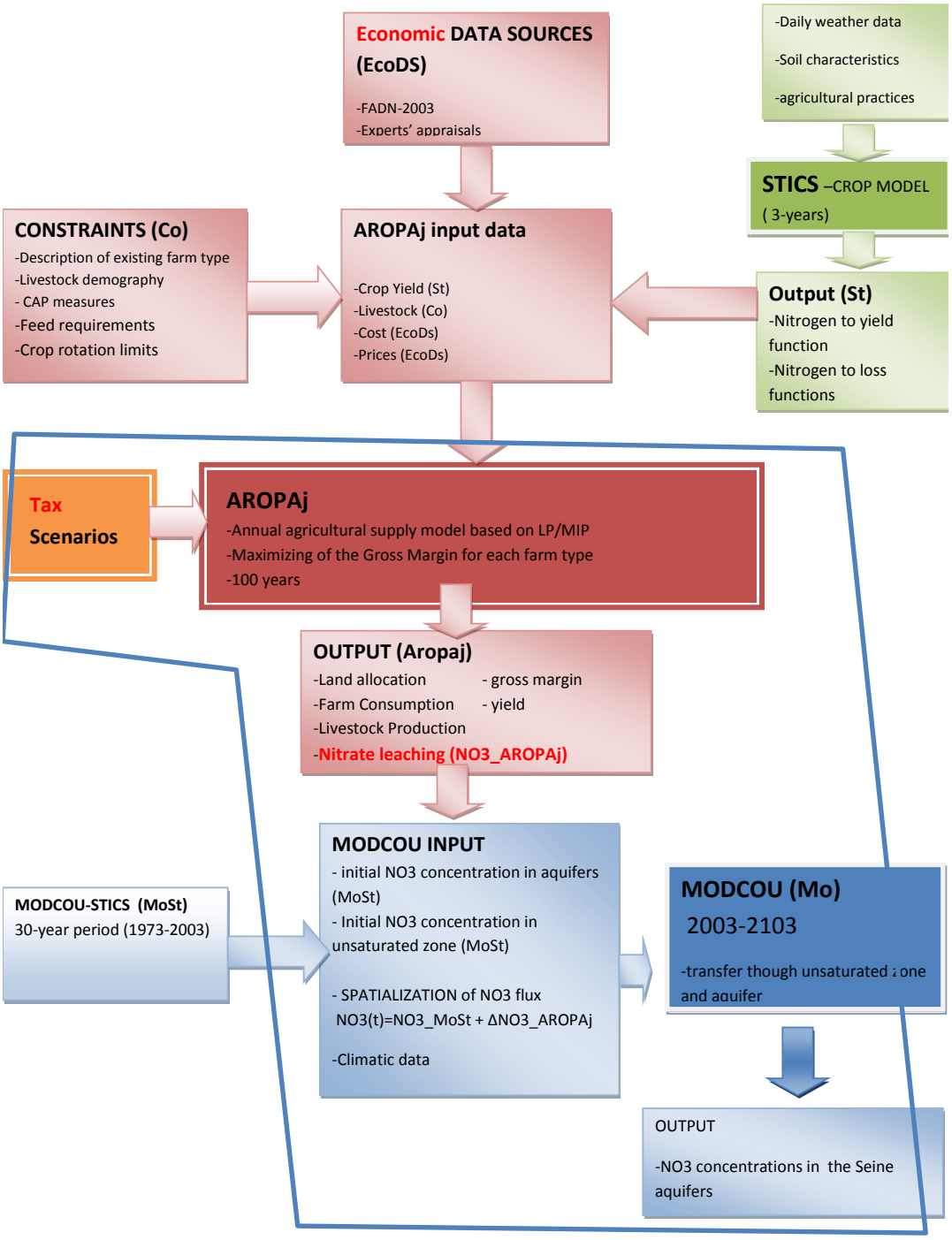


Figure 31. Schematic representation of the coupling methodology. The colors refer to the themes: green for agronomy, blue for hydrology and red for economics. Darker colors refer to models, and lighter colors to inputs/outputs. The method for building the tax scenarios is described in the box below. The information which underpins this work is depicted by the blue line, some *AROPAj* input data and the *MODCOU* initial conditions being given from previous works.
Source: **?**

input or emission tax and/or a subsidy to environmentally-friendly crop (existing in *AROPAj* like miscanthus, grassland). We can also modify the livestock adjustment. For the economic analysis, the NO_3 concentration in the aquifer is calculated as the annual median value for the entire chalk aquifer.

9.5 *GLOBIOM*

Key-points:

Attempts to integrate price and yield feedbacks related to clearing markets at a global scale

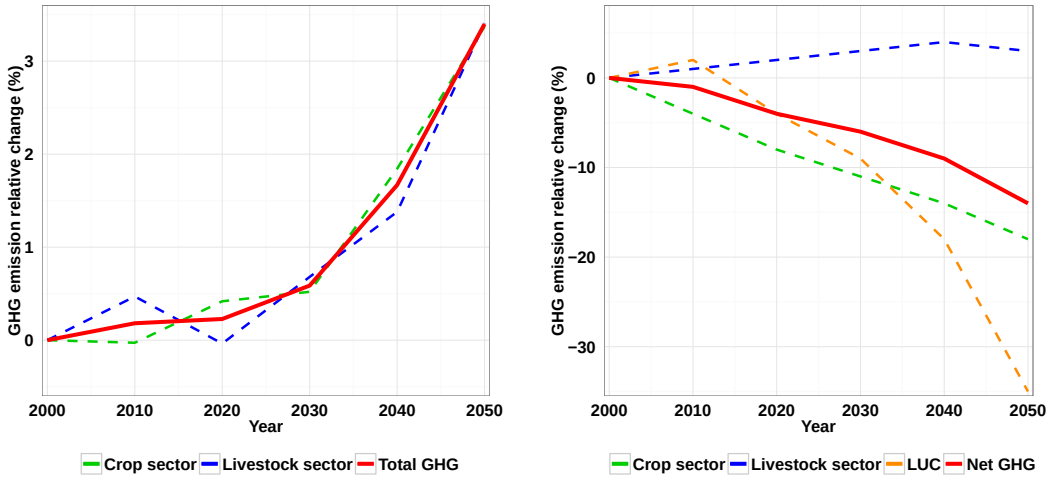
Focusing on crop yields and prices more than fertilizers and feed at a meso-economic scale (countries or sets of countries within the *EU*)

The *GLOBIOM* model is a global, recursive and dynamic, partial equilibrium model designed to aid policy analysis of issues involving land use competition among the major land-based production sectors, particularly agriculture, forestry, livestock and bioenergy (**?**). Moreover, one of the main particularities of *GLOBIOM* is that it makes intensive use of biophysical data.

GLOBIOM is an optimization model based on the maximization of global welfare, defined as the sum of consumers' and producers' surpluses. For each output product defined in the model (i.e. each crop or forest product), a reference level is defined that is subject to an exogenous increase over time on the basis of population and gross domestic product (*GDP*) growth. Besides, the demand for a given product at each time step depends on the product's own price. Prices are homogeneous within each market regions and there about 30 market regions comprising the global scale. Interregional trade, subject to transportation costs and tariffs, is permitted between these regions.

The supply system is based on a detailed representation of land and land allocation. The world is divided into simulation units (*SimU*) that are homogenenous in terms of agro-climatic potential (soil, climate). There are more than 200k land units worldwide when the model is run at its most detailed resolution. Land shares allocated to the different uses within each simulation unit are the main endogenous variables on the supply-side in *GLOBIOM*. Prior to this optimization process, crop yields are determined on the *SimU* level thanks to a biophysical model (in this case the crop model *EPIC*, **?**) fed with SimU-specific climate and soil data. These yields are then introduced as parameters in *GLOBIOM* (the approach is similar for the forestry sector). However, yields are not completely fixed in *GLOBIOM*, they are allowed to vary both exogenously (e.g. exogenous technological progress) and endogenously. As for these endogenous changes, within a given *SimU* and for a given crop, there can be switches between different production systems (namely subsistence, low input, high input, irrigated), yields corresponding to the different systems are all determined by the *EPIC* cop model. The parameters corresponding to the livestock sector are determined on the basis of input-output coefficients provided by the *RUMINANT* model (**?**) for ruminants and from a literature review concerning monogastrics.

The model is dynamic and is run on several time steps on the basis of an exogenous baseline (population, *GDP*, technological progress, climate change...). It is also recursive as land-use changes between two time steps are limited in their extent and therefore depends on the state at the previous time step.



(a) European annualized *GHG* emissions (based on *AROPAj* results)

(b) *ROW* annualized *GHG* emissions (based on *GLOBIOM* results)

Table 31. Difference in European (a) and *ROW* (b) annualized *GHG* emissions (%) comparing scenarios *SSP1* to *SSP3* over time.

There are several points that justify the linkage between *GLOBIOM* and *AROPAj*. At first, endogenous output prices from *GLOBIOM* can be fed as exogenous parameters into *AROPAj*. Then, land-use projections from *GLOBIOM* could be forced into *AROPAj* so as to investigate the impact on both output and non-land input markets. At last, comparing the results from both models may help to identify stylized facts.

The *AROPAj-GLOBIOM* forcing process relies primarily on the application of future yields and price trend variations in order to evaluate economic and environmental impacts on the *EU* agricultural system until 2050. To do so, two macroeconomic scenarios of Shared Socioeconomic Pathways (*SSP1* and *SSP3*), developed under the aegis of the *IPCC-AR5*, have been used.

The combination of the two following assumptions has been used: (i) yield and price variation estimated for two *SSP* scenarios from 2000 to 2050 and provided by the *GLOBIOM* model; and (ii) six levels of livestock adjustment, defining the possibility of farmers to adjust their livestock numbers, from 0% in 2000 to 75% in 2050 (by step of 15% every ten years)[10].

The *AROPAj-GLOBIOM* forcing results show that, although the *SSP1* scenario represents a sustainable world with rapid technological progress towards more environmentally friendly processes, it would provide negative impacts on European grass cover and *GHG* emissions (3.4% in *SSP1* compared to *SSP3*) in the near future (until 2050), as a result of overall developments on both agricultural yields and on the demand side (i.e. increases in population and *GDP*). It seems that lower agriculture product prices in *SSP1* lead farmers to increase cropland so as to obtain higher revenues. In the rest of the world (*ROW*), the *SSP1* may help mitigate climate change and reach emission targets by reducing land use change emissions by 35% and net *GHG* emissions by 15% by 2050 compared to *SSP3* (see Figure **??**).

[10] Aghajanzadeh-Darzi et al, (2015), Global change challenges for the EU agricultural sector: a multi-scale model analysis, submitted

9.6 *GeoMIRET*

Key-points:

Accounting for bio-energy production and bio-energy plants
Benefiting from other modeling experiments at the French scale

We analyse the interactions between cellulosic biomass supply and the cellulosic biofuel demand through the coupling of *AROPAj* and the energy-transport system model *TIMES GeoMIRET*, developed by IFP Energies nouvelles (*Ifpen*). Under public policies, demand and resource availability constraints, *TIMES GeoMIRET* optimizes the French energy mix and determines the location of new biofuel facilities while taking into account all possible energy uses of biomass (heat, electricity or biofuels) and the location of biomass resource basins, refineries, first generation (1^{st}G) existing ethanol facilities, harbours and imported biomass entries on the Belgian and German borders.

The *GeoMIRET* model is a partial equilibrium model applied to the French energy sector and built with *MARKAL/TIMES*, the techno-economic model paradigm developed by *IEA-ETSAP*. It is a long-term, multi-period, techno-economic planning model that covers the energy-transport system. Aiming at providing final energy services at minimum cost, the system encompasses all the steps from primary resources in place to the supply of the energy services demanded by energy consumers, through the chain of processes which transform, transport, distribute and convert energy into services. These processes are mainly driven by the reference demands for energy services (e.g., mobility for the transport sector), the supply of the resources (e.g., amount availability and/or price level), along with public policy constraints, which are provided as exogenous inputs to the model.

Technologies are described in *TIMES* by technical and economic parameters. The choice of the technology generation is conditioned by 4 components : energy service demands, primary resource potentials, policy settings, and description of the technologies set. The demand component includes French aggregated heat and electricity demands, mobility demands for short and long distance for passenger and freight transport and demands for exported products (oil products, electricity). The primary energy supply includes imported fossil energy (crude oil, coal, natural gas), imported bioproducts (palm oil, soybean oil and wood biomass) and locally-produced biomass (sugar beet, wheat, corn, rapeseed, sunflower, miscanthus, switchgrass, willow, forest wood, residues/straw). The energy technology component includes process-detailed oil refining, biofuels facilities ($1^{st}G$ ethanol and bio-diesel, Fatty Acid Methyl Ester (*FAME*) and Hydrotreated Vegetable Oil (*HVO*)), lignocellulosic ethanol and Biomass-to-Liquid (*BtL* or Diesel Fischer-Tropsch), electricity generation, preparation of fuels for transport at blending, and end-use technologies for road mobility (private passenger cars, public transport vehicles, light and heavy duty vehicles). The policy component includes a variety of micro measures (technology portfolios, vehicle fleet with different engine technologies), the national renewable energy action plan and broader policy targets (sectoral carbon tax, renewable energy and fuel quality directives).

AROPAj and *TIMES GeoMIRET* have different spatial distribution scales. While *AROPAj* results can be aggregated at the administrative regional scale, *TIMES GeoMIRET* gathers input data into administrative departments which are clustered into supplying-lignocellulosic biomass basins. These basins are defined as potential regions for supplying second generation (2^{nd}G) biofuel

facilities that correspond to homogeneous facilities in terms of the available biomass type and pedoclimatic conditions. Coupling the two models involves the systematization of the spatial disaggregation of *AROPAj* output to a suitable *TIMES GeoMIRET* spatial scale. We proceed therefore to overlay the *AROPAj* polygons with the administrative department boundaries, by using the Geographic Information System (*ArcGIS*).

The departmental biomass quantities were grouped in supply basins according to an unsupervised classification algorithm, i.e. k-medoids algorithm (*PAM*). Algorithm inputs were represented by department centroids with longitude and latitude coordinates and agricultural and forest availabilities per department. Outputs were 9 partitions of departments, each of which fed biomass to a facility assumed to be probably located at the cluster barycentre or harbour site or 1^{st}G ethanol site. Once the contiguous and homogeneous regions are defined, we calculate the supply distances between the barycentre of each biomass basin, refineries, 1^{st}G ethanol facilities, harbours and entry points in order to estimate the costs of transport, by applying the great circle equation used to determine the straight-line distance between two points.

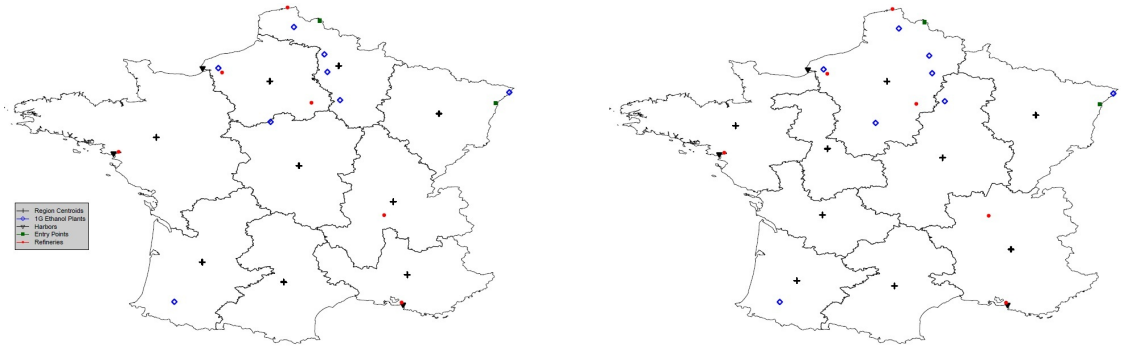


Figure 32. Geographical distribution of biomass resource basins, refineries, ethanol facilities, harbours and imported biomass entries according to different biomass supply scenarios.

9.7 *OURSEurope*

Key-points:

Geographical location and vertical market integration accounting for biomass transformation into biofuels at he European scale

Spatialized three-sector multi-agents viewed through building an operating model cluster

Europe is the first biodiesel market at the global level, and similarly, *AROPAj* is a European agricultural supply model. Therefore, it is interesting to associate it with an agricultural demand model for energy purposes, in order to assess the feasibility and the potential of the biofuel sector in particular. Indeed, the willingness for energy independence, for a boost of the local economy and for a reduction in pollutant emissions has supported the European bioenergy development. Target levels of bio-fuel incorporation in fossil fuels are applicable in most European countries, despite the criticism regarding the environment, the budget constraints, and the arising conflicts on biofuels

and the current economic crisis. In response to environmental criticism, eligibility requirements have also been established and the percentage of certified production is constantly increasing.

The objective of a coupling between *AROPAj* and a model simulating agricultural demand for biofuel production is to estimate the potential of domestic biofuels production, their profitability and the factors limiting their development as well as their environmental and economic impacts. A massive use of biofuels can indeed change the allocation of agricultural land and the profitability of agents. To study the impacts of biofuels on the European agriculture, we have chosen to represent the biofuel production chain, using *AROPAj* as initial link in providing the supply of raw materials.

To implement this approach, our choice was focused on the establishment of a linear mathematical programming model for the entire European biofuels industry. A linear programming model allows us to determine the optimal solution to a problem according to the constraints scripted for a given period and it functions as a decision aid for problems of a large scale. This method actually allows us to use a very large number of variables and equations. Developing a global model is particularly complex and requires, given the variety of stakeholders in the sector, a relatively simplistic representation. To preserve the delicacy of modeling various sectors involved in this industry, we opted for a weak coupling of several sectoral models. Thus, in contrast to a strong coupling, we maintain the independence of the models, which grants them the freedom to evolve.

The methodology here implemented entails coupling three models, each representing one of the key sectors of the biofuels industry. The model of European agricultural supply, *AROPAj*, simulates the domestic supply of raw materials. All or part of its production, as well as the maximum imports of raw materials, is then used by the partial equilibrium model *GIRAF*. This model simulates the European industrial production of different types of biofuels, depending on the supply of domestic and international raw materials, as well as the demand for biofuels. The European refining model *OURSEurope* simulates the production of European refinery products and therefore the local demand for biofuels to meet the objectives of incorporation.

The *AROPAj* model was modified in order to introduce energy crops and crop by-products. The *GIRAF* model was created for the coupling because no existing model met the desired requirements. As for the model *OURSEurope*, it is a European version of the global refining *OURSE* model (developed by the *Ifpen*), and it was created specifically for coupling. In addition, the geographic dimension has been implemented in the three models to spatialize all represented agents (1307 farm-type groups of the *V3-AROPAj* version, 288 biofuel producers and 109 refiners) as well as to highlight the exchanges between them. This allows, in particular, the identification of the European bio-fuel production and consumption areas.

The approach adopted is a partial equilibrium modeling type, so in theory the prices of products considered (oil products, biofuels, etc.) do not affect other products of the economy. Taking into account the spatiality, combined with horizontal dimensions (products, actors) and vertical ones (from upstream to downstream activities), gives a certain advantages to this methodology. The modeling is dedicated to the European zone and is highly parameterized in order to adapt to different types of scenarios, both technical and environmental, as well as economic or political. This allows to test different hypotheses and to analyze their impacts on the production of biofuels (quantity, nature, environmental impacts, co-production and location), associated costs (investment, transportation, raw materials) and the exchanges between agents of the industry. Both conventional (pure vegetable oil, biodiesel, *NexBtL* and first generation bioethanol) and uncon-

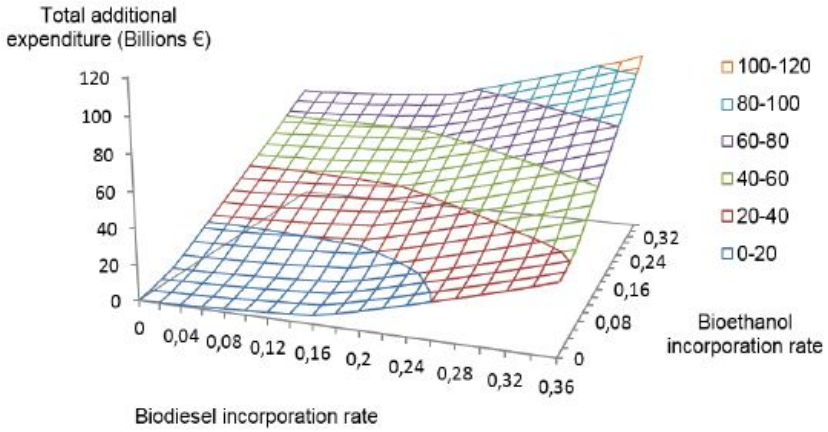


Figure 33. Applying the *AGRAF* model to evaluate impacts of blending rates of biofuel in gazoline and gasoil consumed in Europe.
Source: **?**

ventional (*BtL*, second generation bioethanol) biofuel production are modeled. For each process, technical, environmental and economic data can be changed. However, due to a lack of acces to reliable data on biofuels produced from seaweeds, we have chosen not to represent this production. The modeling is operational and requires that each step be accurately monitored to ensure the proper integration of the parameters in the simulation (**?**).

As illustrative results, Figure **??** shows how incorporating bio-fuels in the European fuel demand may impact the overall cost for the raw biomass transformers and the refiners.

10 Model evolution

10.1 The model over time

Key-points:

Relating published papers with different *AROPAj* versions and *FADN* databases

From *CAP* oriented analysis toward environmental issues

From private company support toward European Universities and institutional public support

Competing for *EC* FrameWork program and tender support (*EUROTOOLS, GENEDEC, INSEA*, specific milk market reform, *CCTAME, AnimalChange, FoodSecure*)

The articles and reports published over the years in relation to the *AROPAj* model refer to different versions of the model. Differences stem mainly from the *FADN* data and from the state of the *EU* enlargement. In addition, series of new blocks and shell were developed regularly. This is mainly due to the change in the area of focus: the model building started with *CAP* analysis and it evolved toward the study of relationship between agriculture and the environment. Nevertheless, we keep the principles of the model intact.

One of the major driving forces of the model is its modular and generic conception. Another driver lies in the evolution of computing capacities and facilities. Thus, we have anticipated the need to overcome computation limits and to enlarge modeling capacities when new scientific problems arise.

The foundations of the model are derived from an early 1990s assessement of the *CAP* reform impacts in France and the UK, which was comissioned by an international grain trader. A major inflexion occurred in the mid 1990s when climate change impacts and *GHG* emissions emerged in the field of economic analysis. Two drivers significantly impact the model evolution: (i) European contacts acting as precursors, through researchers and MSc or PhD students from Universities of Bologna and Parma, and Universidad Politécnica de Madrid; (ii) French programs under the control of environmental agencies and ministries. Successful bids (partnership or coordination) in response to *EC* research calls (*EUROTOOLS*, *GENEDEC*, *INSEA*, *CC-TAME*, *AnimalChange*) strongly boosted the development of the model, in addition to more specific responses to calls for tenders and national incentives. High development and the support of multidisciplinary approaches facilitated the trend from analysis focusing on the *CAP* and economic criteria, to an analysis focusing on the environmental impacts and external effects regulation.

Appendices **??** and **??** provide the list of contributors over time and the list or projects which contributed to funding for *AROPAj*.

10.2 Version list

Key-points:

From the *EU*-15 to the *EU*-27

FADN sourced information, from 1997 untill 2009

Toward a unified version of the model kernel

The European coverage was achieved by the model with the *EU*-12 *EUROTOOLS* program. The *GENEDEC* program allowed us to extend the coverage to the *EU*-15 and the *EU*-25 (except Malta, when the *FADN* was still provisional). The *EU*-27 coverage was realized at the end of the *AnimalChange* program (see Table **??**). When we intend to extend the time coverage (till 2011, as the last *FADN* year available in accordance with the *FoodSecure* program), change in *EC FADN* farm types and individual farm sampling requires to renew all *AROPAj* building steps.

We were forced to deal with delays in accessing *FADN* data as well as delays in validation of *FADN* data by the *EC* statistical services. It must be mentioned how fortunate we are to access the *FADN*, since its counterpart is an individual project that treat data only by providing anonymous parameter estimates.

A version of the model essentially refers to a specific annual *FADN* survey. Such a version is characterized by the typology (i.e. farm groups) and estimated parameter files (`mgg.par` described hereafter in section **??**). At present, there are no, properly speaking, different versions related to the kernel of the model, in the sense that operating routine improvement does not impede running of previous *FADN* versions and calculations.

	*V*1 2001	*V*2 2005	*V*3 2007	*V*4-test 2012	*V*5 2014	*V*5 2017	*V*5 2020
Typology	1997	2002	2004	2009	200(7 8 9)	20(07 - 12)	20(13 - 17)
Irrigation					2014	2017	2020
Estimate	1997	2002	2004	2009	200(7 8 9)	20(07 - 12)	20(13 - 17)
Entropy					2014	2017	2020
Kernel	*Unix*	*Unix*	*Unix*	*Unix*/*Linux*	*Linux*	*Linux*	*Linux*
GHG em.	*IPCC*-1996	*IPCC*-1996	*IPCC*-1996		*IPCC*-2010	*IPCC*-2010	*IPCC*-2016
N-yield		2011			- 2015	- 2018	-
Water-*N*-yield							
Spatial.		2013			2015	2018	-
	EU-12	*EU*-15	*EU*-25	*EU*-27	*EU*-27	*EU*-27	*EU*-28

Table 32. The versions of the *AROPAj* model linked to the added blocks, process and applications (maintained/updated and used, 2020).

Due to the continual growth of the model core and the development of applications, there is significant evolution in the production of results. Model versions differ according to:

- the geographical coverage within the *EU*, referring to the *FADN* year
- the method for farming system clustering (e.g. through the classification method, the used criteria and the selected metric)
- the method for parameter estimation (including or not the *GME* method)
- the activation of added endogenous variables (e.g. yields and *N*-fertilizers when using *N* to yield functions)
- the economic and environmental outputs (e.g. *N*-losses functions associated to *N* to yield functions)
- the accuracy of results (e.g. in line with the *GHG* emission block)

- the geographical downscaling of results (i.e. through the "spatialization process")

Table **??** displays how versions differ according to these items. Major parts of the model construction evolve over time. The structural changes of the model may affect results in terms of type and accuracy. Typology was improved when irrigation was taken into account. Parameter estimation is dramatically changed when the *GME* method is used. The *N* to yield functions allow us to account for the intensive margin effect at the plot level. In addition to these "internal" improvements, the spatialization process allows for new model applications and potential linkages with other models. Finally, let us recall that the evolution of computer capabilities strongly impacts the development of the different stages of model building and applications.

10.3 Adding blocks, programs and tools

> Key-points:
>
> Implementing the calibration process
> Developments and modules related to some versions, ready to be expanded
> About *GHG* emissions, spatialization, dose-response functions, model coupling

The use of additional blocks, programs and applications is in large part autonomous from the model kernel. There is always the possibility of using any version of the model. At the same time activation of some of these additional blocks or applications may not be allowed under some model versions. The *LP* calibration process is a quite specific but emblematic case of external programs linked to the kernel.

***LP* calibration**

Calibration refers to one of the most important steps of the *AROPAj* building, in terms of complexity as well as computation time. The shell dedicated to this step acts as a super routine calling the *AROPAj* kernel as a subroutine. This shell encapsulates the kernel for managing iterated *AROPAj* runs. As explained in Section **??**, the calibration algorithm determines how the *AROPAj* parameters have to change from one run to the next. In this case, the shell is composed of external code components (e.g. *FORTRAN* code, mainly). Additional information, requested for the calculation of the calibration criterion, enters the process through *FADN* sourced parameters. The process runs at the *FG* level and provides partly re-estimated parameters.

***GHG* emission assessment**

Our standard process of estimating *GHG* emissions is based on *IPCC* guidelines as well as national emission factors that impact emission inventories. We need to compute multiple series of national (possibly regional) parameters, separate from the core parameter set and external to the *AROPAj* estimation and calibration steps. This process entails adding a block to the kernel. Related parameter values are then formatted in line with *AROPAj* basic parameters.

Another possible way of assessing N_2O emission types is to activate the *N*-block. This require that *N* to yield functions are available and that *N*-losses are regressed against *N*-inputs (see below).

***N* to yield functions**

This additional block requires the availability of calibrated *N* to yield functions. When such a function is available for a farm group and for a crop, *FADN* sourced yield and input values (at

the plot level) are replaced by computed values. In such case, the computed yield and N-input depend on the price ratio. By default, the model calls for market prices. There is need to change the kernel.

When N to yield and N to N-loss functions exist, there exists the possibility of assessing different N-losses at the farm and wider levels, namely N_2O, NO_3 and NH_3 (what the *STICS* crop model allows for estimation). It must be noted here that, regarding N_2O, results may vary drastically from those provided by the *IPCC* method.

Organic N and LT effect

Accounting for LT effect assessment of organic N spreading in *AROPAj* requires the availability of N to yield functions. This kind of effect affects the N amount requested by a crop to reach a given yield. The magnitude of the LT effect depends on the average quantity of organic N spreading at the farm level. The computation requires the use of an external routine and the iteration of *AROPAj* runs leading to the convergence of *AROPAj* outputs. This process concerns two kinds of parameters, some of them provided by the *Carbo-PRO* model, and others re-estimated according to iterative *AROPAj* runs. The *LP* case study shows fast convergence of the iterative process.

The replacement of market prices by shadow prices

This process may be implemented when activating the nitrogen block in line with N to yield functions. Mathematically, shadow prices linked to the use of crop production should replace market prices for maximizing the gross margin at the plot scale in some cases. This occurs when, at the optimum, the crop production is entirely re-used on-farm (e.g. for livestock feeding), or when no N-input comes from the market (e.g. manure sourced). In practice, the problem is solved by iterated *AROPAj* runs with shadow prices replacing the prices used at the previous iteration for the yield function derivative. The process quickly converge in most of *AROPAj LP*s, except for in a few cases where solutions have periodic forms over iterations, with slight amplitudes.

Spatialisation of farm groups and downscaling of *AROPAj* outputs

There exist the possibility of *FG* locating, in terms of probabilistic distributed contribution of the agricultural activities over a geographical grid. This is obtained through the spatialization module (see Section **??**) available for the *V2* version, and is in progress for the *V5* version. Once contributions of all *FG*s within a region are computed, there is no difficulty in displaying any *AROPAj* output related to this region at a fine resolution level. The size of the database and the overlapping data used to pose problems for our computing capacities, specifically at the first step of the spatialization process (e.g. organize data for spatial econometrics). Currently, most of the spatialization steps are performed on our workstation thanks to open-source *GNU* softwares, except for the *GCE* procedure which calls for *GAMS*.

The combination of these blocks and tools is displayed on Table **??**.

10.4 Comparison tables and *AROPAj* sourced biases

Key-points:

Farm group representativity and *FADN* based allocation of agricultural lands over the years and the *AROPAj* versions

Calibration results and land allocation bias

	Database	clustering / sample farms	estimation & calibration	*LP* design	*FG* spatialisation
base	*FADN*	farm groups (*FG*)	basic parameters	kernel	
FG's *LP* calibration			re-estimated parameters	external routine	
GHG emissions	*IPCC* & nat. inventories	regionally distributed	added parameters	GHG block	
N to yield functions	*STICS* - relat. requirement	added parameters related to *FG* & *FADN* sourced data		internal subroutine	
N to *N*-loss functions		parameters related to *N* to yield function calibration			
LT effect of organic *N*		added & re-evaluated parameters / *N* to yield funct.		external routine	
shadow prices convergence			re-evaluated parameters	external routine	
outputs downscaling	*CLC* & other physical data	probabilistic *FG* location			external program

Table 33. Impacted *AROPAj* steps and files when adding blocks or managing applications.

From the real world up to the stylized calibrated representation of farming systems, there are successive gaps leading to an aggregation bias. Upstream from the *AROPAj* filter, we have to account for:

- the European surveys (with possible truncation, regarding very large farms as well as small systems managed by part-time farmers),
- their statistical treatment by the *EC* services (with possible difficulties due to heterogeneous quality of national surveys),
- the clustering of sample farms leading to *FADN* categories (*FT*, economic size, altitude), which strongly impacts the *AROPAj FG* building, even in terms of excluded farming systems (e.g. ones for which *FT* refers to wine or horticulture).

Obviously, as for any model, the *AROPAj* model results do not perfectly fit with observations (here viewed through the *FADN*). The *LP* design, the *FG* typology and the estimation and calibration phases all contributed to the distortion of the raw *FADN* estimates of production and allocated area among crops. In addition, comparison is made more difficult when our *LP* objective does not conrrespond exactly to a *FADN* variable.

In order to illustrate the quality of the calibration phase, we present histograms that show the distribution of *FG*s' square root of sum of squared deviations between *FADN* based informations and computed terms (see Figures **??**, **??**, **??**, **??**, **??**, **??**). Regarding the possible *FADN* sourced bias, the figures displayed in Table **??** show significant sources of bias. Bias may occur because of the *AROPAj* typology (see the *UAA* deviation between the *V4* and *V5*s versions for Estonia or Romania, among others *EC* countries). Bias also occurs because of the *FADN* itself, as is envident when comparing the national *UAA* among the years 2007, 2008 and 2009, all of which were estimated through a common algorithm of *FADN*-farms clustering (e.g. for Italy and Poland, between 2007 and 2008 or 2009).

Stated as a quasi-fix factor, the *UAA* is not optimally re-computed at any level in *AROPAj* (from the *FG* to the *EU*). However, calibration impacts the land allocation within the *UAA*, as illustrated by Tables referring to often used *V2* and *V3* versions, and to the 2007 to 2009 *V5* ones.

Table 34. Comparison of land allocation between *FADN* estimates and *AROPAj* calibrated outputs, aggregated at the member State level and for different *AROPAj* versions (millions hectares) - (*AROPAj* runs were realized on March 2015).

MS	Land use	V2		V3		V5-2007		V5-2008		V5-2009	
		FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.
bel	cereals	0,306	0,332	0,303	0,321	0,333	0,313	0,379	0,382	0,381	0,379
bel	oilseeds & proteins	0,006	0,003	0,005	0,006	0,041	0,041	0,044	0,041	0,047	0,054
bel	roots & indus.	0,156	0,161	0,171	0,190	0,161	0,155	0,141	0,146	0,153	0,151
bel	pastures	0,689	0,560	0,574	0,420	0,611	0,648	0,615	0,621	0,626	0,671
bel	fodders	0,254	0,255	0,271	0,291	0,175	0,171	0,196	0,182	0,202	0,151
bel	fallows & set-aside	0,021	0,121	0,015	0,111	0,012	0,004	0,005	0,007	0,002	0,005
bgr	cereals					1,637	1,720	1,548	1,712	1,733	1,820
bgr	oilseeds & proteins					0,660	0,670	0,702	0,680	0,756	0,743
bgr	roots & indus.					0,016	0,039	0,015	0,043	0,027	0,054
bgr	pastures					0,344	0,223	0,390	0,202	0,309	0,247
bgr	fodders					0,065	0,020	0,051	0,011	0,055	0,012
bgr	fallows & set-aside					0,029	0,080	0,075	0,132	0,105	0,109
cyp	cereals			0,087	0,072	0,042	0,054	0,083	0,004	0,060	0,055
cyp	oilseeds & proteins			0,000	0,000	0,003	0,010	0,006	0,008	0,006	0,010
cyp	roots & indus.			0,004	0,011	0,004	0,008	0,004	0,012	0,003	0,008
cyp	pastures			0,026	0,027	0,010	0,001	0,004	0,000	0,005	0,021
cyp	fodders			0,014	0,017	0,038	0,001	0,022	0,001	0,031	0,001
cyp	fallows & set-aside			0,006	0,009	0,009	0,034	0,006	0,100	0,007	0,017
cze	cereals			1,664	1,795	1,618	1,795	1,510	1,693	1,511	1,633
cze	oilseeds & proteins			0,319	0,357	0,392	0,434	0,377	0,431	0,422	0,425
cze	roots & indus.			0,141	0,204	0,103	0,076	0,088	0,070	0,101	0,058
cze	pastures			1,037	0,693	0,782	0,665	0,815	0,732	0,823	0,785
cze	fodders			0,249	0,229	0,376	0,205	0,358	0,167	0,354	0,277
cze	fallows & set-aside			0,020	0,153	0,006	0,101	0,005	0,060	0,012	0,044
dan	cereals	1,530	1,416	1,485	1,591	1,455	1,879	1,518	1,633	1,553	1,705
dan	oilseeds & proteins	0,081	0,039	0,133	0,139	0,195	0,286	0,189	0,226	0,174	0,224
dan	roots & indus.	0,106	0,103	0,094	0,107	0,088	0,122	0,086	0,087	0,078	0,078
dan	pastures	0,187	0,246	0,164	0,172	0,389	0,018	0,424	0,400	0,500	0,325
dan	fodders	0,320	0,257	0,335	0,177	0,198	0,140	0,214	0,098	0,216	0,217
dan	fallows & set-aside	0,212	0,376	0,185	0,210	0,160	0,042	0,089	0,077	0,068	0,041
deu	cereals	6,461	5,667	6,717	7,523	6,488	7,191	6,823	7,511	6,683	7,218
deu	oilseeds & proteins	1,011	0,838	1,189	1,178	1,511	1,807	1,627	1,841	1,735	1,867
deu	roots & indus.	0,681	0,902	0,724	0,905	0,679	0,671	0,632	0,507	0,648	0,731
deu	pastures	4,452	3,560	4,568	2,864	3,861	4,425	3,873	4,376	3,931	4,484
deu	fodders	1,278	0,698	1,387	0,590	1,755	0,600	2,021	0,896	2,157	0,935
deu	fallows & set-aside	0,845	3,063	0,829	2,354	0,636	0,238	0,328	0,175	0,257	0,175
ell	cereals	0,876	0,824	0,966	0,991	1,383	1,544	1,499	1,631	1,470	1,593
ell	oilseeds & proteins	0,022	0,019	0,012	0,021	0,084	0,115	0,085	0,116	0,105	0,144
ell	roots & indus.	0,053	0,044	0,331	0,317	0,381	0,368	0,383	0,370	0,345	0,336
ell	pastures	0,156	0,138	0,181	0,071	0,058	0,066	0,059	0,078	0,059	0,103
ell	fodders	0,061	0,159	0,078	0,185	0,245	0,147	0,263	0,171	0,261	0,166
ell	fallows & set-aside	0,052	0,035	0,083	0,066	0,205	0,116	0,155	0,077	0,183	0,081
esp	cereals	5,769	5,948	5,942	6,999	6,917	8,162	7,032	8,073	6,521	7,863
esp	oilseeds & proteins	0,628	0,292	0,582	0,664	0,996	1,083	0,900	1,046	0,956	0,825
esp	roots & indus.	0,193	0,454	0,299	0,376	0,269	0,208	0,244	0,224	0,242	0,273
esp	pastures	2,463	1,825	2,228	1,293	2,718	2,193	2,372	2,157	2,815	2,437
esp	fodders	0,434	1,933	0,373	1,180	0,657	0,981	0,754	0,822	0,941	1,051
esp	fallows & set-aside	1,842	0,878	1,909	0,821	2,009	0,939	1,706	0,686	1,778	0,805
est	cereals			0,232	0,419	0,301	0,381	0,325	0,390	0,336	0,369
est	oilseeds & proteins			0,039	0,046	0,091	0,105	0,081	0,080	0,085	0,096
est	roots & indus.			0,006	0,053	0,005	0,038	0,004	0,035	0,004	0,026
est	pastures			0,163	0,009	0,410	0,292	0,435	0,331	0,424	0,382
est	fodders			0,189	0,091	0,010	0,007	0,008	0,004	0,013	0,006
est	fallows & set-aside			0,040	0,051	0,020	0,014	0,016	0,027	0,024	0,006
fra	cereals	9,251	10,956	9,283	10,916	9,120	9,987	9,650	10,249	9,486	10,155
fra	oilseeds & proteins	1,770	1,925	1,902	1,788	2,704	2,789	2,543	2,767	2,679	2,772
fra	roots & indus.	0,571	0,507	0,585	0,910	0,585	0,504	0,559	0,502	0,578	0,456
fra	pastures	5,964	4,986	6,092	4,696	9,975	8,662	10,087	8,812	10,176	8,949
fra	fodders	5,318	2,735	5,431	3,870	1,529	1,195	1,593	1,298	1,610	1,420

continued on next page –

– continued from previous page

MS	Land use	V2		V3		V5-2007		V5-2008		V5-2009		
		FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.	
fra	fallows & set-aside	1,167	2,933	1,028	2,141	1,152	1,928	0,776	1,580	0,589	1,366	
gbr	cereals	3,689	3,660	3,431	3,765	2,991	3,527	3,296	3,423	3,227	3,412	
gbr	oilseeds & proteins	0,439	0,491	0,567	0,664	0,856	1,084	0,764	0,878	0,840	0,938	
gbr	roots & indus.	0,395	0,302	0,349	0,385	0,254	0,225	0,266	0,366	0,258	0,332	
gbr	pastures	4,852	4,922	4,078	3,471	5,437	5,052	5,476	5,334	5,691	5,554	
gbr	fodders	1,663	1,098	1,364	1,067	0,225	0,166	0,240	0,135	0,235	0,151	
gbr	fallows & set-aside	0,609	1,175	0,547	0,984	0,563	0,272	0,281	0,187	0,354	0,219	
hun	cereals			2,309	2,683	2,360	2,762	2,355	2,708	2,356	2,662	
hun	oilseeds & proteins			0,495	0,393	0,752	0,691	0,797	0,757	0,790	0,702	
hun	roots & indus.			0,075	0,074	0,055	0,042	0,018	0,033	0,024	0,037	
hun	pastures			0,393	0,082	0,218	0,094	0,227	0,097	0,234	0,153	
hun	fodders			0,151	0,153	0,293	0,090	0,296	0,095	0,295	0,140	
hun	fallows & set-aside			0,053	0,091	0,047	0,044	0,033	0,036	0,040	0,043	
irl	cereals	0,304	0,312	0,350	0,411	0,259	0,301	0,298	0,297	0,306	0,322	
irl	oilseeds & proteins	0,001	0,001	0,001	0,002	0,008	0,017	0,006	0,013	0,007	0,008	
irl	roots & indus.	0,039	0,036	0,047	0,070	0,014	0,016	0,010	0,015	0,006	0,007	
irl	pastures	2,531	2,272	2,380	2,049	3,634	3,359	3,696	3,484	3,807	3,577	
irl	fodders	0,289	0,330	0,269	0,395	0,047	0,028	0,050	0,014	0,059	0,033	
irl	fallows & set-aside	0,035	0,247	0,041	0,161	0,031	0,271	0,011	0,249	0,011	0,249	
ita	cereals	3,417	4,163	3,646	4,358	3,658	4,515	3,618	4,625	3,403	4,822	
ita	oilseeds & proteins	0,264	0,260	0,222	0,186	0,570	0,675	0,511	0,529	0,590	0,620	
ita	roots & indus.	0,244	0,161	0,239	0,290	0,178	0,155	0,157	0,170	0,146	0,173	
ita	pastures	2,891	1,243	0,434	0,384	4,115	3,197	0,060	0,260	0,122	0,548	
ita	fodders	0,488	0,777	3,338	2,303	0,233	0,247	3,217	1,288	3,152	0,911	
ita	fallows & set-aside	0,248	0,947	0,247	0,578	0,355	0,320	0,285	0,977	0,571	0,907	
ltu	cereals			0,687	0,833	0,905	1,032	0,931	1,021	0,889	0,949	
ltu	oilseeds & proteins			0,097	0,119	0,199	0,211	0,184	0,186	0,197	0,214	
ltu	roots & indus.			0,053	0,118	0,037	0,059	0,029	0,049	0,030	0,047	
ltu	pastures			0,063	0,065	0,642	0,572	0,635	0,534	0,641	0,617	
ltu	fodders			0,394	0,306	0,030	0,008	0,021	0,015	0,021	0,005	
ltu	fallows & set-aside			0,177	0,028	0,094	0,025	0,090	0,085	0,090	0,037	
lux	cereals	0,030	0,034	0,027	0,021	0,029	0,023	0,030	0,025	0,030	0,023	
lux	oilseeds & proteins	0,003	0,005	0,004	0,005	0,005	0,005	0,005	0,005	0,005	0,005	
lux	roots & indus.	0,001	0,004	0,001	0,002	0,001	0,001	0,001	0,002	0,001	0,001	
lux	pastures	0,081	0,064	0,077	0,067	0,063	0,073	0,062	0,076	0,062	0,083	
lux	fodders	0,013	0,008	0,014	0,021	0,027	0,019	0,027	0,015	0,029	0,010	
lux	fallows & set-aside	0,002	0,014	0,001	0,008	0,001	0,004	0,000	0,003	0,000	0,004	
lva	cereals			0,409	0,599	0,485	0,592	0,495	0,614	0,492	0,555	
lva	oilseeds & proteins			0,066	0,061	0,103	0,121	0,087	0,108	0,093	0,111	
lva	roots & indus.			0,034	0,080	0,021	0,054	0,017	0,038	0,017	0,051	
lva	pastures			0,133	0,096	0,405	0,346	0,426	0,299	0,439	0,420	
lva	fodders			0,175	0,147	0,018	0,023	0,015	0,029	0,014	0,029	
lva	fallows & set-aside			0,182	0,017	0,109	0,006	0,116	0,069	0,122	0,012	
mlt	cereals					0,000	0,000	0,000	0,000	0,000	0,000	
mlt	oilseeds & proteins					0,000	0,000	0,000	0,000	0,000	0,000	
mlt	roots & indus.					0,000	0,000	0,000	0,000	0,000	0,000	
mlt	pastures					0,000	0,000	0,000	0,000	0,000	0,000	
mlt	fodders					0,001	0,000	0,002	0,000	0,001	0,000	
mlt	fallows & set-aside					0,000	0,001	0,000	0,001	0,000	0,001	
ned	cereals	0,209	0,217	0,168	0,187	0,195	0,211	0,204	0,279	0,203	0,183	
ned	oilseeds & proteins	0,002	0,011	0,003	0,006	0,074	0,072	0,068	0,077	0,069	0,083	
ned	roots & indus.	0,335	0,258	0,234	0,195	0,224	0,264	0,219	0,205	0,223	0,213	
ned	pastures	0,676	0,756	0,588	0,661	0,978	0,284	0,982	0,904	0,963	1,194	
ned	fodders	0,240	0,122	0,209	0,114	0,229	0,024	0,245	0,003	0,252	0,004	
ned	fallows & set-aside	0,049	0,147	0,031	0,072	0,017	0,860	0,011	0,260	0,008	0,041	
ost	cereals	0,675	0,743	0,732	0,788	0,705	0,835	0,736	0,832	0,737	0,802	
ost	oilseeds & proteins	0,086	0,041	0,082	0,084	0,123	0,090	0,122	0,109	0,119	0,095	
ost	roots & indus.	0,069	0,120	0,065	0,090	0,063	0,083	0,064	0,094	0,064	0,070	
ost	pastures	0,797	0,665	0,811	0,702	0,619	0,571	0,634	0,610	0,636	0,662	
ost	fodders	0,074	0,043	0,084	0,081	0,213	0,161	0,213	0,130	0,200	0,106	
ost	fallows & set-aside	0,086	0,176	0,080	0,107	0,056	0,040	0,038	0,033	0,035	0,056	
pol	cereals	0,000	0,000	4,830	5,267	7,361	7,815	7,974	8,476	7,934	8,426	
pol	oilseeds & proteins	0,000	0,000	0,555	0,561	1,313	1,448	1,278	1,420	1,316	1,487	
pol	roots & indus.	0,000	0,000	0,591	0,753	0,649	0,882	0,577	0,597	0,592	0,629	
pol	pastures	0,000	0,000	0,248	0,055	2,590	2,201	2,663	2,281	2,752	2,298	
pol	fodders	0,000	0,000	1,474	1,043	0,556	0,126	0,638	0,369	0,652	0,375	
pol	fallows & set-aside	0,000	0,000	0,008	0,026	0,125	0,122	0,124	0,111	0,127	0,159	
por	cereals	0,390	0,577	0,312	0,502	0,212	0,566	0,279	0,564	0,247	0,565	
por	oilseeds & proteins	0,021	0,020	0,016	0,014	0,042	0,048	0,043	0,051	0,044	0,072	
por	roots & indus.	0,033	0,201	0,024	0,054	0,020	0,016	0,019	0,019	0,021	0,020	

continued on next page –

– continued from previous page

MS	Land use	V2		V3		V5-2007		V5-2008		V5-2009		
		FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.	FADN	calibr.	
por	pastures		0,501	0,497	0,424	0,250	0,693	0,665	0,842	0,892	0,916	0,589
por	fodders		0,211	0,222	0,238	0,197	0,349	0,275	0,396	0,247	0,317	0,248
por	fallows & set-aside		0,833	0,473	0,703	0,374	0,715	0,462	0,689	0,493	0,579	0,630
rou	cereals						4,538	5,386	5,947	7,180	5,997	7,160
rou	oilseeds & proteins						1,348	1,020	1,469	1,528	1,532	1,586
rou	roots & indus.						0,133	0,345	0,180	0,499	0,145	0,356
rou	pastures						1,278	0,855	1,365	0,553	1,768	0,712
rou	fodders						0,539	0,040	0,722	0,112	0,655	0,177
rou	fallows & set-aside						0,080	0,166	0,417	0,227	0,086	0,190
suo	cereals		0,953	1,070	1,043	1,193	1,031	1,177	1,116	1,201	1,092	1,200
suo	oilseeds & proteins		0,067	0,002	0,086	0,086	0,115	0,127	0,096	0,114	0,119	0,147
suo	roots & indus.		0,067	0,127	0,062	0,117	0,046	0,053	0,049	0,062	0,056	0,067
suo	pastures		0,007	0,010	0,002	0,001	0,629	0,614	0,646	0,660	0,661	0,618
suo	fodders		0,611	0,655	0,600	0,555	0,016	0,005	0,014	0,003	0,017	0,000
suo	fallows & set-aside		0,183	0,025	0,185	0,026	0,188	0,048	0,163	0,043	0,165	0,078
sve	cereals		1,057	1,022	1,105	1,242	0,979	1,023	1,061	1,197	1,038	0,930
sve	oilseeds & proteins		0,055	0,054	0,089	0,087	0,153	0,175	0,144	0,178	0,143	0,122
sve	roots & indus.		0,090	0,158	0,090	0,086	0,069	0,108	0,069	0,047	0,067	0,092
sve	pastures		0,386	0,726	0,328	0,442	1,072	1,131	1,067	0,966	1,118	1,215
sve	fodders		0,829	0,325	0,721	0,378	0,030	0,012	0,033	0,034	0,042	0,034
sve	fallows & set-aside		0,223	0,353	0,227	0,325	0,179	0,034	0,097	0,050	0,088	0,104
svk	cereals				0,707	1,038	0,774	0,955	0,769	0,874	0,748	0,857
svk	oilseeds & proteins				0,178	0,239	0,280	0,392	0,286	0,383	0,298	0,361
svk	roots & indus.				0,041	0,086	0,026	0,032	0,021	0,026	0,027	0,013
svk	pastures				0,597	0,205	0,588	0,352	0,575	0,441	0,540	0,475
svk	fodders				0,068	0,116	0,236	0,119	0,223	0,123	0,229	0,120
svk	fallows & set-aside				0,170	0,077	0,021	0,076	0,025	0,051	0,012	0,027
svn	cereals				0,060	0,096	0,077	0,096	0,085	0,097	0,085	0,091
svn	oilseeds & proteins				0,001	0,000	0,008	0,007	0,006	0,006	0,006	0,003
svn	roots & indus.				0,009	0,012	0,004	0,002	0,004	0,005	0,003	0,011
svn	pastures				0,311	0,253	0,312	0,290	0,309	0,274	0,316	0,286
svn	fodders				0,049	0,032	0,035	0,035	0,030	0,030	0,028	0,027
svn	fallows & set-aside				0,001	0,038	0,001	0,007	0,000	0,022	0,000	0,019

Former versions exhibit some significant deviations between *FADN* estimates and calibrated results, with some countries showing larger discrepancies than others. This could be due to difficulties in *FADN* land use identification (with likely misinterpreted land use codes on the *FADN* data warehouse). On the whole, cereals and set-aside are overestimated and pastures are underestimated, even though fodders partly compensate the underestimated pastures in several countries. The fifth version is able to minimize the discrepancy between *FADN* estimates and calibrated results due to progressions in the typology, cost estimates, and the calibration process itself. The calibration process could be even further improved with the ability to achieve even more iterated runs by our computing system.

11 Computing tools

11.1 Hardware and software

> Key-points:
>
> From *Unix* to *Linux*
>
> Preferably *GNU GPL* applications, the *FORTRAN* compiler, and the *GAMS* software
>
> From the *MGG* kernel to enlarged *MGG* code transformed into *GAMS* code
>
> *AROPAj* file architecture

11.1.1 Operating system, software and runtime

Until 2013, all computing tool implementation were implemented within the *Unix* operating system. Now the *AROPAj* model core and applications are used via the *Linux OS*. These softwares are quite convenient for long term secured applications launched by batch programs.

Over the time, the model has been developed and implemented on multi-core workstations. Protection and security tools are managed by the *INRA* "Direction des Services Informatiques". Thus, databases, including the *FADN*, are protected and secure. This final point is of high interest, the ability to manage datasets provided by public institutions is of paramount importance.

The running workstation is of a 16-core class. In observing the progression from the first working tool to the 2014 working tool, we see a computer core scale example of "Moore's Law." In other words, what was once a binding time limit became much less binding (e.g. for implementation of the calibrating process). Multi-core systems multiply the gain in runtime, so what required several weeks of runtime 10 years ago now only needs a couple of hours.

11.1.2 Software and runtime

Year after year, we have chosen to use *GNU* softwares in the improvement of the *AROAPj* model, not only for their relatively low financial costs, but also as a means of facilitating the maintenance and evolution of our programs and applications.

The starting command and the processes calling for major applications are written in `csh` *CShell* language. Intermediate commands and applications are written in *FORTRAN* or *CShell* languages.

The core of the *AROPAj* model is written in *FORTRAN*, even if the basic file `aro.mgg` remains coded in the former *MGG* code, close to *FORTRAN*. It is currently written and used regardless of the former *MGG* software. Programs that transform the core files and deliver the *GAMS* code for *MP* solving combine *CShell* and *FORTRAN* processing tools.

Basic statistical calculations are now realized through *R* code programs. The *R* language is also used for spatialization modules, which deliver maps. The *POSTGreSQL* system is used for data processing and management, especially in the case of *STICS* based applications. Major batch interfaces between databases and the *STICS* crop model are made of *R* code programs. In addition, some processing tools are now written in *Python* code.

The *GAMS* software is used for cumbersome optimization calculations. That applies for *AROPAj* simulations as well as for *GME* and *GCE* estimating programs. The *GAMS* code is

generated for *AROPAj* from the kernel files described hereafter (Section **??**). This is automatically done when the user calls the model for simulation runs.

Some specific applications may call for other languages and softwares. For example, we used *Mathematica* for time based scenario development. This occurs when applying the Faustmann rule to the calculation of the optimal perennial crop rotation over years. These softwares are used upstream from *AROPAj* simulations, in an external way, without direct coupling to *AROPAj*. In addition, *AROPAj* outputs may be formatted for treatment by usual office softwares.

At last, we combine some of the previously mentioned *GNU* languages to feed information documents into LaTeX format. This is the case with some of the elements of this paper, such as the list of parameters and *LP* variables.

11.1.3 The *AROPAj* files and shell

The core of the *AROPAj* model is essentially composed of 4 files, as required by the use of *MGG* and *SCICONIC* software. We updated these files and transform them into *GAMS* language via *Unix CShell* and *FORTRAN* programs. This transformation is automatically processed when the user calls the model via the *AROPAj* call shell. These files are:

- **aro.mgg**, which designs and describes the basic linear structure of *AROPAj* by listing and defining activities, constraints and the objective function, as well as the link to the model parameters. The coding language is similar to *FORTRAN*. This file allows us to build the linear programming matrix associated to any parameter set. We keep this method for matrix generation operational;
- **intern.f**, which codes the relationship between parameters (in **mgg.par** format) and the matrix generator **aro.mgg**. This file is made of *FORTRAN* instructions;
- **aro.dat**, which is useful for writing reports and naming primal and dual variables. It also plays a role in flagging variables to activate and deactivate them;
- **mgg.par**, which is made of the comprehensive list of parameters that is read by **intern.f** and which is required by **aro.mgg** to design the matrix. This is a *text* format file made of fixed format lines referring to the type of information (parameter category, e.g. *rjj* for crop yield), the characteristics inside the category (e.g. *bl* for soft wheat), the farm group number when appropriate, and the parameter value. When information exists on several lines of the file, the operating value retained for the matrix is the last one.

The model kernel can be included in a calculation chain designed by the user. To do that, the user has to combine the core with external files with an eye to her/his goals (see the next section). One important illustrative case is the calibrating process (see Section **??**).

The usual way to implement a series of simulations embedded in a *scenario* is to select and design the following files ("scen" is of user's choice):

- **scen.par**, which identifies the parameters of interest from those listed in the **mgg.par** file (the **scen.par** file keeps the **mgg.par** format);
- **scen.var** which focuses on *LP* (primal and dual) variables of interest for the *scenario*. The format is designed to allow the *LP* outputs to become entries of the calculation program;

- **scen.f** which codes (in *FORTRAN*) the *scenario* and manages the writing of the final output (by default close to *csv* format); final outputs may become entry variables of analytic functions of parameters of interest (belonging to the **scen.par** list) and *LP* outputs (belonging to the **scen.var** list). This kind of program may be used to insert *AROPAj* as a subroutine of a wider model, when *AROPAj* parameters are viewed as variables of the global model (e.g. the *AROPAj* calibrating process);
- **scen.gar**, which is a list of 8-character chains that refer to final outputs. These outputs are delivered per farm group for each "country" to which the *scenario* applies. Each of **scen.gar** entry must be linked to a *FORTRAN* instruction in **scen.f**, which designs a formatted output as a part of the output file (**resultIDk**, see below Section **??**);
- **scen.ini**, which, depending of the **scen.f** programming file, allows us to select a series of sub-scenarios from the *scenario* class. A typical example is when a user wants to examine *AROPAj* results while changing one, two or series of parameters within a predefined range of values;
- **scen.don**, which is optional and refers to any external data which is possibly required by the *scenario* (e.g. when nitrogen-yield functions depend on climate change);
- **scen.csh**, which organizes the linkage between the kernel files (and more precisely the *AROPAj* shell *Unix* scripts allowing the user to call *AROPAj*) and the previous *scenario* files. The *AROPAj* shell allows the user to follow the calculation steps in real time, and it provides identifying numbers, automatically based on the computer clock, which relate to the output files. All other files named **scen** may be built and/or copied (by default "cloned" files) and transformed through **scen.csh** by the user.

The user may utilize other files as inputs or outputs, and these files may be built, called and transformed with the support of the **scen.csh** script as well as of the **scen.f** routine. Some processing rules apply when building from **scen.f** and **scen.csh**. These mainly entail format consistency, (possibly forbidden) comment lines, and consistency with predefined *FORTRAN* variables.

Different steps of the *AROPAj* building are performed upstream by the *AROPAj* administrator, namely:

- the *AROPAj* shell conception and the development of codes (*CShell*, *FORTRAN*, *R*), widely enlarged by users' contribution
- the *AROPAj* shell implementation (**miraj** is the starting command launched by the administrator and by users)
- the modeled farm design and the *LP* matrix generating conception
- the secure transformation of *FADN* databases
- the farm group typology and the *FADN* farms clustering
- the parameter estimation step which aims at feeding the *AROPAj* databases (**mgg.par**)
- the re-estimation of parameters (calibration process)
- the building of "*AROPAj* country" parameter databases, i.e. the **mgg.par** files which contain all farm group information referring to the *FADN*-year based scenario
- the conception of simulation and *AROPAj* run management

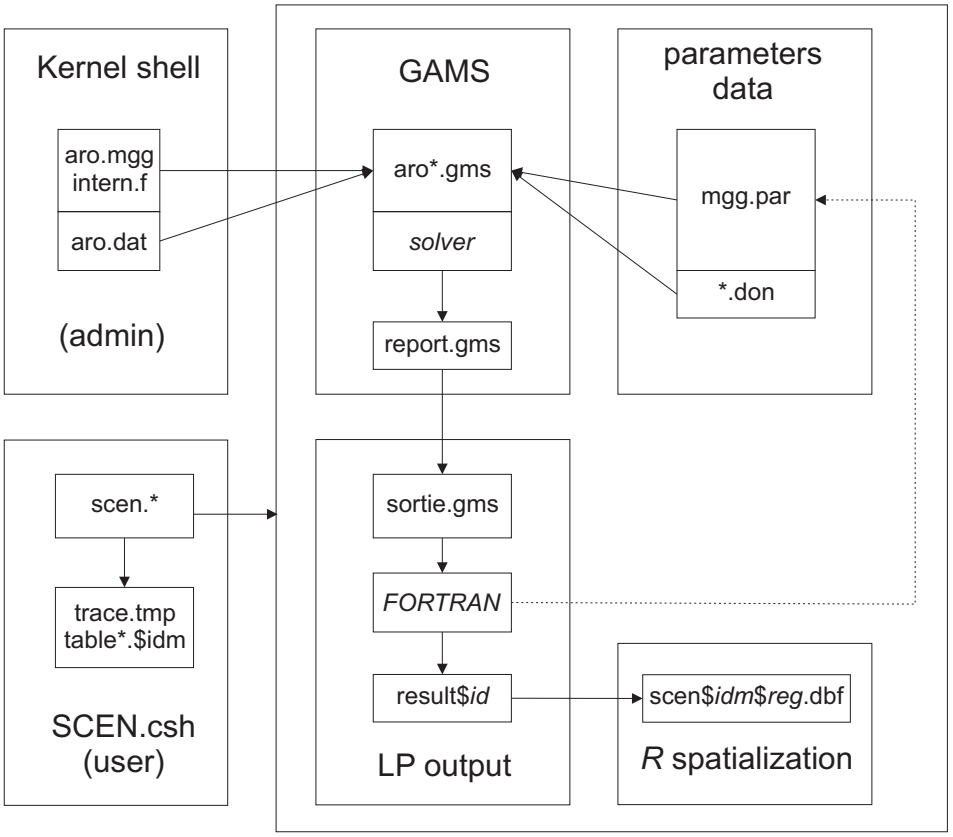


Figure 34. Overview of the file and program structure related to the use of *AROPAj*.

Up until now, the N-to-yield functions building and the spatialization of farm groups are realized outside of the **miraj** environment. However, the "nitrogen-yield" block and the use of spatialization probabilities are included in the **miraj** menu.

The *AROPAj* shell itself is made of scripts written in *CShell*, *awk*, *sed*, *FORTRAN*, and *R* codes. The general application is launched by the **miraj** command, thereby allowing the user to access nested menus. An overview of the file and program structure is presented in Figure **??**.

11.2 Basic use and call of the model

Key-points:

Users, users' rights and the administrator
Some coding rules and links through the *FADN*, parameters, matrix elements and *MP* results
Exporting *AROPAj* outputs and spatialisation
Changing or adding blocks, parameters, outputs for graphics and tables, farm groups, *MS*

11.2.1 Parameter value transmission through *AROPAj* steps and files

Let us recall how a constraint is mathematically expressed (see the general optimization program in Section **??** and Table **??** in section **??**). Let us consider the i-constraint of the k-farmer's program labeled by its related dual value λ_{ki}, expressed as it follows:

$$\lambda_{ki} \ : \ \sum_j E_{kij}(\theta_k, \phi)\ x_{kj} \leq Z_{ki}(\theta_k, \phi)$$

where x_{kj} refers to the j-activity of the k-farmer's program, z_{ki} refers to the *RHS* i-value, and ϕ and θ_k refer to general and to k dependent parameter vectors. The terms E_{kij} and Z_{ki} are matrix elements which depend on basic information provided by the estimation process (see Section **??**).

For example, let us consider the element `RRR` used by the constraint `s(j,k)` of the crop block. This element depends on the `j`-crop yield (per hectare of `k`-farm's crop) which is labeled `rjk(j,k)` (see Section **??**). [11] The core of the modeling chain (referred to as `aro.mgg`) is written in a language close to *FORTRAN* and uses the vector form of parameters, which is `rjk(j,k)` in the case.

As for data, the *FADN* sourced parameters as well as all *AROPAj* parameters are fed into the model in a variety of ways. The *IPCC* sourced parameters, completed by data referring to feed and fertilization, are examples of non *FADN* sourced parameters. All parameters are clustered per *AROPAj* country[12] into files labeled `mgg.par`. We base the *AROPAj* data feeding on the "on file line - one information" principle. When a parameter is referenced several times in the `mgg.par` data file, the due value that is ultimately retained in the *LP* matrix and solved program is the last occurrence in the file.

The typing of parameters is based on a mnemonic criterion (French at the beginning), on alphanumerical character parsimony, and obviously on the discriminating power of character assembling. This leads to replacing numerical indexes by alphanumerical ones (especially for crops

[11] Yield values enter the *LP* matrix either directly from estimated values provided by the *FADN* or via nitrogen to yield functions (see Section **??**). This latter case leads to using an intermediate function called by the element `RRR`.

[12] See Section **??** about farm group numbering.

and animals, see Table **??** for correspondence between files and types). The correspondence between the matrix generator support (`aro.mgg`) using vectors and the data file (`mgg.par`) is realized via the `intern.f` *FORTRAN* subroutine. Generally, the 2-characters strings `jj`, `jk`, `er` and `es`, when located in columns 2-3, refer to crop data (and strings referring to crops appear in columns 4-5). The 2-characters strings `ii`, `ik`, `er` and `es`, when located in columns 4-5, refer to animal data (and animal categories appears in columns 2-3). Concentrated feed and raw feed are identified by 2-character strings `a1`, `a1`, `a1`, `a1` and `ag`, when located in columns 4-5. Parameters referring to effluent categories are codes in the same way. The 5 animal sourced effluents are presented on Table **??**. Urban and compost sludges as well as fermentation digestates will refer to added categories.

As an example, lines of the `mgg.par` referring to crops yields are of the type:

```
...
rjjoh1  :        7.824480
rjjpt1  :       36.363918
rjjtb1  :        1.766110
rjjbl2  :        9.086650
rjjbt2  :       76.504700
...
```

The *FORTRAN* code-lines referring to the yield example are summarized by something like:

```
if(¨1-character¨.eq.'r'.and.¨2-3-character¨.eq.'jj'.and.  ...)  rjk(j,k)=value
```

In the present numerical case, the *FADN* estimate of the 2-farm group soft wheat yield is 9.08665 (t/ha). This value aims at feeding the 2-farm group's *LP* matrix (for one or more member State of the EU). As another example, the following line refers to the raw energy contend in the 3-type concentrate intended to dairy cow of the 41-farm group (for one or more member State of the EU):

```
...
bvla341 :    4600.000000
...
```

The use of this process is due to prior development of *AROPAj* as well as the pressing need for *FORTRAN* variables in *AROPAj* parameters for model coupling and for using *AROPOAj* as a subroutine of a broader program. This is the case when the *AROPAj* calibration is processed via computation tools (see Section **??**). This technical process holds for any kind of constraints and for the objective function.

Changing or adding *GHG* parameters: use of external data This is an example of linking structured external information to *AROPAj* parameters. Regular modifications occur either when we need to update the *IPCC* based parameters referring to the *GHG* block, or when *AROPAj* accounts for new Member States. It also occurs when parameters change in status, for instance when they become crop or animal indexed. Such a change may affect both the upstream *IPCC* sourced database (`ipcc.dat.csv`), the file code `ipcc.awk` and the kernel files (`aro.mgg` and `intern.f`).

Changing parameters value: an example We consider the case of crop price change to be managed through the parameter datafile `aro.mgg`. We extract a series of 2009-FADN sourced farm groups'

Table 35. Crops and animals typing in the parameter file `mgg.par` (2nd and 5th columns) and referred index in vectors used in the core file `aro.mgg` (1st and 4th columns); effluent cateogries below.

crops			animals and bought feeds		
1	bd	durum wheat	2	ve	8-days slaughtered calves
6	bl	soft wheat	3	vg	2-months slaughtered calves
7	oh	barley	4	vm	6-months calves (field)
12	av	oats	5	vt	calves reported on-farm to be young bulls
17	ca	other cereals	7	vb	calves reported on-farm to be more 2-yr bulls
18	se	rye	8	nl	female calves from dairy herd non reported on farm
19	rz	rice	9	nv	female calves from breeding herd non reported on farm
20	ma	maize	10	fr	female calves reported on farm
21	bt	sugar beet	11	jt	male calves bought to be young bulls
22	ba	A-sugar	12	jb	male calves bought to be bulls (> 2 years)
23	bb	B-sugar	13	jl	female calves bought to renew the dairy herd
24	bc	C-sugar	14	jv	female calves bought to renew the breeding herd
25	tb	tobacco	21	te	18-months bulls
26	ct	cotton	23	me	1-yr males from "vb" or "jb" to be bulls
27	li	flax & hemp	24	ma	1-yr males bought to be bulls
31	cz	rapeseed	25	fl	1-yr female on-farm dairy herd
32	tr	sunflower	26	fv	1-yr female on-farm breeding herd
35	sj	soja	27	hl	1-yr female bought for dairy herd
36	pr	proteins	28	hv	1-yr female bought for breeding herd
40	lg	field vegetables	31	be	2-yr males on-farm > beef
41	fr	fallow	33	rl	dairy heifers
42	gl	1988 set-aside	34	rv	bred heifers
43-47	...	various set-aside*	35	vl	dairy cows
49	ja	fix set-aside or forest	36	vv	suckler cows
51-56	ms,...	various SRC**	41	ov	sheep
61	pt	potatoes	45	cp	goats
67	lf	fodder vegetable	51	po	pigs
71	bf	fodderbeet	55	vo	poultry
72	mf	fodder maize		ag	bought raw feed
73	lz	alfalfa***		a1	simple energy-rich concentrated feed
74	sp	protein fodder		a2	simple protein-rich concentrated feed
81	sh	permanent meadows		a3	compound energy-rich concentrated feed
82	af	temporary grassland		a4	compound protein-rich concentrated feed

* refers to the 1992 *CAP* reform

** short rotation coppice such as miscanthus and poplars

*** for some *FADN* years only, not maintained

Effluent categories (animal-dependent or urban waste)

1	i1	swine & bovine	slurry
2	i2	all except poultry	"solid" manure
3	i3	poultry	manure
4	i4	poultry	slurry
5	i5	poultry	droppings
11	cu	urban	compost
12	dv	urban	liquid

prices related to soft wheat in one of the *EU* country as follows. In addition we consider the relative price parameter allowing to changes all soft wheat individual prices.

```
...
pjjbl   : 1.
pjjbl 2 :     0.23000
pjjbl 3 :     0.15563
pjjbl 4 :     0.12505
pjjbl 9 :     0.12743
pjjbl10 :     0.17510
pjjbl11 :     0.21901
pjjbl13 :     0.23490
...
```

As explained above, parameter code refers to combinations of 1-, 2- or 5-character strings. In the present case, the first column character (`p`) refers to the price category among parameters, the 2-3 column characters `jj` means that information contained in the datafile line refers to crops, and the 4-5 column characters `bl` refers to soft wheat as detailed in Table **??**. When no other indication exists, the parameter refers to the relative price which is 1. as default value. Reversely figures in columns 6-7 refer to the farm group numbering. It has to be noted that " 1", "1 " and "01" are of equivalent interpretation. Nevertheless when dealing with other data files of `mgg.par`-like format, the user should select the "1 " or "01" forms (this point refers to the `scen.par` file mentionned in Section **??**). The parameter value should be registered in columns 10-25 (right from the ":" character) in standard FORTRAN-like format (namely f16.6).

When the model's user wants to change all soft wheat prices in a common ratio (e.g. increasing by 50%), the value of the non-numbered parameter `pjjbl` has to be changed from 1. in 1.5, either in the last line of the date file containing this parameter or in an additional line registered as:

```
pjjbl   : 1.5
```

We recall the rule of "last read value - stated value in the *LP*".

11.2.2 Reporting

In addition to the *GAMS* files provided by the kernel file transformation, another *GAMS* file (`report.gms`) is produced with the aim of listing the *LP* comprehensive solution in terms of primal and dual variables. Based on other *GAMS* files and additional flags or descriptors (see the `aro.dat` file), this file is automatically built by a *FORTRAN* routine when the user launches *AROPAj*.

FORTRAN routines provide formatted output files designed for downstream treatment.

By default via minimal use of the *AROPAj* shell

The organization of calculus, as illustrated by Figure **??**, and the organization of files, as illustrated by Figure **??**, refers to the standard manner of using *AROPAj*. In this organization, most of user's file names are indexed by the computer clock system in two ways.

Firstly, the names of general information files related to the simulation process management use the 12-digits index "YYMMDDhhmmss" (denoted by "$*idm*"). Since both trace files and the

main table file use this format, the user can easily follow the process evolution over *AROPAj* steps and over *AROPAj* countries. This table file is named `tableau.$idm` when "$idm" accounts for the 12 digits.

The second way of indexing files uses the 8-digits index "DDhhmmss" format, which is also based on the system clock provided at the *AROPAj* country scale (denoted here by "$*id*"). Most of intermediate and final output files are renamed in accordance with this index. Standard comprehensive outputs provided by the *GAMS* solver (`sortie.gms`) are reported on files named `mggf$`*`id`*`.sortie`. Extracted results formatted through the *FORTRAN* routine (`scen.f`) are reported on files named `result$`*`id`*`$`*`k`*, k denoting the farm group. The same principle is used for intermediate files used, for instance, in iterated *AROPAj* runs, when parameter values at the iteration n depend on *LP* results provided at the iteration $n-1$.

Applications are developed to display aggregated and tabulated results at the farm group level as well as at the national level. The user accesses these files in accordance with the repertory hierarchy illustrated by Figure **??**. On the upper level, results may be compiled in "txt" or "csv" format for "countries" for which *AROPAj* runs are realized. Additional two-dimensional indexation refers to run iterations that are conducted on a two-dimension grid (like when two parameter ranges are explored). The file name is of the type `table.compil.$scen.$n.$m.txt`.

Changing output structure

The user always has the ability to change output files and their structure. The question here is not to discuss what can be done from default outputs freely exported (default outputs are easy to manage through personal computer tools). In the case of structural change upstream from *GAMS* files produced by the kernel, the reporting is automatically adjusted. In other words, the *LP* results are comprehensively displayed in the `output.gms`. The user may effectuate downstream change through his/her own shell script `scen.csh` in order to produce *FORTRAN* codes and related files. The *AROPAj* shell requests only to have a shell script file in due code. In order to fully benefit from all the applications around the *AROPAj* shell, we advise keeping basic structure of file and output formats.

Output spatialization

The spatialization process can be use when the user created a file of the type `table.compil.$scen.$n.$m.txt`, and when the farm group location probabilities have been previously estimated. Let us recall that these probabilities are only computed for a few *AROPAj* versions and *MS* (Sweden and Finland are excluded, for the *V2* version). The process implicitly calls for *R* code files which are included in the *AROPAj* shell. Attention should be paid to the output file size. Outputs are displayed in compressed "dbf" format files. These files are provided and indexed according to *FADN* region. File names are of the type `Arc.$scen.$reg.$n.$m.dbf`.

11.2.3 Inserting new items, linking parameters and outputs

Extending the *AROPAj* coverage

The coverage extension refers to different kinds of *AROPAj* enlargement, wether it be in terms of geography, crops, environmental issues, policy tools, or model applications.

MS inclusion is straightforward (no change in kernel files):
in particular, typology, parameter estimation and calibration remains unchanged.
Country data from the *FADN* is a prerequisite.

crops (or other agricultural activities):
introducing cotton, hoax or tobacco for instance requests *FADN* sourced information
For perennial crops, such as miscanthus, the user must calculate the discounted values of (added) parameters

LP variables (activities and constraints, e.g. a new block):
changes have to be made in the kernel files (`aro.mgg, intern.f, aroN.dat`).
Related parameter estimates have to be added accordingly with the requisite format of the `mgg.par` file.

It is also possible to call external data, such as *N* to yield functions. In this case, `lisscal.don` files are read by the `intern.f` routine. Yield function parameters are used in each *AROPAj* run as long as the user calls them through the shell script for that scenario (`scen.csh`). This allows the user to change yield function parameters from one run to another (e.g. in case of climate adaptation).

Parameters valued through *LP* outputs

We have briefly exposed several cases relevant to projects and papers mentioned in previous sections. The user must adequately prepare the `scen.csh` file in order to process the desired application.

Referencing *GHG* quotas Instead of the "primal" approach based on *GHG* emission pricing (through tax), the "dual" approach entails using the emission quota shadow price. We must set these quotas, based on a reference situation without binding limits. To do so, we compute reference emissions and we transform these outputs in quota parameters (`qopes(k,l)`).

Valuating the economic *LT* effect of *OM* on crop nutrient This effect enters the model through the `aponlt(k)` parameter and lowers crops' *N*-requirement. This *FG* indexed parameter depends on annually renewed amounts of *OM* spread on arable lands. In the model, the available *OM*, be it on-farm sourced (manure) or imported (*UOR*), is optimally allocated to various crops and fodders with respect to regulatory limits. In other words, this amount per crop area is endogenously calculated by the model, thus impacting the `aponlt(k)` value.

Fixing crop areas In some cases there is an interest in constraining certain crop areas on order to, for instance, assess the part of different sources of autonomous adaptation in economic or environmental impacts. The user mus thus set reference areas, i.e. the `xjk0(j,k)` parameters, which are in turn evaluated by the reference *AROPAj* outputs.

Shadow prices substituting market prices By adding new parameters, the user can easily process the iterated substitution of market prices by relevant shadow prices in the yield and *N*-input calculation. On the "crop output side", new parameters (`rojk(j,k)`) refer to the dual values related to the `s(j,k)` constraints. We had not added yet the relevant parameters on the "*N* input

side", because there are very few cases of solely organic N input dedicated to a crop by a farm group.

The *MODANI* model and equilibrium prices This application is based on an iterated *tâtonnement* process that attempt to compute the converging prices associated with balanced markets. In this case, at each iteration, study-relevant market prices are changed in line with the excess of demand.

The *AROPAj* calibration The *AROPAj* calibration process is one of the primary instances in which *AROPAj* acts as a subroutine of a wider program. As explained above, the process is based on MonteCarlo and gradient algorithms, which require a large amount of *AROPAj* runs. Each run is iteratively based on re-evaluated parameters.

11.2.4 Starting and managing simulations

There are several prerequisites for starting *AROPAj* and calling a simulation scenario. First, a user account has to be created on the workstation that host *AROPAj*. In addition to the usual change of the "PATH" variable associated to the user's account, some small changes have to be reported in the "login" scripts. The *AROPAj* administrator has to report the user's account characteristics in the *AROPAj* shell scripts.

AROPAj simulations are launched through the `miraj` command script, thereby supplying nested "menus". Initiating a simulation only requires a `scen.csh` script from the user.

The menu allows the user to undertake many operations. It also allows the administrator to undertake basic *AROPAj* steps such as farm group clustering, parameter estimation and calibration. The user can:

- access information files referring to scientific projects (named `cahier.txt`)
- successively or automatically initiate the different steps relating to the version choice and parameter complements, for different "*AROPAj* countries" covered by the selected version
- iterate the *AROPAj* runs as designed by the script file `scen.csh`
- aggregate the outputs at the country level
- tabulate the outputs at the country level
- export the output compilation in "txt/csv" format files for countries covered by simulations and for selected iterations (the user is free to select a subpart of iterations among all of them)
- undertake all the previous steps once the simulation is initiated ; this entails selecting the version and the script `scen.csh`, and initiating the process for just one country
- display selected results through the spatialization process for *FADN* regions covered by the simulation when the spatialization is available; for this step, the user is free to select outputs and iterations among those scheduled by the script `scen.csh`.

During the computation process, as explained above, several files are reported in the working directory and in the subdirectories as shown on Figure **??**. Process visualization is briefly developed in the Appendix **??**.

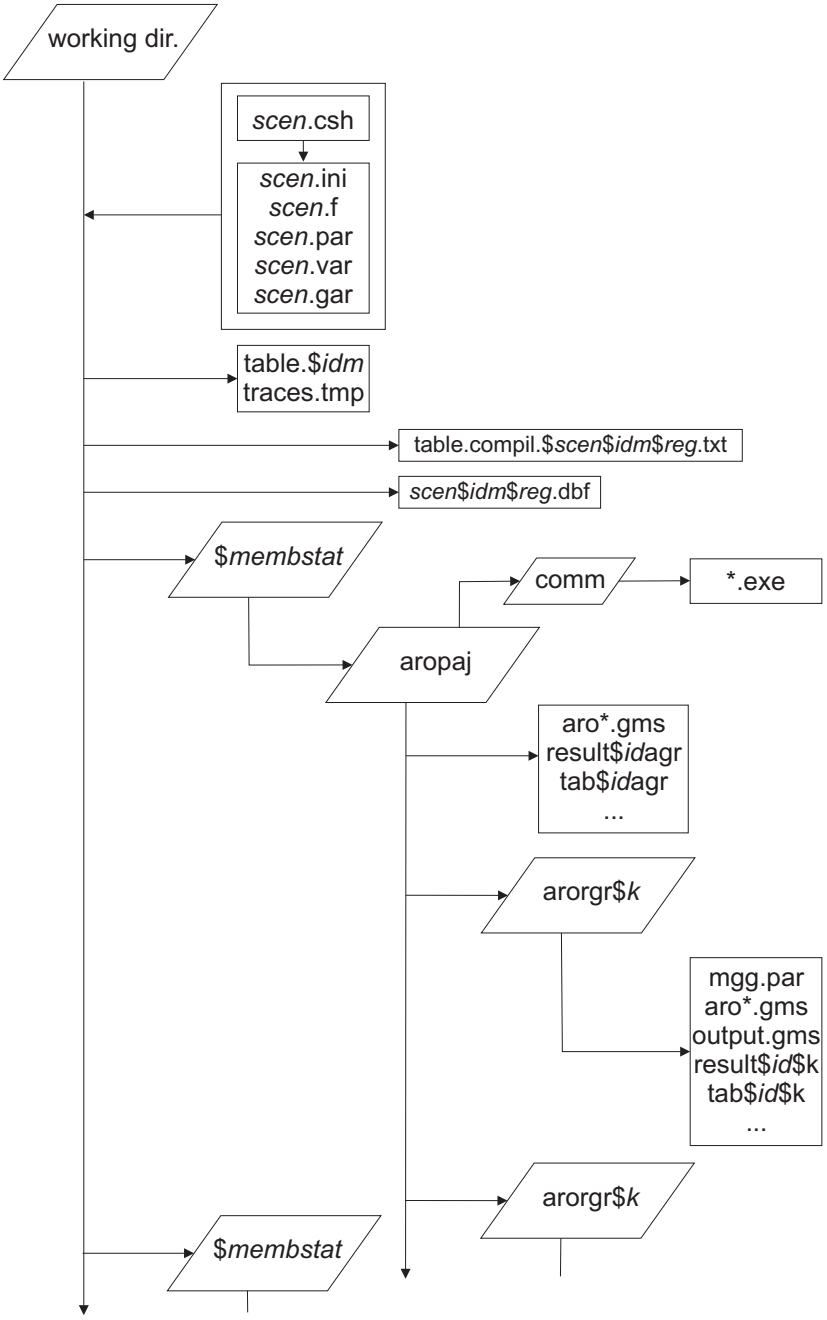


Figure 35. Overview of the file and program structure related to the use of *AROPAj*.

12 Applications

We explore some of the results and the applications of *AROPAj* capabilities for a wide range of issues. These findings mainly refer to published papers. We add a few results enlarging this range or providing renewed estimates across *AROPAj* versions.

12.1 Agricultural policy analysis

> Key-point:
>
> The *CAP* as a major driver and scenario provider

Analysis of the 2003 *CAP* reform

A strong challenge for the model occurred with the 2003 *CAP* reform (namely the "Luxembourg agreement") and its analysis through the European *GENEDEC* project. Technical information available in addition to published papers is collected in **?** (see also **?**). Policy tools introduced by the agreement were in large part implemented in the model as they were applied in to the *EU*-15 (see Section **??**).

The first step was to assess the various levels of support coming from the former policy ("Agenda 2000"). This was used to calculate "historical" reference support per hectare and per farm (more precisely per farm group in *AROPAj*). This was used also because *MS* kept the right for a few years to (partially) recouple some supports. This is detailed in **?**. We recall here the policy tools explicitly taken into account in the analysis.

Full decoupling of beef production is one of a number of options that are available at the *MS* level. The other two options are either 1) to retain up to 100% of the suckler cow premium, 100% of the slaughter premium for calves and 40% of the slaughter premium for other animals (the choice of Belgium, Spain, France, Netherlands, Austria and Portugal), or 2) to retain up to 75% of the special beef premium and 100% of the slaughter premium (the choice of Denmark and Sweden). For crops, the intervention price remains unchanged. Apart from full decoupling, the options are to either retain up to 25% of the arable crop payments in the coupled form (independent of set aside payments, as France and Spain have chosen), or retain up to 40% of the durum wheat payment in a coupled form. For sheep, apart from full decoupling, the available option is to keep up to 50% of the ewe premium (the choice of Denmark, Spain and France). As of 2005 (2007 at the latest), the new milk premia can be part of the single farm payment. Mathematical programming models make it possible to compute the single farm payment, *SFP*, as well as compare farmers' gross margins before and after the reform.

The computation of individual or regional decoupled payments in the *AROPAj* model is based on the results obtained using the Agenda 2000 policy as input. In the model, the single year of simulation is representative of the 3–year reference period. The prior *AROPAj* supports have been broken up according to different items related to possible decoupling combinations:

1. basic support for cereals, oilseeds and proteins
2. specific support for durum wheat
3. specific support for proteins

4. set-aside support
5. extensification support related to livestock
6. support for milk
7. support for ovine, caprine, and generally for "small" herd
8. support for suckler cows
9. support for male bovine
10. support for slaughter calves
11. other slaughter supports;
12. other supports, possibly excluded from decoupling (taxes, sugar regimes)

The *AROPAj* modeling of the decoupling reform depends on the *FADN* regions and farm groups within regions. Thus, it is possible to compute the *SFP* for each farm group and to compute the unique regional entitlement. Some supports from the Agenda 2000 policy are currently not in the model. The main reason is the general lack of data. For example, the *FADN* does not supply information about product destination, and it does not distinguish between some types of crops, e.g., starch and food potatoes. Consequently, neither does the model. Nevertheless, the total European budget devoted to agricultural policy is well represented in *AROPAj*. The regionalisation option can be implemented without any difficulty in the *AROPAj* model when the regions covered by this option correspond to *FADN* regions upon which the *AROPAj* model is based (this is the case for Germany). Otherwise, some hypotheses on farm groups' location within a *FADN* region are necessary (the case for England).

The changes in the *CAP* related to the Luxembourg agreement are implemented in the model as explained in Table **??**. Values of policy tool parameters are provided by the previous step of reference estimates. We introduce a "full decoupling" scheme through the unique payment (`psuni`), which is estimated per *AROPAj* farm group.

We focus on two series of results which appear especially interesting. We consider three scenario changes between the "reference" and the full decoupling scheme. The scenarios are defined as follows:

AG00 refers to the reference scenario, based on the calibrated 2001-*FADN AROPAj-V2* version

AG15 only the livestock adjustment parameter is changed (from 0% in `AG00` to 15%)

LX15 the Luxembourg agreement is implemented in the model, keeping the *LS* adjustment at the previous level

FD15 there is no partial recoupling and the farm group *SFP* applies, keeping the *LS* adjustment at the previous level

Allocation of agricultural lands We present series of three-figure sets showing relative change in area allocated to (i) cereals (**??**), pastures (**??**), set-aside (**??**) and fallows (**??**). [13] Changes in areas are computed at the *FADN* level. It is clear that they differ compared to the relatively slight average changes at the European level.

[13] Fallows are freely chosen by producers, meanwhile set-aside is requested by the *CAP*.

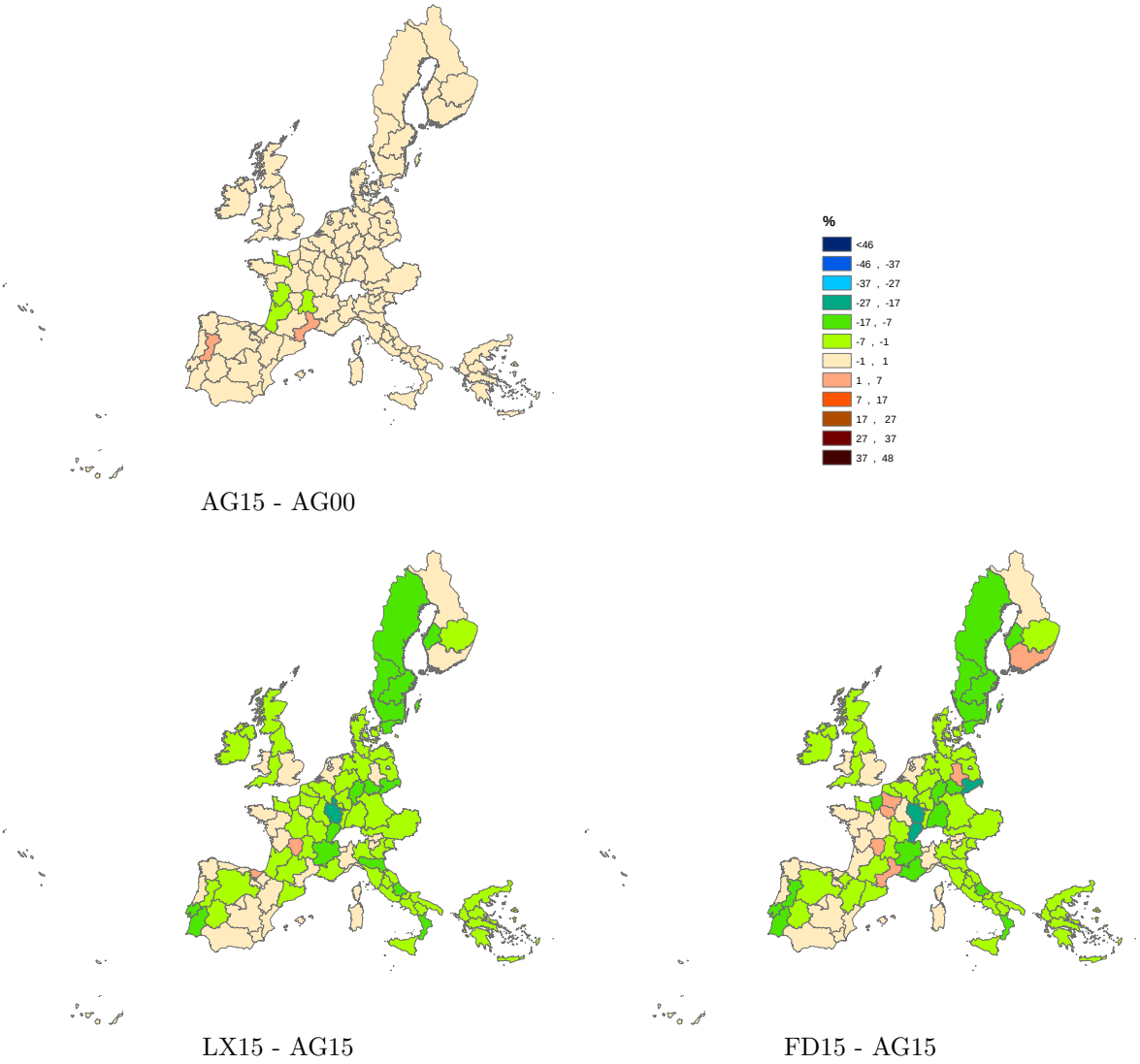


Figure 36. Variation in area devoted to cereals between the scenarios AG00, AG15, LX15 and FD15 (in % of the regional agricultural area).

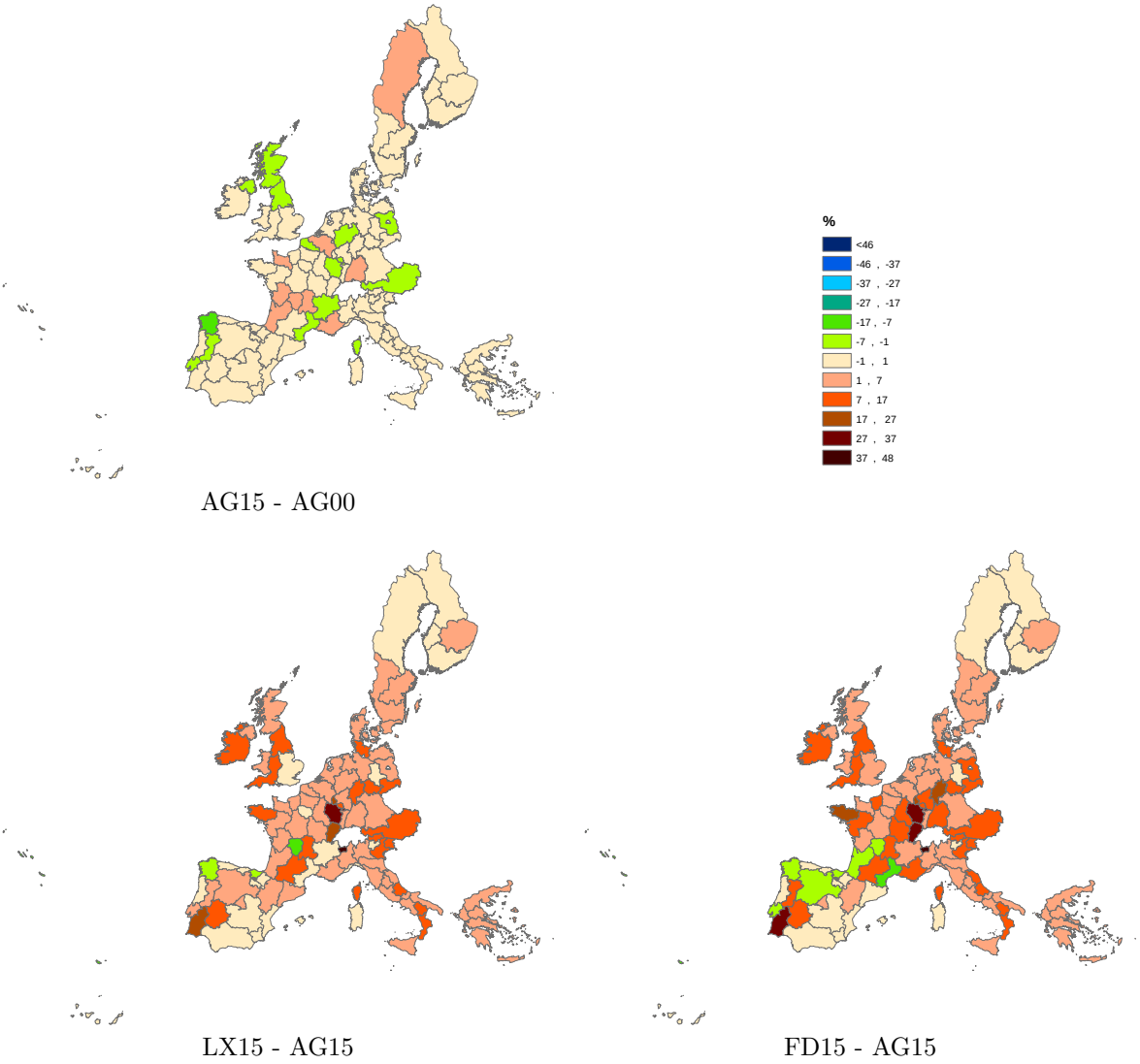


Figure 37. Variation in area devoted to pastures between the scenarios AG00, AG15, LX15 and FD15 (in % of the regional agricultural area).

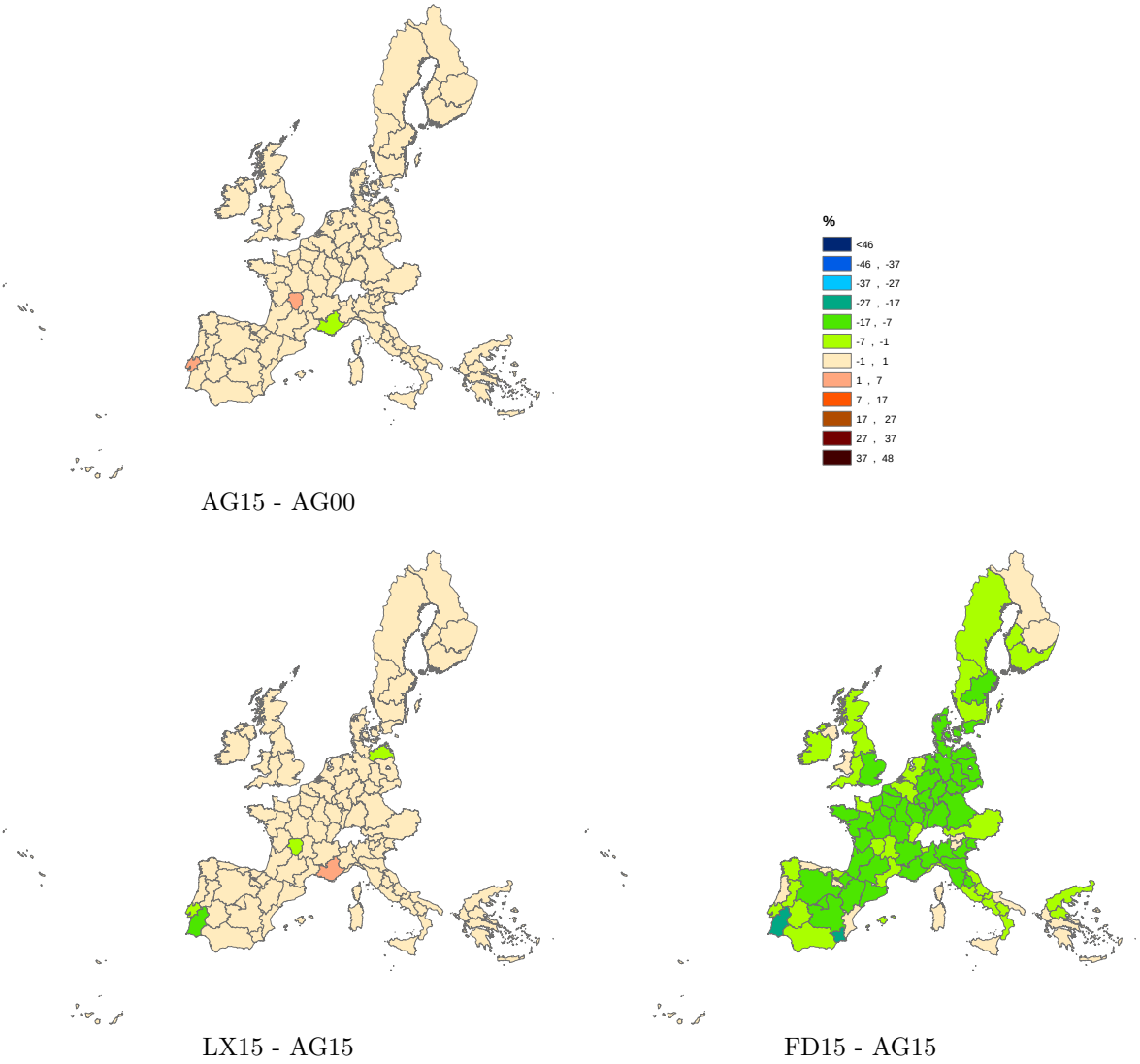


Figure 38. Variation in area devoted to set-aside between the scenarios AG00, AG15, LX15 and FD15 (in % of the regional agricultural area).

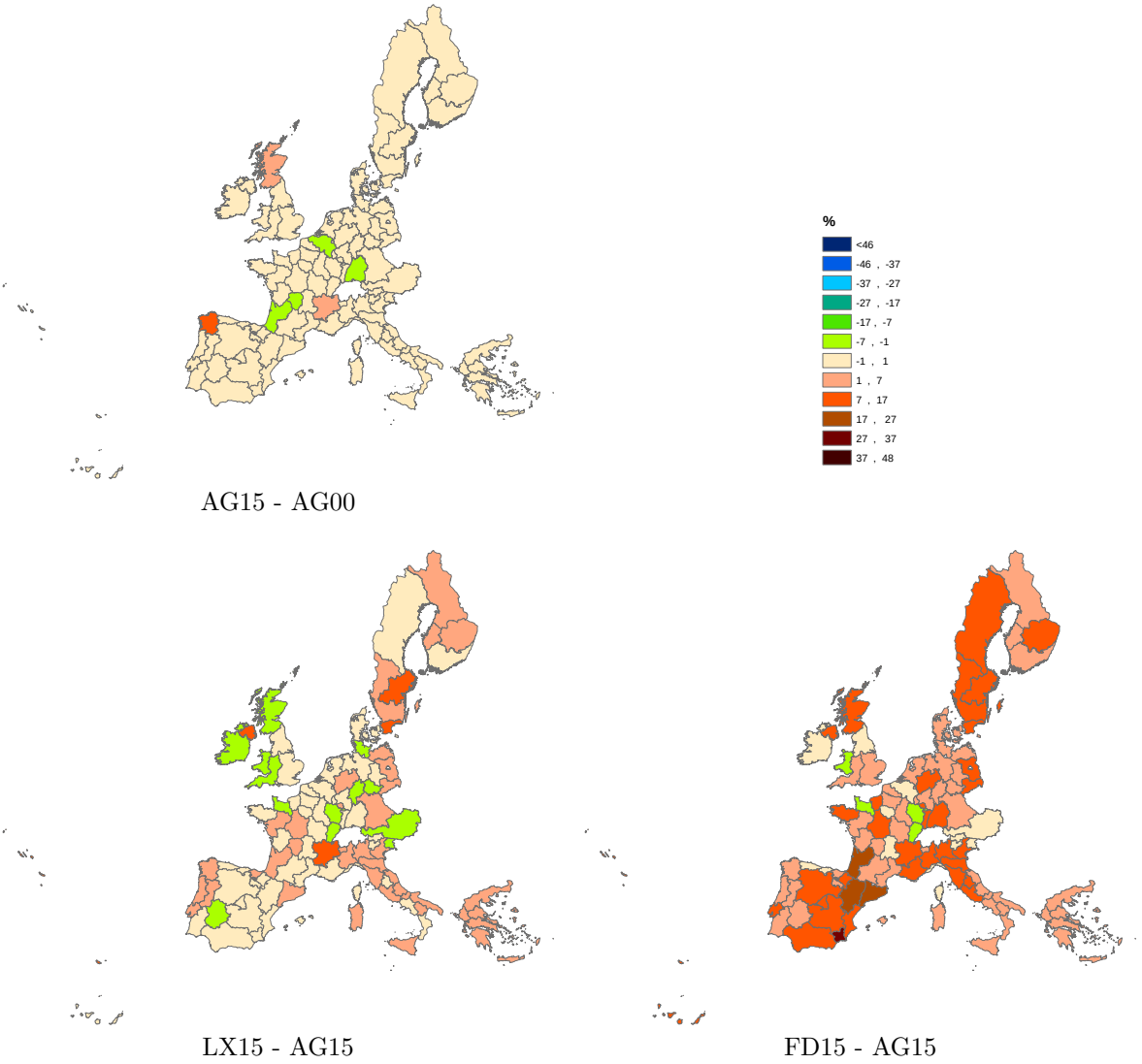


Figure 39. Variation in area devoted to fallows between the scenarios AG00, AG15, LX15 and FD15 (in % of the regional agricultural area).

Land shadow costs The scenarios previously defined now apply for providing estimates of the land shadow cost. We introduce a slight change in the implementation of the scenario when the unique *SFP* payment is paid to the whole farm, as a lump-sum transfer. Considering this transfer with the 0-value is equivalent to suppressing the *SFP* ("no support" scenario). Since it is a "constant" element in the mathematical programming problem, the complete solution in case of "lump-sum support" should not differ from the `FD` scenario with the *SFP*, except for the shadow land cost. This is due to the fact that, when reporting the shadow cost form the *LP* output file, we take account of a unique additional term in the gross margin at the plot level. In the first case (*FD*), the *SFP* is added to each plot's gross margin, while in the other case ("lump-sum support"), the support stands for a "fix" term. Figures **??** show the changes in the land shadow cost among 4 scenarios, namely `AG15`, `LX15`, `FD15` and `no support`.

Managing new *CAP* tools: the case of milk market organization

The *AROPAj* model was asked to provide insight on different policy changes with regards to existing milk quotas and to the European milk market. In this study initiated by the *EC*, various options were introduced in the model:

- add and modify a milk premium
- create a simple quota market
- create a double market (fixed quotas and market dedicated to out of quota quantities)
- suppress the *CMO* (no quota, no subsidy)

No results can be displayed here for confidentiality reasons. For our current purposes, we must only focus on the method employed to introduce a quota market allowing for exchanges between farming systems at different scales.

We put annual quotas into consideration, which is in line with the basic *AROPAj* period. Introducing the rights to buyl or sell quotas is very simple in a *LP* model. We need to add:

- a parameter related to the quota exchange price (`pqlai`)
- a primal *LP* variable indexed on the farm group and related to the sale or purchase of milk quotas (`xolqm(k,l)`, considered as a "free" variable which can be positive or negative)
- a dual *LP* variable related to the quota demand excess (`qolqt`), which is of interest for aggregation at the scale desired by the user

The market clearing price at the desired scale is easily approximated by *AROPAj* runs realized for series of `pqlai` prices valued in an adequate range. This "one dimension" problem has a unique solution for each aggregation level (i.e. for each of the *MS* and for the *EU* as a whole).

The introduction of a "second market", dedicated to out of quota milk production, can be easily added to the milk quota market. Parameters and *LP* variables are listed in the *CAP* blocks, which refer to "second market" in their description.

Post-2007 *CAP*

The *CAP* is a major economic framework shaping the European agriculture for many years now. However, the decoupling scheme initiated in the 1990s is currently applied with less and less

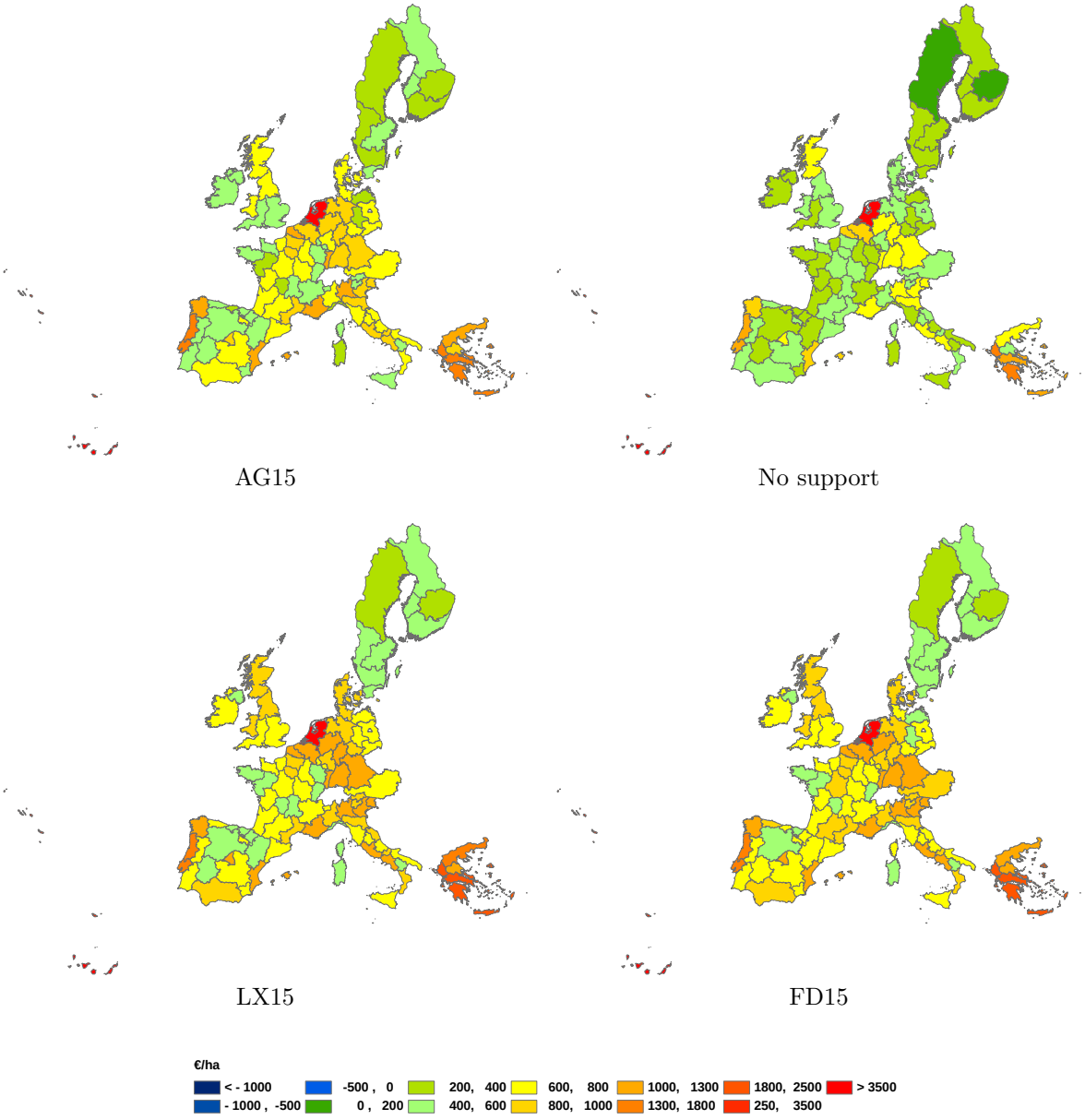


Figure 40. Shadow costs of land within the scenarios AG15, LX15, FD15, and in the case of a "no support" policy, for land and activities covered by the AROPAj model.

national- or product-specificities. The default setting in *AROPAj V5* policy tools refers to the full decoupling scheme with an updated quota system (no dairy product or sugar quota). In our model, we can easily reactivate any combination of *CAP* tools as necessary for future numerical explorations.

The *AROPAj* model has been updated with environmental policy tools to tackle rising environmental issues. Besides the *CAP* policy tools, additional economic regulation related to the environment and bio-products need to be taken into account. As a matter of fact, in addition to the *CAP* framework tools, the *AROPAj* model is including tools related to nitrogen pollutants, greenhouse gas emissions, climate change but also bio-fuel production and natural resource preservation. At last, *AROPAj* remains a supply-side model in which market prices are exogenous and parametrized as we should do in case of regulated prices.

12.2 GHG emissions and abatement

Key-points:

GHG emission assessment initiating *AROPAj* evolution and supplying original analysis

Estimation of marginal abatement cost curves (*MACC*)

In order to derive marginal abatement cost curves in this type of model, an emission price (expressed in €/tCO_2eq) is introduced in the objective function of the model. It affects each farmer's revenue according to the total amount of emissions emitted at the representative farm level. By construction, for a given emission tax t, emissions are such that the marginal loss of income due to an additional reduction equals t at the individual optimum for any farm k. By letting t vary in a given range, one thus obtains the quantity of abatement that each farmer is ready to supply at this price or, equivalently, the *MACC* pertaining to this particular farm. Note that, by using the same level of tax for all representative farms, one obtains the cost-effective repartition of abatement across farms. Several applications of the model have used this procedure to assess *MACC* at various levels of aggregation (**???**).

Illustrative examples

Abatement costs are a fundamental determinant of the role that agriculture could play in meeting efficiently the *EU* commitment to reduce its *GHG* emissions. Two broad dimensions have been examined by **?**: the magnitude of abatement costs in the agricultural sector and their heterogeneity.

Regarding the magnitude of abatements from agriculture for plausible carbon prices, the range is found to be substantial. Despite rather conservative assumptions – such as the fixed number of farms, fixed total area, fixed crop yield response to nitrogen, no adoption of specific abatement technology – an additional potential abatement of about 4% of 2001 agricultural emissions is obtained at a price of 20 €/tCO_2eq. The drop of abatement is of 8% at a price of 56 €/tCO_2eq. In the design of economic instruments, one faces an important trade-off between accuracy and observability. In this respect, the *IPCC*-based economic instruments, which we analysed, provide an interesting balance. On one hand, complex emission processes are only imperfectly captured

because of the use of simplified relationships between activities and emissions. On the other hand, the *IPCC* framework is recognized as an international reference and is based on data that is relatively easy to collect. The study was based on a preliminary *V2* version of *AROPAj* applied to 15 *EU* Member States (*FADN* 2002).

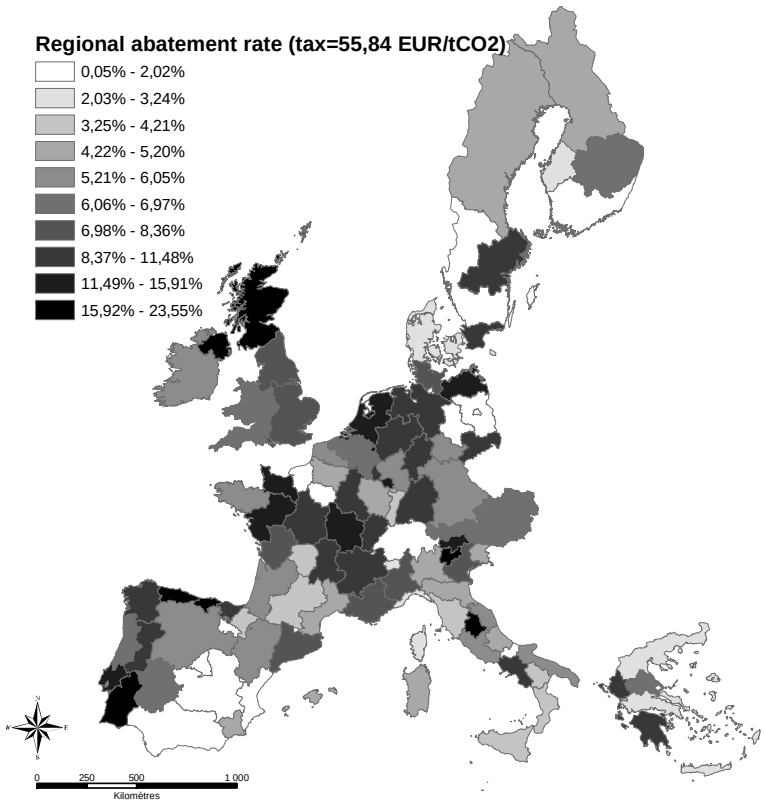


Figure 41. Abatement of *GHG* emissions at a price of 56€/tCO_2eq averaged at the *FADN* region level. Source:**?**.

According to this study, abatement cost heterogeneity comes from a variety of sources. The farm type approach captures some of these sources of heterogeneity in the results, such as those related to farms' size, crop yields and area allocation, total animal numbers, input use, and *CAP* support. Some are only captured at the country-level (e.g. *IPCC* emission factors). And, due to the lack of data, some of the sources of heterogeneity have been ignored (e.g. spatially disaggregated emission factors). Nevertheless, the results indicate a wide variability of abatement costs (Figure **??**). This has two broad implications for policy purposes. First, the impacts of incentive-based instruments on income and environmental performances vary widely from one farmer to another. Second, the cost savings permitted by market-based instruments relative to uniform standards are substantial. This means that, if mitigation policies are to make use of quantity-based instruments, substantial savings can be drawn from tradable (or at least differentiated) emission allowances. Clearly, this also means that such policies will have to go beyond uniform or region-specific standards commonly used in agricultural policies (e.g. set-aside rate or regionally differentiated reference yields). At last, this underlines the importance of the inclusion of agriculture in the European Emission Trading Scheme (*ETS*).

According to **?**, the assessment of marginal abatement costs (*MAC*) is key to the issue of cost-

effectiveness. The study carried out a quantitative assessment of marginal abatement costs of *GHG* emissions from the European agriculture and analyzed the implications of the *EU* burden sharing agreement (*BSA*) for this sector. To do so, a generic specification of *MAC* curves was proposed and applied to the *V3* version of *AROPAj* (the *EU*-25 except Malta, *FADN* 2004, with 1307 farm groups).

Their findings are threefold. First, the heterogeneity of *MAC* across Member States stands out as an important feature. As the agreed targets under the *BSA* do not reflect the heterogeneity of agricultural *MAC* at the Member State level, the use of flexibility mechanisms may provide substantial cost-savings compared to the strict implementation of each individual country's target. Second, the equilibrium price range at which the *EU* 10% reduction target can be reached in European agriculture (32–42€/tCO_2eq) is in line with results from other analyses covering all non-*ETS* sectors and lower than that found in previous studies focusing on *EU* agriculture (56 €/tCO_2eq in **?** for an 8% reduction target). This suggests that the agricultural sector could play an important role in meeting the overall *EU* target in a cost-effective manner. Third, the use of the *BSA* targets as a basis for allocating allowances in a cap-and trade system for agricultural emissions may have important redistributive impacts.

For the actual simulations, the *IPCC* parameters for 27 *EU* Member States have been updated (*FADN* 2009). As well as the *IPCC* parameters, the prices of fertilizers have been updated.

Using the MACC from AROPAj to asses the benefit from an emission tax when it is costly to administrate

The magnitude and the heterogeneity of the abatement costs are of great concerns when the mitigation policy is costly to administrate. In the context of climate policy, these costs are due to the costly monitoring, reporting, and verification (MRV) procedures that ensure the compliance with the policy objective. Indeed, as soon as the policy involves MRV costs that are increasing with the number of agents, it may well be the case that including the whole agricultural sector within the mitigation policy could lead to greater MRV costs than the net benefits from the emissions abatted. Therefore, it can be better not to implement any policy than fully covering the sector to mitigate agricultural GHG emissions.

The work in **?** studies an instrument design that bypasses this issue. As is the case in some environmental policies (the EU-ETS or the CAP concerning the 'green' payment), some exemptions are granted for some firms, e.g. the small emitters. This paper explores the implications of such exemptions in the case of an emission tax, and in the presence of monitoring, reporting, and verification (MRV) costs. They develop an analytical framework capturing the trade-off between the cost-effectiveness of a broader tax base, and the savings on MRV costs enabled by a partial coverage. Second-best partial coverage is defined by a threshold value of some characteristic of the firms (e.g. size as measured by initial emissions) below which firms are exempted. It then allows to characterize the optimal threshold and discuss its welfare implication. The optimal threshold and the welfare implied greatly depend on the magnitude and the heterogeneity of the abatement costs, alongside with the magnitude and heterogeneity of the MRV costs and the marginal damage of emissions. Indeed, the more efficient are the bigger emitters in reducing their emissions, we will expect the coverage to be smaller and then save a significant amount of MRV costs on the smaller emitters exempted from any mitigation effort. However, in practice there is not a monotonic

relationship between the abatement cost and the size of the firms.

Then, the paper uses the MACC from AROPAj to assess the welfare implication of such a scheme. The study is based on the version V.5. of AROPAj applied to the EU-27 with 1802 farm groups. According to this simulation, the agricultural GHG emissions are rather concentrated, since the top 25% emitters emit almost 80% of total emissions as depicted by the Lorenz curve in Figure **??**.

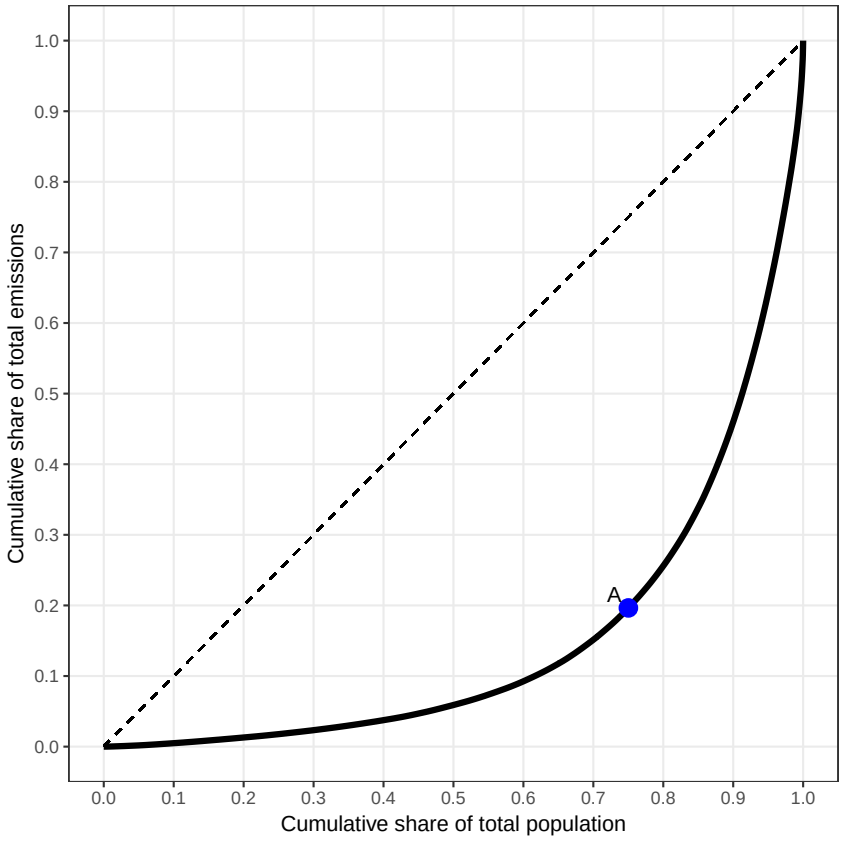


Figure 42. Lorenz curve of initial emissions. *Source:* **?**

Thus, targeting only the top 25% emitters (those to the right of point A in Figure **??**) saves the MRV costs associated with the remaining 75% of farmers, while still covering almost 80% of total agricultural emissions. Of course, the setting of the optimal coverage will all the more depend on whether the small emitters are less efficient at reducing their emissions than the large emitters. Therefore, the welfare implications of partially covering the agricultural sector within the climate policy is dependant on the heterogeneity of abatement costs and their distribution among agents. The more concentrated are the emissions, and the more efficient are the big emitters in mitigating their emissions, the smaller the coverage and the higher the welfare gains (for a given MRV costs level). Thus, they use the AROPAj model to assess the relationship between the abatement efficiency and the size of the farm. They find an abatement elasticity with farm size of 1.1 (which means that when the initial emissions of the farm increases by 1%, its abatement efficiency increases by 1.1%). This abatement cost elasticity is greater than the MRV cost elasticity. Thus, it ensures that the bigger the farms, the greater the net social benefit from its inclusion in the emission tax.

In their benchmark case (MRV costs increasing and concave per farm, averaging 275 €.farm^{-1}, or 2.50€.t$CO_2{}^{-1}$eq and maginal damage of GHG emissions set at 30€.t$CO_2{}^{-1}$eq), the top 10% of the bigger farms are covered (see Figure **??**). It leads to a net social benefit of 124 M€and abatement costs of about 184 M€(to be compared with a sector-wide gross margin of about 140 billion €). Thus, their empirical application shows that including the agricultural sector within the European climate policy is not an all-or-nothing solution, since exempting the small emitters may provide significant savings on MRV costs compared to the full-coverage, while still incentivizing cost-effective reductions in emissions.

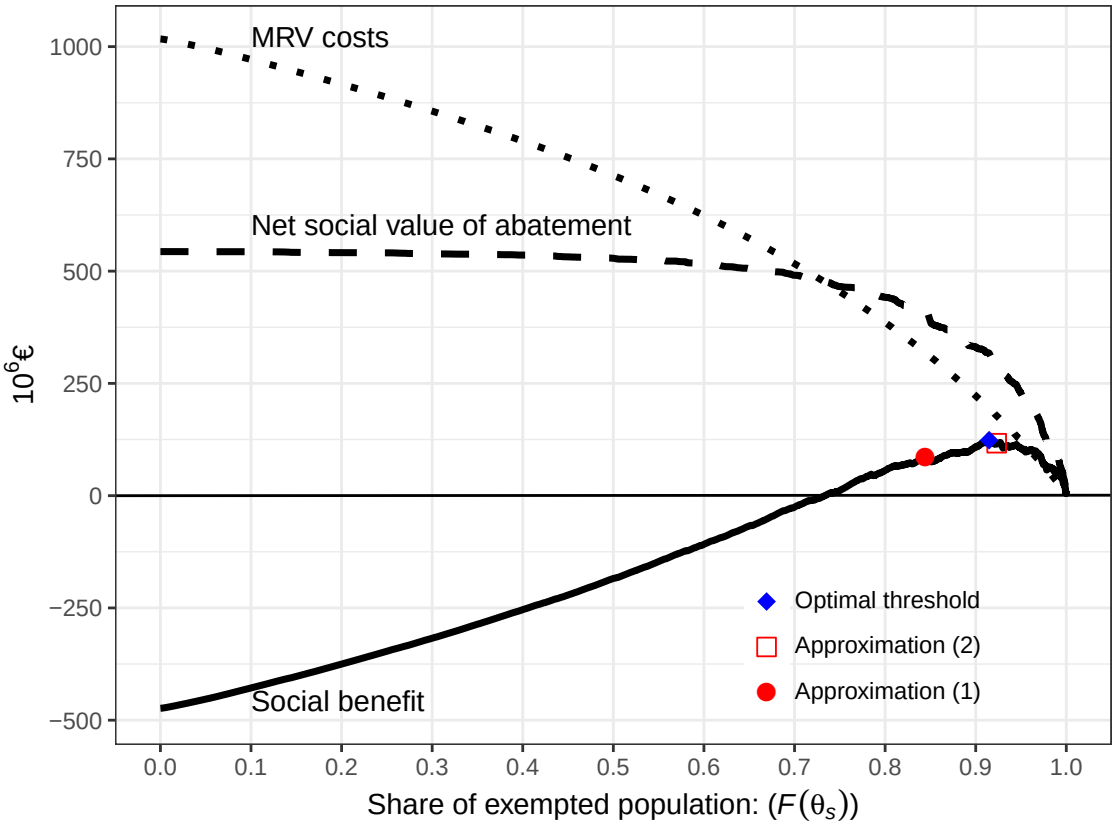


Figure 43. Total social benefit (solid curve) and its components in the benchmark configuration. The net social value of abatement is directly derived from the AROPAj simulation results *Notes:* $\delta = 30$ €$/tCO_2eq$, *medium, increasing and concave per-farm MRV costs. Source:* **?**

In terms of mitigation policy design, this work could be extended in several directions. First, the analysis of an emission tax could be adapted to examine a cap-and-trade mechanism. Although the fundamental mechanisms at work would remain, this would require to take into account the costs related to the trading of allowances. Since these costs depend on the level of abatement, this would introduce a wedge between marginal abatement cost and the emission price. Second, the simple second-best approach developed here could be compared to a more complex mechanism design aimed at revealing individual information. The empirical model used in this paper could serve as a basis for quantifying the associated information rent. Third, the introduction of a partial coverage might cause leakage effects, and/or induce strategic behavior from firms in response to implementation of partial coverage. For instance, we might observe firms to bunch below the initial

emissions threshold in order to be exempted.

12.3 N-sourced pollutions

> Key-points:
>
> Environmental policies and policy mixes accounting for multiple activities, multiple agents, multiple pollutants and widely diverse spatial conditions

Possible counterproductive effect of input taxation

The first empirical application of the coupling between *AROPAj* and *STICS* was to investigate the effect of a nitrogen tax on agricultural production in France under different policy scenarios (**?**). Additionally to estimating yield response functions for every crop and farm-type combination, linear functions of nitrate losses at the root zone were also estimated, using the output from the nitrogen balance module in *STICS*. The model was able to reveal the uncertainty that characterizes second best instruments, such as input taxes, and showed that the N-tax could actually lead to the completely opposite results than those expected in theory.

More precisely, it was demonstrated that the tax can force a "crop substitution effect", which is caused by the change in the relative profitability of the different farming activities; as the N-tax increases, and due to the different yield functions of each crop, the changes in crops' gross margins lead to an increase in area allocated to the most profitable ones, which, in turn, emit more nitrate than the original crop. These paradoxical results indicate that the decrease in nitrogen use can be sometimes overlapped by extensive margin changes that involve an increase in land allocated to more profitable, but also more polluting crops.

This paradox is illustrated on Figure **??**.

Joint pollutions

? assess the cost-effectiveness of a mixed policy in attempts to reduce the presence of three nitrogen pollutants: NO_3, N_2O and NH_3. The regulation of joint agricultural pollutants remains a challenge for applied economics and modelling techniques when pollutant emission factors differ strongly and depend on soils and farming practices. Thus far, policies have addressed these issues on a pollutant-by-pollutant basis. However, the land-use and crop reallocation induced by each policy (alone) may increase the regulation cost for other pollutants.

This paper improves the understanding of the relations between nitrogen pollutants and the gains associated to regulating these different pollutants together by a combination of instruments. An integrated economic and agronomic approach is proposed. *AROPAj* is combined with *STICS* in order to capture the soil and crop heterogeneities in fertilizer response (in terms of yields and emissions) as well as to compute nitrogen pollutions. The policy under review combines a tax on nitrogen input and incentives that promote perennial crops (Miscanthus), which is assumed to require low levels of input.

A key point appears regarding the two kinds of pollutants, a soil pollutant (NO_3) and the atmospheric pollutants (NH_3 and N_2O). While there is no significant cost-efficiency gain to promote

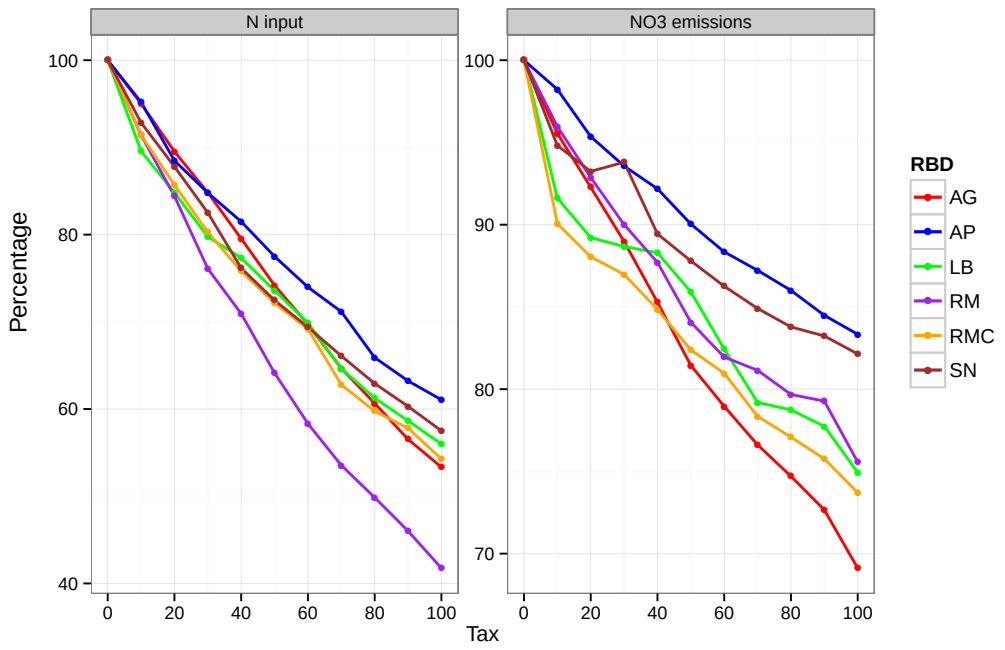


Figure 44. N-Input tax impact on the input consumption and on the pollutant emission: a possible paradoxical effect (on the Seine river basin - SN). The tax represented on the x-axis corresponds to the percent increase in nitrogen price.

the land-use policy for gas pollutants, the policy mix strongly decreases the abatement cost for nitrate pollution.

For instance, for a target of 20 % abatement, the mixed policy decreases the NO_3 abatement cost by 0.8 percentage points (from 2.2% to 1.4%) and does not significantly change the N_2O abatement cost. The explanation lies in the fact that, unlike gas pollutants, NO_3 is very soil-crop dependent and the land reallocation in favour of miscanthus leads to higher additional abatements than those obtained with the N-tax alone. Thus, the land use induced by N-tax may be relatively far from the optimal one.

However, for N_2O and NH_3 emissions, the abatement gain obtained from a mixed policy is low because the land use induced by N-tax is relatively close to the optimal one for gas pollutants. On the basis of the parsimony criterion, the fertilizer tax would be therefore preferred.

Another advantage of a mixed policy is the possibility of choosing the combination of instruments that brings the largest abatement for a given agricultural income loss and maximizes the abatement for a given revenue. In particular, for a given net tax receipt (tax receipt minus crop subsidies) between 0.6 and 0.8 MN, the abatement range is between 7 and 35% in the case of NO_3 and between 16 and 50% in the case of N_2O in France. This is illustrated by Figures **??**.

Moreover, because soils, crops and emissions are different in each region considered, we provide an analysis by river basin. These differences are reflected in the abatement cost curves. Also, we demonstrate that the mixed policy, especially the land-use policy, should be adapted per river basin in order to improve the cost-effectiveness of regulation.

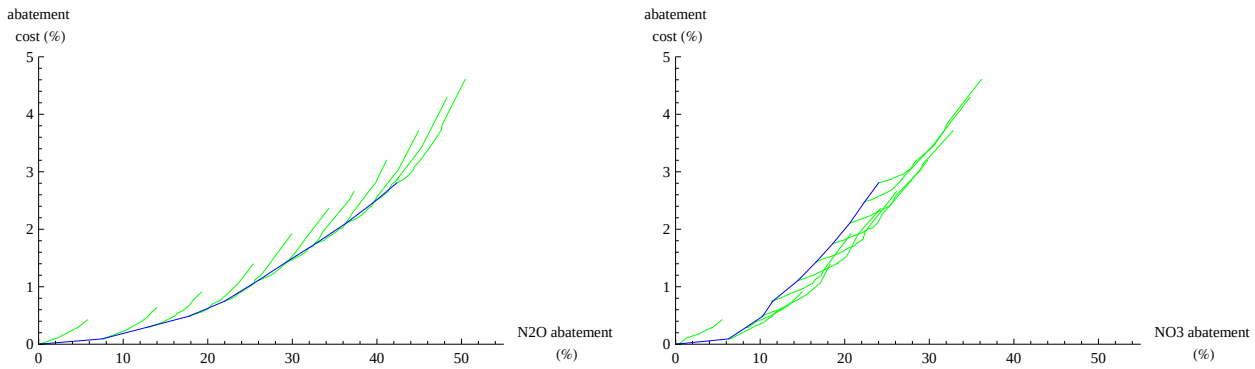


Table 36. Abatement costs to N-losses under NO_3 and N_2O at the French level (source: **?**).

Effect of spatial differentiation of nitrogen taxes

? assess the impacts of the spatial differentiation of input-based nitrogen tax on farmers' gross margins. In theory, this differentiation should lower the compliance costs but the empirical studies on the matter report different magnitudes for the associated economic gains. The assessment's methodology combines a statistical model for nitrate pollution of French water bodies and the *AROPAj* model. Thus, that accounts for the heterogeneity of French farms and the disparities in the spatial distribution of water pollution. A major contribution to the literature on tax differentiation is the inclusion of tax on livestock units as a source of animal manure. The results obtained allow to estimate the potential efficiency gains associated with the spatial differentiation of pollution taxes.

It is argued that the only economically viable input-based taxation scheme addressing nitrogen pollution from agriculture is spatially differentiated, since the estimated losses for firm-specific and water body-specific pollution fees equal respectively 5.8% and 32.5% of the gross margin, whereas the uniform policies studied (at the river-basin district and national level) result in substantial economic losses and the discontinuation of agricultural activity. Scenarios with fine-scale spatial differentiation are associated with lower tax rates in areas subject ot lower environmental pressure which, in turns, results in lower compliance costs for the farmers. Such policies are better targeted and avoid needless restriction on agricultural activity. This is illustrated by maps in figure **??**.

As in previous studies based on the *AROPAj* model and focused on nitrate pollution and input-based taxation, an adverse effect of the taxation is observed in some cases (see also **??**). This adverse effect is due to a crop switching process resulting from farmers' adaptation to taxation since newly adopted crops require less nitrogen but are also less efficient in fixing it and thus, ultimately, are leading to more nitrate pollution than in the business-as-usual case. Even if the extent of this adverse is reported to be rather limited, there are numerous cases of it because of the fine-scale resolution of the study. However, results suggest that input-based taxation should be accompanied by additional regulation limiting the crop switching process.

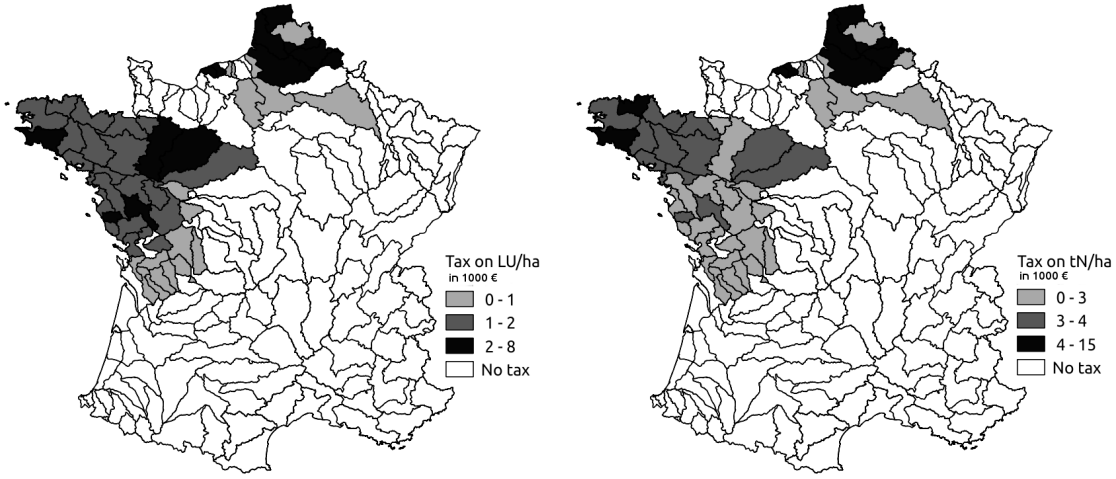


Figure 45. Geographical distribution and rates, in € 1000 per livestock unit (on the left) and per ton of mineral fertilizing nitrogen (on the right), for the "Tax hydrosectors" scenario (source: **?**).

12.4 Biomass provision from perennial crops

Key-points:

Biomass and bio-fuel related issues as yet open questions

Land use impact of perennial crops introduction

The impacts of producing cellulosic biofuels from dedicated energy crops influence agricultural and environmental outputs at local, regional and national scales. Thus, we attempt to assess the trade-offs between food, feed and (bio)fuel production by focusing on the interactions between the cultivation of lignocellulosic biomass, i.e. miscanthus, and its joint implications on agricultural activities and non-CO_2 emissions, i.e. CH_4 and N_2O at the European scale. To this end, miscanthus, which is a annually harvested perennial crop, was integrated into *AROPAj*. Introducing this crop in a one-year period model requires the development of a two-step procedure to feed the model with appropriate information. The first step involves selecting a continuous time-yield function, which is calibrated and adjusted to the yield of traditional crops. The second step is a computation step based on a Faustmann dynamic approach aiming at the estimation of the rotation duration, the average yield and the discounted annual costs.

The average estimated yield of miscanthus among *AROPAj* farm-groups is typically about 23 tDM/ha. Yields between 20 and 30 tDM/ha are very frequent and the highest and lowest yields are 35 and 5 tDM/ha respectively. In order to test the sensitivity of *AROPAj* results to possible variations in miscanthus yield, yield scenarios were evaluated. They correspond to the introduction of miscanthus with 10 levels of possible yield potential. We proceed to a homogeneous reduction of miscanthus yield from 0 to 100% by 10% intervals over all farm groups. The rate of reduced potential is denoted by ψ. This potential reduction is applied at the plot level after the Faustmann optimizing step. We assume that farmers first follow advice in terms of perennial crop management, and then observe yield reduction. Because of the vivid "Food *vs* Fuel" debates, we assume that the part of the farm group's UAA devoted to miscanthus does not exceed 20%.

Table **??** shows that arable land and grassland are reduced at the same rate when miscanthus is progressively introduced. Areas devoted to oilseeds and cereals decrease by 20% and 18% respectively. When the miscanthus potential increases (ψ up to 1), grasslands are slowly but continuously reduced up to 11% (full potential). This is also the case for fallow land, from a 0.5% decrease when the potential is low, to 21% when it is high. The decrease in cropland and grassland areas leads to a decrease in animal and feed production. Indeed, the quantity of on-farm consumed cereals decreases by 4% and marketed concentrates increase by 3% when miscanthus potential is 100%. With a maximum of potential, the computation of the livestock units per hectare of grassland shows an increase of 10%. Given these results, it is evident that a 12% decrease in grassland leads to an intensification of animal rearing.

	without	with miscanthus								
$\psi =$	0	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
miscanthus	0	0.0	0.3	1.4	4.7	8.8	10.8	13.1	14.0	15.3
cereals	38.7	38.7	38.6	38.1	36.6	34.7	33.8	32.7	32.4	31.8
sugarbeet	3.0	3.0	3.0	3.0	3.0	2.8	2.8	2.8	2.7	2.7
oilseeds,proteins	6.0	6.0	6.0	5.9	5.7	5.3	5.1	5.0	4.9	4.9
potatoes	3.6	3.6	3.6	3.6	3.5	3.4	3.3	3.3	3.2	3.1
fodders	11.0	11.0	11.0	11.0	10.5	10.4	10.1	9.9	9.8	9.6
fallows	10.2	10.2	10.2	9.7	9.5	8.8	8.5	8.3	8.2	8.1
grassland	27.4	27.4	27.3	27.2	26.6	25.8	25.5	24.9	24.7	24.5

Table 37. Share of land among major groups of agricultural activities (% UAA) when miscanthus yield potential ψ is increasing - estimates provided by the $AROPAj$ model for the EU-15. (**?**).

Introducing an environmentally-friendly crop has positive effects on GHG emission abatement. The replacement of crops characterized by high nitrogen (N) consumption (e.g. cereals) by miscanthus, which requires low N input, is certainly advantageous for the abatement of N-pollutants, among them N_2O. Results show that the introduction of miscanthus leads to a significant reduction in N_2O losses. The figures rise to 10% when miscanthus reaches its full potential. With regard to CH_4 emissions, the increase in purchased concentrates leads to a reduction, but this reduction (of 2%) remains low because of the intensification of animal rearing.

Sensitivity analysis

In another study, we address the question of biomass uncertainty in continuous time for a perennial resource, in our case miscanthus, in order to analyse the sensitivity of the supply of this crop in France (**?**). By applying two Faustmann models, we derive harvesting rules to deal with the optimal rotation age and the value of miscanthus when its growth function is accounted for in a deterministic way as well as when it is governed by a stochastic process. To manage the random yield of miscanthus, we develop a simple stochastic model in which the yield process is based on a popular distribution, the so-called beta distribution (see Figure **??**). Because of its versatility, this distribution has been used to model a variety of uncertainties. It is bounded on both sides and so it is often used for representing processes with natural lower and upper limits while it has the flexibility to accommodate both positively and negatively skewed data. The $AROPAj$ model is used to highlight the large scale impact of annual yield randomization.

This study provides an analysis of one factor, i.e. yield, which comes into play when the farmer is given the possibility of adopting miscanthus. We highlight the differences between the deterministic yield expectations and stochastic yield expectations of miscanthus in terms of profitability, land-use

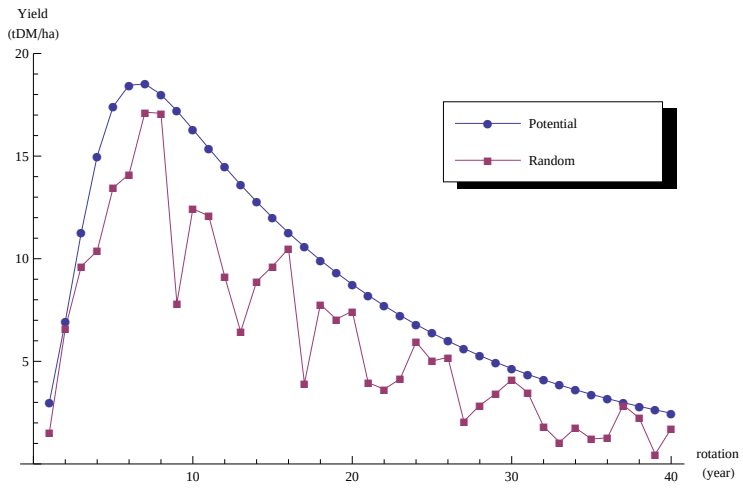


Figure 46. Miscanthus yield distribution: an example of generated yield sample.

allocation and N-losses. Results show that planting this crop requires an important investment which will be returned as of the 4th year in the deterministic case and as of the 10th year in the uncertain case. Concerning the economically optimal rotation age, it is about 16 years in the deterministic case. The introduction of a random factor in the growth function delays the optimal rotation for 5 years, up to 21 years.

In addition to the long-term profitability problem, the farmer has to deal with questions related to land use change. Favourable yield potential incites the farmer to adopt miscanthus despite the high degree of uncertainty. In fact, introducing yield uncertainty delays the adoption of miscanthus until its potential is sufficiently high to motivate the farmer's decision. More specifically, in the deterministic case, miscanthus is adopted when 30% of the yield potential is reached. In the stochastic case, this occurs when 60% of the yield potential is reached.

Simulations show a strong decrease in the food-crop areas when miscanthus is introduced. When the miscanthus yield reaches its full potential, cereals' areas decrease by 25% in the deterministic case, and by 7% in the stochastic one. At the same time, grassland decreases by 13% and 7.5%, in the certain and uncertain cases, respectively. Moreover, abandoned areas decrease to 28% in the first case, and to 14% in the second one. These changes in land use have a substantial effect on the N-input demand and N-losses. Indeed, we notice a decrease of fertilizer demand by 22% in the deterministic case and by 13% in the stochastic case when the miscanthus yield reaches its full potential. This decrease in N-demand leads to a 26% reduction in nitrate losses in the first case and 13% in the second one.

Through a sensitivity analysis, we notice that yields, renewal cycle costs and the discount rate interact with yield randomization and significantly affect the future profitability of miscanthus.

12.5 Bio-economy oriented studies

Key-points:

Opportunities in complex bio-economic model based analysis combining all *AROPAj* modules

Farming systems negatively impact by troposheric ozone

Ozone impacts

The *AROPAj* model was used in a study aiming at estimating the potential detrimental impacts of surface ozone (O_3) on French crops in the context of climate change (**?**). Surface ozone is an air pollutant which causes damage to leaves and reduces photosynthesis efficiency. Its atmospheric concentration is predicted to increase in the future and could significantly impair agricultural productions.

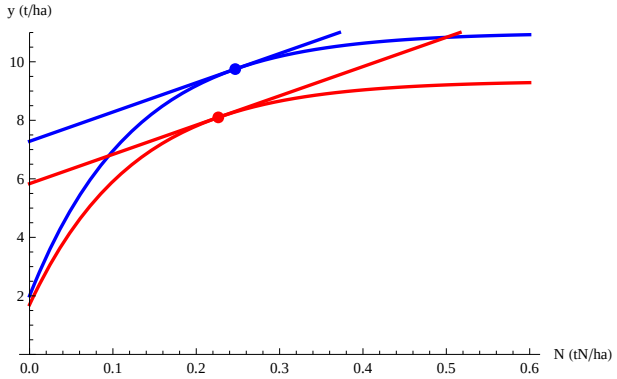


Figure 47. *AROPAj* yield response function (blue) and the impact of ozone (red), in this case a 10% decrease in yield. The economical optimum is affected and the nitrogen quantites are not the same after the impact of ozone. (source **?**)

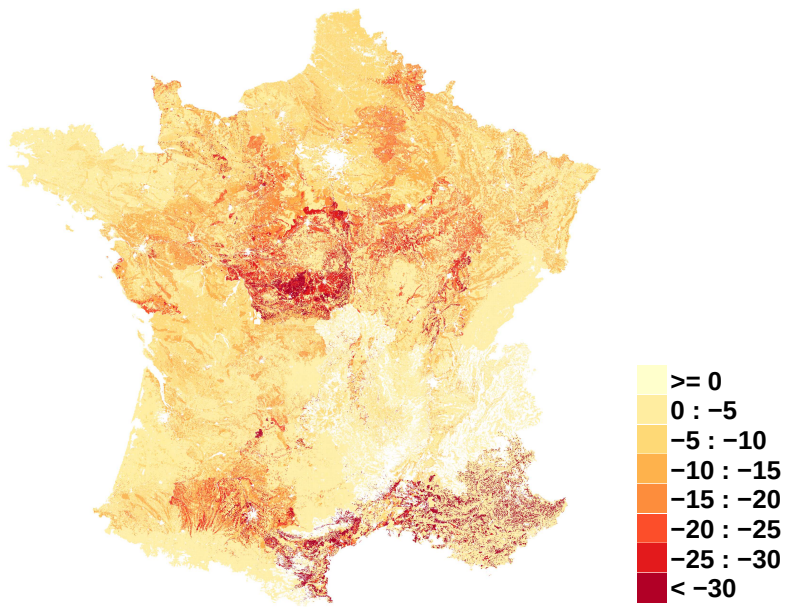


Figure 48. Losses in agricultural gross margin (in %) in France according to the worst case climate scenario. (source **?**)

Since the impact of ozone depends on the considered crop, the developed methodology had to take this heterogeneity into account. To do so, the authors used a literature review compiled

by **?**, who performed regression analyses on more than 700 articles and 19 crops to extract the coefficients of impact as a function of the ozone concentration.

The data for future ozone levels in 2030 were taken from climatic studies and represented in a 50km per 50km grid. The amount of ozone was calculated as the sum of daily values for the three months preceding the harvest date of the crop, as estimated by the *STICS* model, enabling for a fine consideration on both spatial and temporal variability in ozone concentrations. Then for each crop and each cell of the grid, the ozone impact coefficients were applied to the yield response function as presented in Figure **??**.

The impacts on gross margin are significant and can reach more than 5% in some regions based on the worst case climate change scenario, as presented in Figure **??**. The authors also observed a shift towards the use of more ozone resistant crops such as barley for which acreage increases to the detriment of sensitive crops such as wheat.

Impacts of crop inoculation by plant growth promoting rhizobacteria

? assessed the economic and environmental impacts of crop inoculation by plant growth promoting rhizobacteria (PGPR) in the Rhône-Alpes region. The inoculation practice with these beneficial bacteria, including the genus Azospirillum, has been shown to sustainably enhance cereal growth (**?**) and has attracted a lot of attention over the past 40 years. However, effects on crop yield, including maize, are still uncertain. As a result, no economic analysis of the effect of inoculation by Azospirillum has been conducted so far. Moreover, no previous studies have addressed the possible environmental effect of such inoculation, in particular regarding GHG emissions.

For the economic and environmental analysis, we used a combination of *AROPAj* and outputs generated by the crop model *STICS*. The main proxies used to assess the effect on farmers profit were the possible influence of the PGPR on maize yield and its inoculation cost. Since neither previous reports nor our field trials results provided clear and consistent inoculation effect on maize yield, we created a range of randomized realistic scenarios of the possible effect of inoculation on yield using the dose

$$N$$

-response function generated by *STICS*. The environmental analysis of inoculation based on nitrous oxide N_2O emissions from cropland soil was assessed firstly through field trials. Our results show that N_2O production was stimulated by inoculation up to $> 80\%$ at sites with high C content, but was slightly decreased by inoculation down to 8% at sites where C content was lower (**?**). These results were used to elaborate the scenarios for the possible effects of inoculation of soil N_2O emission according to soil type.

In the case of a low soil C content (**??**c) and for a given set of parameters ($N0$ and τ) for the dose-response function (**??**a), N_2O emission was reduced (green line, **??**c) at low fertilization rates. In contrast, N2O emission increased in response to inoculation in the case of a high soil C content (**??**d).

Compared to the benchmark (red line, **??**), gross margin increased on average for maize when the inoculation cost was low, it was close to neutral for a cost of €20/ha, and it dramatically declined for higher costs (up to €20/ha). Similar results were obtained on average for marketed maize production, maize on-farm use and fertilizer demand (**??**) but not for maize area, for which

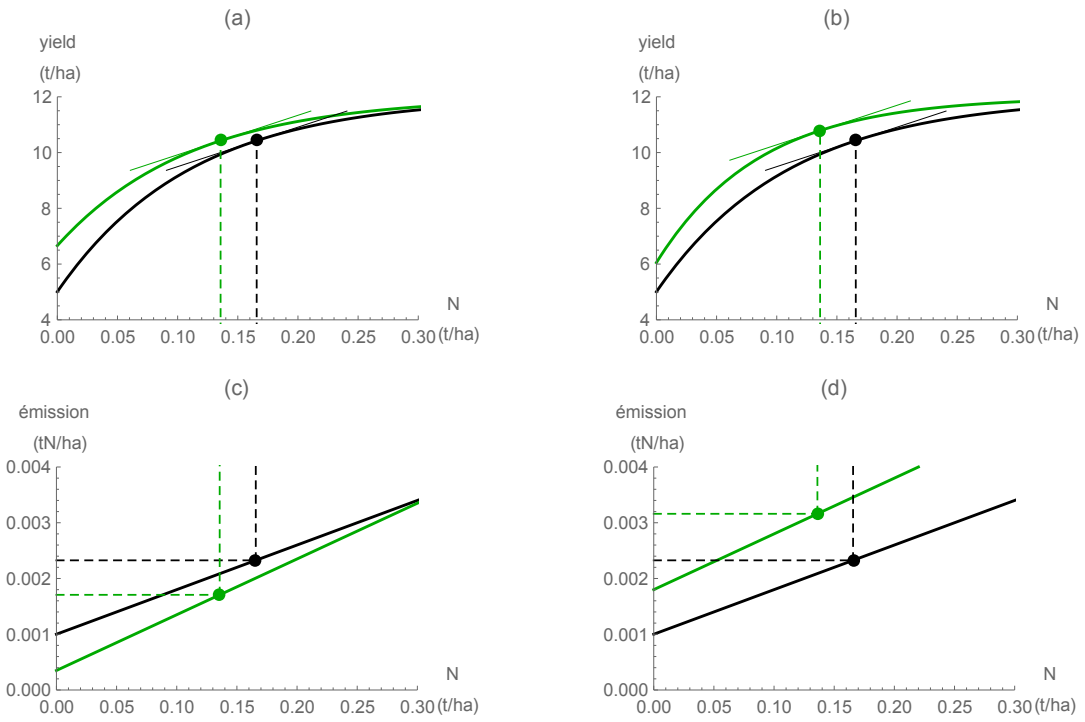


Figure 49. Changes in yield (a, b) and N_2O emission (c, d) in response to maize inoculation for farmland with low (a, c) and high (b, d) soil carbon content. (source ?)

density functions strongly differed with different inoculation costs. In all cases, synthetic fertilizer demand decreased, leading to a decrease in GHG emissions. Furthermore, improved emission abatement in low C content soils was reversed and deteriorated when soil C content increases. In this case, private economic benefits in farming systems were offset by increasing GHG emissions.

12.6 Adaptation to climate change

Key-points:

Combining the crop model capability of refining technical cropping routes and the optimal agricultural land allocation provided by *AROPAj*

Using *AROPAj* to feed spatial econometrical analysis of climate induced land use change

Autonomous adaptation of farming systems

? investigated the impacts of climate change on main European crops, the potential autonomous adaptation of farmers, and the final impacts on European agricultural supply. This is the first study to estimate in a consistent manner what would be the impacts of climate on main European crops and the gross margin and production levels of European farmers, while accounting for a certain type of adaptations. These autonomous adaptations cover actions that rely on routine farmer's knowledge of how climate influences production activities, such as shifting sowing date by up to 3 weeks, using a cultivar of lower precocity, altering the area allocated to various crops, or applying irrigation to previsouly rainfed crops.

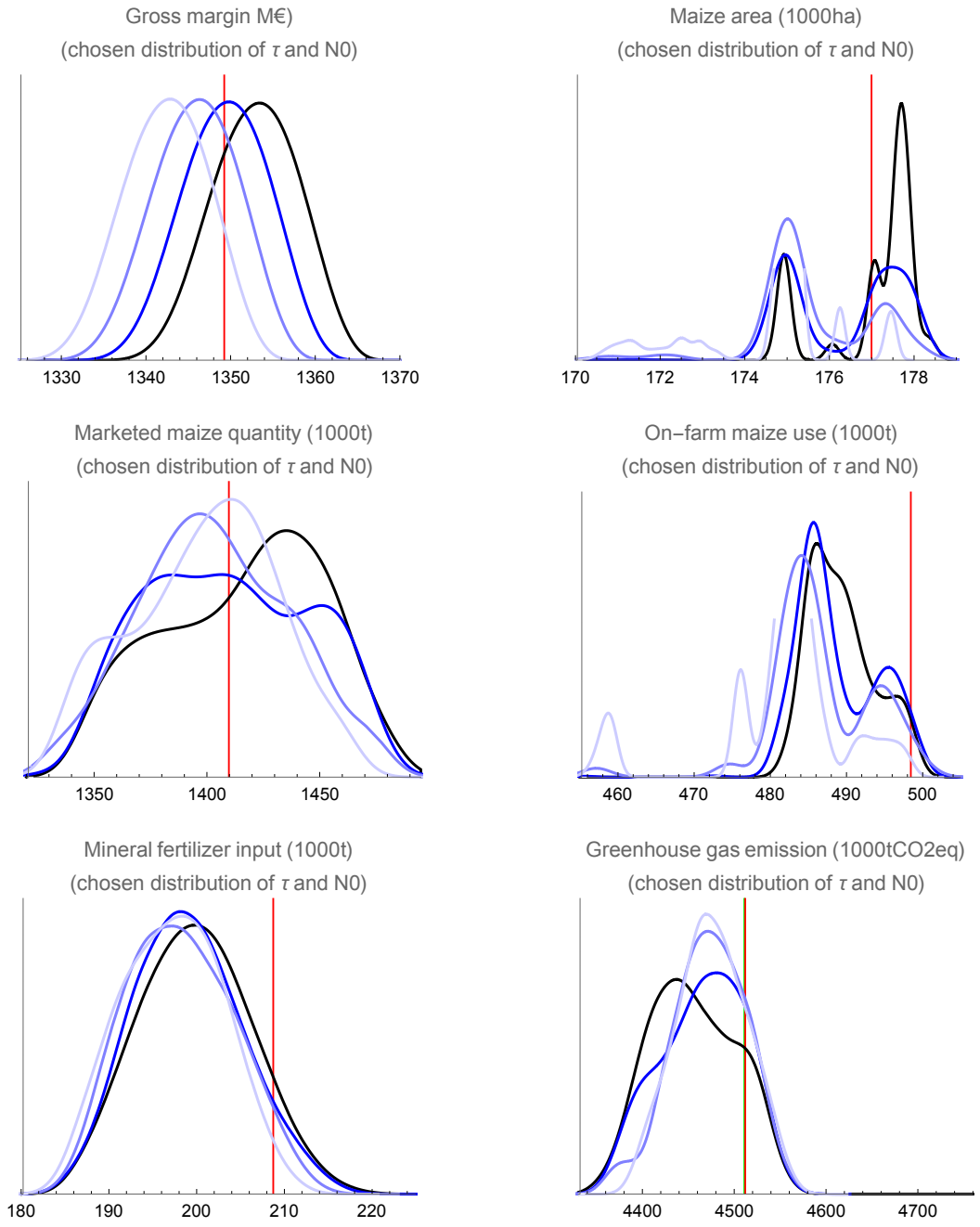


Figure 50. *AROPAj* based probability density function for gross margin, maize area, marketed maize production, maize on-farm use, fertilizer demand and GHG emissions, at the region level. Blue lines correspond to four inoculation costs, from 0€ (dark blue) to 60 €/ha (light blue). The red line represents the non-inoculated maize control case. (source ?)

The methodology builds on the *AROPAj* model (version *V2*) coupling to the *STICS* crop through *N*-to-yield dose-response functions on the *ArTiX* plateform (see section **??** and **??**).

First, we generated a first set of dose-response functions using present climate information at the scale of *EU*-15, as simulated by a regional climate model (*RCA3*, **?**) driven by simulations from the *ECHAM5* global climate model for the time period 1976-2005. The effects of using this set of dose-response functions, as compared to using the initial yield and *N* rates contained in *FADN* data (i.e. no coupling), as well as to using dose-response functions simulated with a more accurate source of weather data for present climate was evaluated in an article under review **?**.

In a second step we simulated four new sets of dose-response functions under future climate with the *STICS* crop model using the same source of climate data (*RCA3* regional climate model driven by *ECHAM5* global climate model) for the time period 2071-2100. These four sets of dose-response functions differed in terms of the assumed radiative forcing level (*SRES* scenario *A2* or *B1*) as well as by virtue of wether or not farmers used an alternative sowing date or cultivar. These four sets thus govern the effects of autonomous farmers' adaptation. An additional scenario, which permits for the possible land reallocation in the *AROPAj* model, further allowed us to estimate the impacts of crop substitution. This was achieved by introducing additional parameters (`xjk0(j,k)`, see **??**) referring to the land allocation provided by a reference scenario, and additional constraints (`fx1(j,k)`, `fx2(j,k)`, see **??**).

Overall, we found results similar to previous literature for climate change impacts without adaptation: aggregated impacts over Europe range from -13% to +15% depending on the crop, with high spatial variability and Nothern (resp. Southern) Europe being more positively (resp. negatively) impacted. We found that autonomous adaptation - and in particular adjustments in field scale management practices - could have large benefitial impact: for instance, accouting for them would change the impact of climate change on aggregated European cereal production from moderately positive (+5-13%) to highly positive (+53-55%). This would, however, entail the high cost of a sharp increase in water consumption, in particular because autonomous adaptation increases crop yields without increasing their water-use efficiency. Our results show that autonomous adaptation is very important to account for when estimating climate change impacts, and is likely to trigger price feedbacks from markets as well as high pressure on water resources.

Grass monetary value and its sensitivity to climate change

Grassland (e.g. **?**) and livestock (e.g. **?**) systems have a essential role in both ecosystem services and agricultural economy. Livestock (**?**) as well as grassland (**?**) are impacted by climate change and the overall impacts are region-dependent (**?**). Implementation of practices for climate change adaptation and mitigation could reduce climate change damage to agricultural systems (e.g. **??**) and also open up the possibility of increasing the resilience of agricultural systems.

A large proportion of produced forage and grass is not marketed and therefore their economic impacts are not considered in farm costs. To overcome this issue, the economic value of grass production has been calculated, across fifteen member states of the European Union (*EU*-15), using the *AROPAj* agro-economic, which then has been coupled with *PaSim* biogeochemical grassland model to address both economic and environmental evaluation of European farming systems, taking into consideration the potential future impacts caused by climate change till 2100 (See section **??**).

The shadow price of five types of grass and fodders implemented in *AROPAj* has been calcu-

lated. Our results show that grass shadow value depend significantly on climatic trends (See Table **??**). According to our results, grass cover increase to a greater extent under *B1H2* scenario, which may rely on a small increase in atmospheric CO_2, a slight increase in the temperature, along with precipitation levels substantially higher than *A2H2* scenario (**?**). However, applying farm-scale adaptation practices (*A2H2-A* and *B1H2-A*) resulted in an important increase in cropland cover, at the expense of grassland (see Fig. **??**). As a result of these changes, grass shadow prices will increase more significantly under the two adaptation scenarios due to the potential decrease in grassland area (see Table **??**).

Table 38. Economic value of five varieties of grass under different climatic scenarios (expressed in €/t) and the relative shadow price difference (expressed in %) between two future adaptation scenarios (source:**?**).

	Shadow prices (€/t)					Europe	
	CTL	A2H2	A2H2-A	B1H2	B1H2-A	A2H2-A	B1H2-A
Beef forage	3.9	4.0	4.0	3.8	3.9	0.6	2.6
Maize forage	19.0	18.9	19.5	17.6	18.5	3.0	5.4
Permanent grassland	14.3	14.1	15.0	13.0	14.2	6.8	9.8
Temporary grassland	7.7	7.7	8.2	6.9	7.5	5.7	8.5
Other meadows	10.2	9.8	10.3	8.4	9.1	5.2	8.5

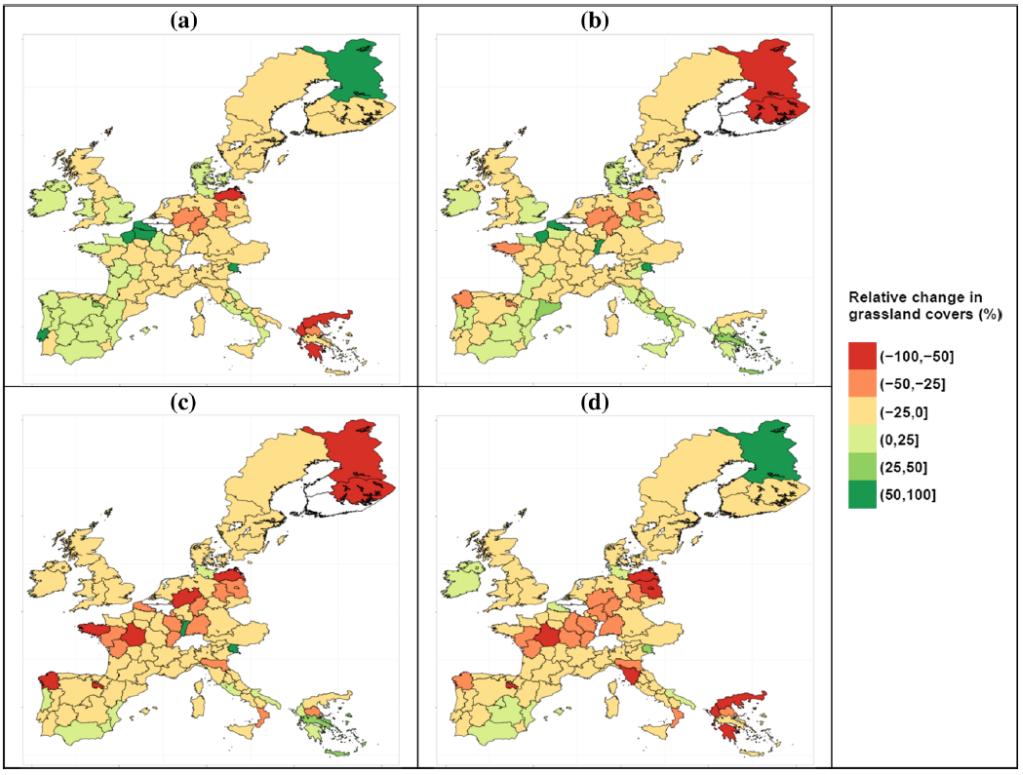


Figure 51. Relative variation in land covered by grass (expressed as a %) across the EU under future scenarios compared to the present (CTL) a. A2H2, b. B1H2, c. A2H2-A and d. B1H2-A (source:**?**).

Adaptation and mitigation of climate change in agriculture: accounting for possible land use changes

Following the study by **?**, we go a step further in the possible adaptation measures to climate change. Since climate change is going to modify future agriculture incomes, it will also impact the returns to agricultural land and the spatial equilibrium between land based economic sectors such as agriculture and forestry. In order to account for these effects, we developed a methodology where the results of two sector-specific models (AROPAj for agriculture and FFSM++ for forestry) are used into a spatial econometric land use shares model (**?**). We investigate then the effects of climate change adaptation of farmers on land use along with the effects of a mitigation policy (tax on greenhouse gas emissions from agriculture). We find that the mitigation costs for farmers are lower when the land use feedback is accounted for (**?**).

The land use shares model is based on the hypothesis that landowners maximize the income of their land or, in other words, their land rent. We model the following land uses: i) agriculture and pasture, ii) forest, iii) urban, and iv) other land uses. For each land use we have an explanatory variable or variables to proxy their respective land rent. Agricultural land and pasture are approximated by the land shadow price estimated by AROPAj. Since we cannot specifically distinguish land rent for crops and land rent for pastures, in the application to climate change of this methodology, we model crops and pasture as one land use in the econometric part, and distinguish them later on the basis of their respective share in the production mix of farmers. Returns to forestry are obtained by the partial-equilibrium forestry model FFSM++ (**??**). The proxy for the urban rent is the population density and incomes.

The land use shares model is also accounting for the spatial spillovers through the error terms and also through the explanatory variables. We simulate the effects of climate change by modifying the land rents for each of the modeled land uses. For the urban land use, we use the projected demographic evolution by IIASA downscaled to France based on the French national statistical institute (INSEE) projections. The impact of climate change on forestry returns has been captured through two indicators: tree growth and tree mortality. Also, forest manager are allowed to change tree species if another specie is more profitable under future climate conditions.

Since we are building on the results from **?**, we simulate the A2 and B1 climate change scenarios and the baseline, CTL. We also introduce a tax on greenhouse gas emissions with values from 0 to 200€/tCO_2 eq. Figure **??** presents the land use changes resulting from climate change and/or the mitigation policy. Under current climatic conditions (CTL scenario), the mitigation tax results in an increase in forest and pastures area at the expense of agricultural crops. Agricultural crop area is nevertheless increasing under both climate change scenarios although this tendency is somewhat restrained when greenhouse gas emissions taxes are applied.

Figure **??** presents the obtained results in terms of greenhouse gas emissions and tax levels for the three climate scenarios. The tax impacts agricultural profits and, as expected, agricultural crops area is decreasing when taxation applies. Thus, the tax has both an impact on the intensive margin (lower emissions levels per hectare), and on the extensive margin (less area in agricultural activity). The effect on the extensive margin is given in figure **??** by the gap between the two emission curves with and without the land use change feedback effect.

Greenhouse gas emissions from French agriculture are to be reduced by 12% in 2028 comparing

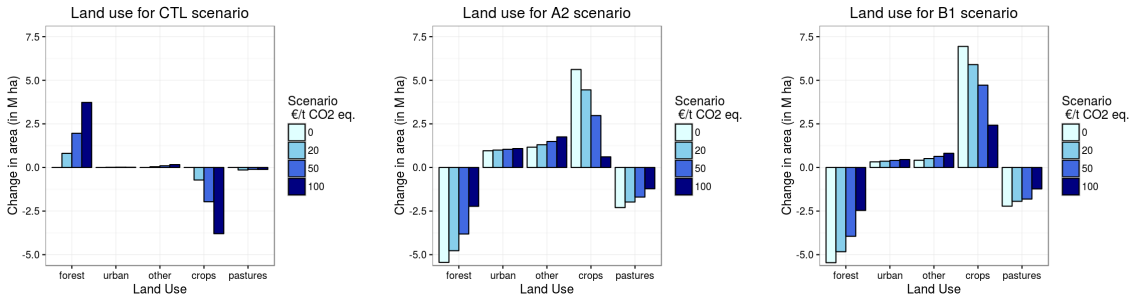


Figure 52. Land use change for the three climate change scenarios and for different greenhouse gas tax levels (source **?**)

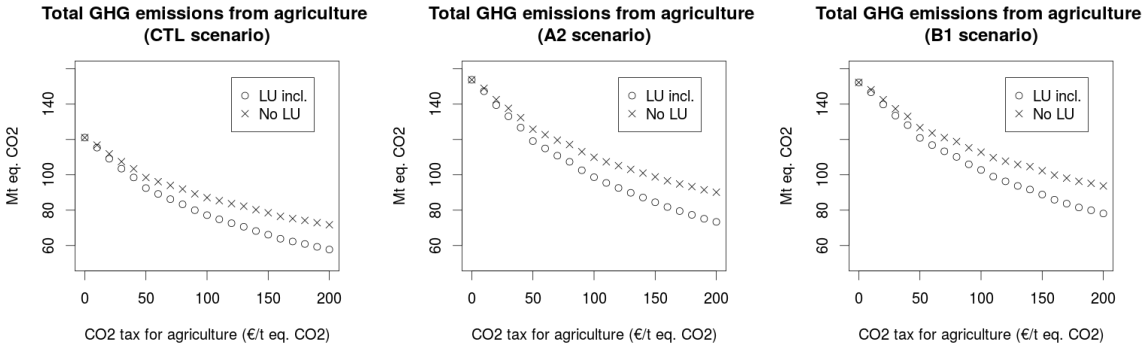


Figure 53. Greenhouse gas emissions for different levels of the mitigation tax (source **?**)

to 2013 levels. This objective is attained at 30€/tCO_2 eq. in the CTL case (40€/tCO_2 eq. when land use feedback is ignored); 90€/tCO_2 eq. in the A2 case (120€/tCO_2 eq. when land use feedback is ignored); and 90€/tCO_2 eq. in the B1 case (130€/tCO_2 eq. when land use feedback is ignored). We can thus conclude that climate change potentially would impact favorably agricultural profits, but can have negative impacts in terms of greenhouse gas emissions (and other pollutants associated with intensive agriculture such as nitrates), and can result in the loss of permanent grasslands and forests.

13 Perspectives and further developments

Key-point:

Some challenging issues under study

This section deals with recently initiated or discussed research on *AROPAj* capabilities. That results from demand expressed by scientific or societal partners, or from researchers' initiatives.

In 2015, the most important challenge involves the water issues of increasing and competing agricultural demand for water, and moreover, environmental issues and quality problems. The increase in demand would result from climate change and climate change adaptation, when agriculture will face water stress conditions more often and with greater intensity (see the caveats and the conclusions following the paper of **?**). Environmental issues related to water have been approached, mainly with regards to nitrates, and their analysis could be improved when accounting for water scarcity. Both qualitative and quantitative water issues will be key points in the future.

Questions about organic and "low input" farming systems are also very often adressed. There is a possible discrepancy between what consumers and policy makers expect for inputs and yields and what organic farmers are yielding. Indeed statistics are limited. We need to pay attention to the risky extrapolation of the favorable figures due to the management provided by pro-active, motivated and well-qualified bio-farmers to traditional growers and breeders converted in bio-farmers. Nevertheless, a prospective analysis demands that we be innovative in terms of scenarios and potential changes so that we evaluate even the optimistic of situation.

Many other issues are put on the table as well. *AROPAj* could be coupled with other bio-physical models to account for phosphorus, pesticides, and other major inputs that acts as major pollutants. At present, there is a large gap between what we require as information and data and what is provided. Moreover, it would be necessary to deal with other models than those already experimented.

Finally, we must develop a way to validate not only the *AROPAj* model, but also the entire modeling chains. This critical point is usually mentioned by articles' readers and addressed by potential users. Nevertheless, we must remember that *AROPAj*, as a normative, economic and prospective model, cannot be validated through repeated observations in controlled conditions as it is done with physical, chemical or biological models. The calibration step is a part of the response to these issues. Publication of peer-reviewed papers and comparison of models are another part. Finally, the improvement of the model by including broader phenomena and processes allows us to "check with reality". We present below some future extensions of the model.

13.1 Water issues

The adaptation of agriculture to climate change and growing demand for water by a range of users entails the question of water availability and pricing. A large part of renewable water (from rainfalls) used by agriculture circulates outside the market. Irrigation is and will become an even more pressing problem. But even in this case, costs related to water extraction (pumping and other investment or variable costs) remain higher than those related with the resource itself.

An innovative mix of crop modeling and economic analysis is under progress with the aim of providing two-input yield response functions. These two inputs are nitrogen (through mineral fertilizers or manure sourced organic matter) and water. Water is taken into consideration as the command variable from the farmer's economic point of view. Thus, the producer may call for irrigation when rainfalls do not provide water in sufficient quantities in terms of the potential costs and benefits of additional water.

$$y(W,N) = (1-e^{-\sigma(W_0+W)})\,(1-e^{-\tau(N_0+N)})\,\phi$$

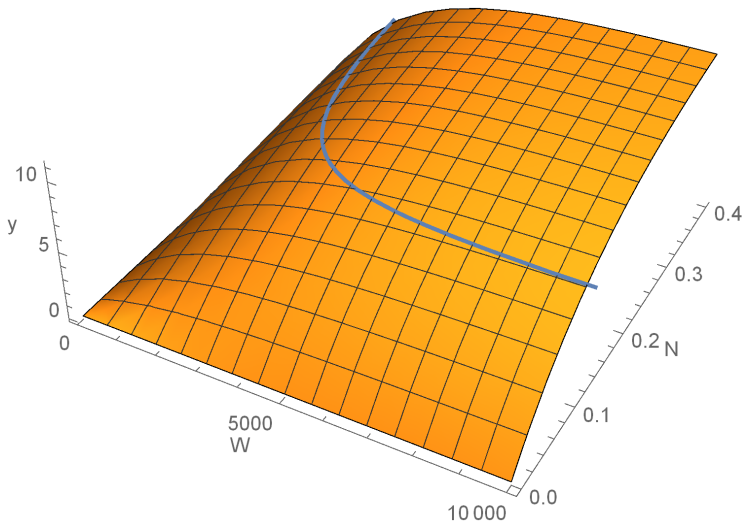


Figure 54. The yield function as annual response to water and nitrogen inputs; water and nitrogen refer to the total annual quantities including natural resource; the blue line refers to an isoquant in yields.

The link between the crop model sourced functions and the economic analysis is complex. The demand for water (expressed in cubic meters or *mm* height per hectare, 1*mm* being equivalent to 10 cubic meters) is translated in water stress for the crop model simulation. For one period, the two-input function is illustrated by Figure **??**. The method of response function building follows the method initiated by **?** and expanded by the E. Galko's and D. Leclère's PhD work. The two-input approach is much more complex to develop, but difficulties are compensated by database availability and increased computing capabilities. It is implemented and explored in the P. Humblot's PhD thesis (May 2016). Methodological aspects and implementation are described by **?**.

13.2 Low input agriculture

The first attempt to analyze the performances of low input farming systems was carried out at a small watershed level, at which only a few organic farming systems currently exist. Economic performance was assessed through the standard criterion of gross margin. The environmental performance was judged through subsoil nitrogen losses. However, this kind of experimental approach suffers from a lack of reference points of comparison with traditional farming systems.

At this point, the first difficulties which arise concern the definition of organic farming, based on product labeling and farming practices. This indeterminacy of definition may be reflected in

existing databases. Secondly, the European economic databases, such as the *FADN*, still provide too little information for parameter estimation.

As for climate change and other long-term issues, the use of models might provide very useful tools for overcoming the lack of experiments and observations. There is room and demand for the combination of validated bio-physical models and economic analysis, for instance, in the case of using large amount of organic matter as nitrogen intakes for crops. Such possibilities were tested by the *Carbo-PRO* and *STICS* based analyses (see Section **??**). Nevertheless, in this case, organic sourced nitrogen acts as a supplement for mineral fertilizers. The analysis could be completed with simulations that exclude mineral nitrogen intakes, but bio-farming entails more than organic nitrogen sources. It refers also to other restrictive practices which need statistical surveying. Renewed databases and model based results remain a limiting but exciting factor for progress in this direction.

13.3 Facilitating *MP* solutions involving interacting agents

The *AROPAj* model is based on individual mathematical programming *LP* which are solved independently of each other as long as individual activities and constraints are separated (in other words in case of "orthogonal" individual matrices). When considering that resources are shared by several agents, there is possibility to find the aggregated optimal solution of the aggregated mathematical program, for instance by trial-and-error process until dual values related to the resource constraints are equal. Nevertheless in many cases such processes are of quite high computation cost.

Software are now able to solve *MP* problems involving huge size matrices. Even when considering that computation time increases with matrix size much more than linearly, there is strong potential gain to solve directly this kind of mathematical programs. The challenge is to preserve the benefit of existing "matrix generator generator" and to keep operating applications developed around the *AROPAj* model. The locking point is the indexation of farm groups' elements, which is reset for each sub-country.[14]

There is now the possibility to aggregate farm group matrices emanating from several *EU* countries, up to the *EU* at a whole. This is of interest when *MP* constraints involve activities from different farm groups. We define a specific $\boldsymbol{T}$-block operating in the kernel of the model. (see Section **??**) The particular design of such constraints is that the related *RHS* elements are not farm group-indexed. Examples of this are targeted limits in agricultural production (e.g. minimal food-calorie amount), or environmental issue (e.g. maximal amount of emitted pollutants). There is a real gain by doing so when the problem of interest implies at least two kinds of not linearly correlated trans-farm group constraints. Otherwise, the standard individual *MP* should be better solved for all involved farm groups when parameterizing the "price" of the bounded resource. But even in this case, there is the advantage of an easier implementation of the "dual" approach, which directly provides the dual price of a bounded resource utilized by several economic agents (implicitly clearing a quota market).

A processing application is encoded in *FORTRAN* reseting automatically the list of suffices and the indexation of farm group-indexed parameters, independently from member-States involved in

[14]Indeed the point arises when one *EU* member State is subdivided into two or more sub-countries, which is automatically performed in the farm group clustering step for countries with large sample size *FADN* database.

a given calculation. By doing so there is no need to change the kernel of the model (i.e. no required re-writing of basic blocks of *MP* activities and constraints). The counterpart is the need to manage the complete *MP* solution provided by *GAMS*. Nevertheless the *MP* result exploitation is facilitated by indexation specifying the sub-country and the farm group within the sub-country.

This application has been tested in solving the *MP* made of all *EU* farm groups in the case of three trans-group constraints. Solution is exhibited in reasonably required computer time and space limits on our multi-core workstation.

13.4 Anchoring the use of the *STICS* crop model and the calculation of probability farm group location to the core of *AROPAj* computing applications

The two aspects presented here are run-time costly, and consequently require friendly-user and repeatedly utilized applications which need to be carefully carried out.

Building dose-response functions

The production of input-yield functions for a wide set of crops and farm groups in different climatic scenarios mobilizes computation resources when dealing with million runs of the *STICS* crop model. This is two-fold dimension problem referring first to databases and second to the use of the crop model by itself. Computing interfaces are developed to manage the different steps leading to *STICS* run management, yield function adjustment and selection in present time and in future climate context (see the section **??**).

Applications made of *Python* encoded programs were tested for the *V5-AROPAj* version applied to France and the 2009-*FADN*. Refining and running the applications is conducted for the EU and *FADN*-years 2012-2018. Yield functions are provided as well as nitrogen-loss function, losses referring to the three pollutants N_2O, NO_3 and NH_3.

Mapping farm groups at an infra-regional resolution level

The method presented in Section **??** is now tested for the whole of the *EU* and for the 6 *FADN*-years of the *V5-AROPAj* version. The link to the menu accessing the model use and surrounding computer applications is now validated (but reserved to the model's administrator). Applications are encoded in *CShell*, *R* and *GAMS* scripts.

Issues still pending refer to different kind of aspects:

- from the region list and to crop location priors in all *FADN*-year cases: a few regions are not properly referred to basic spatial information, e.g. 580-Canarias (Spain), 650-Açores e Madeira (Portugal), extra-European regions (205-Guadeloupe, 206-Martinique, 207-La Réunion), German "State-towns" (20-Hamburg, 40-Bremen, 110-Berlin)
- from the region list and to crop location priors for one *FADN* year in case of region merging in *EC* statistics: 610-Entre Douro e Minho/Beira litoral and 620-Tras-os-Montes/Beira interior merging into 620-Norte e Centro (Portugal, 2007); 761, 762 and 763 merging into 768-Dunántúl and 760, 765 and 766 merging into 768-Alföld (Hungary, 2012)
- from crop location priors to crop location posteriors: all regions in Germany, except 100-Saarland (10, 30, 50, 60, 70, 80, 90, 112, 113, 114, 116; all years); regions 10 and 20 merging into region 15 for 2018

- from crop location posteriors to farm group location: 204-Corse (France, all years), 221-Aosta (Italy, all years), all regions in the United Kingdom (411, 412, 413, 421, 431, 441; all years), 380-Ireland (all years), 100-Saarland (Germany, 2008-2012)

Given the region and the year, one step failure means that the farm group probability set is empty for this region and this year.

Managing some of these issues is still in progress in the encoded applications. At the present time, March 2018, between 24 and 30 regions are missing in terms of farm group location probabilities (see Table **??**).

14 Glossary, communication, references

Glossary

A2, A2H2, B1 refer to *SRES* (now *RCP*) proposed by the *IPCC* (*AR4*)

ADEME the French Environment and Energy Management Agency

AR4, AR5 Assessment Report (*AR4* and *AR5* standing respectively for the 4th and 5th *IPCC* reports)

ArcGIS geographical information system software, formerly used in some cases

AGRI4CAST refers to interpolated meteorological data provided by the *JRC MARS* unit.

AROPAj acronym of the agro-bio-economic model ("Agriculture, recomposition de l'offre et politique agricole", "*j*" accounting for mnemonics)

AROPAj version refers to one or several yearly *FADN* sample(s), characterized by specific farming systems clustering, parameter estimating and *LP*s calibrating

ArTiX set of applications and database dedicated to *STICS* run management and designed to provide *AROPAj* with N and H_2O to yield functions

awk is an interpreted programming language designed for data and text processing (*sed* is another programming language widely used by the *AROPAj* shell)

BSA burden sharing agreement (refers to the European climate policy)

BtL biomass to liquid (thermo-chemical process)

CAP the Common Agricultural Policy (applying to all *MS* of the *EU*)

CAPRI refers to the Common Agricultural Policy Regionalised Impact model, funded by several *EU* projects

Carbo-PRO models the links between soil carbon storage and organic matter application, developed by *INRA* and the Veolia-VERI company

CC climate change

CE cross entropy

CH_4 methane

CLC the CORINE Land Cover database

CMO Common Market Organization (refers to the *CAP*)

COP stands for cereals, oilseeds and proteins crops

DEM Digital Elevation Model

DM dry matter

dual variable refers to the implicit variable related to a constraint in a mathematical programming model (also: Lagrange multiplier, see also *primal variable*)

EC the European Commission

ECHAM5 one of the *GCM*s

EFOM Energy Flow Optimization Model

EPIC a crop model: http://blackland.tamu.edu/models/epic/

ESDB the European Soil Database

ESIM "European SImulation Model" (developed by the ERS-USDA and used by the Institute for Prospective Technological Studies, IPTS -Seville)

ETS the (European) emission trading scheme

EU the European Union

EU expected utility

Eurostat is the statistical office of the *EU* situated in Luxembourg

EUROTOOLS research projet funded by the *EC* 5th FrameWork program

E-V refers to mean-variance models ("mean" or "esperance")

FADN the farm accounting data network, based on national surveys managed by all *EU* member States, and managed bu the *EC*

FAME Fatty Acid Methyl Ester

FAO Food and Agriculture Organization of the United Nations

FAOStat the statistic division of the *FAO*

farm group refers to *AROPAj* clusters of *FADN* surveyed farms

farm type first refers to the *FADN* classification of the surveyed farms; it belongs to clustering criteria providing the *AROPAj* farm groups; it may be used instead of *farm group*

FT see *farm type* or *OTEX*

GAMS software used for implementation and solving of optimization programs

GCE generalized cross-entropy

GCM General Circulation Models or Global Climate Models

GDAL open source translator library for raster and vector geospatial data

GDP gross domestic product

GENEDEC acronym of one of research framework programmes granted by the *EC*, which supported many of *AROPAj* developments

GeoMIRET name of the *TIMES* model developed by *Ifpen* (acronym from "GEOgraphical Model Integration for (Renewable) Energies and Transport")

GHG greenhouse gas, knowing that major agriculture sourced *GHG* are N_2O (nitrous oxide) and CH_4 (methane)

GIRAF model of the European biofuel transformation sector

GIS geographical information system

GLOBIOM partial equilibrium model focusing on agriculture and forest land uses developed at *IIASA*

GME generalized maximum entropy

GNU Unix-like computer operating system intended to be Unix-compatible. The acronym stands for "GNU's Not Unix"

GPL The General Public License designed to deal with *GNU* products

GTOPO30 a US geographical survey model used for *DEM*

GWC green waste compost

GWP global warming potential, defined for any GHG in terms of CO_2 equivalent given the time horizon

HVO hydrotreated vegetable oil

IEA-ETSAP International Energy Agency - Energy Technology Systems Analysis Programme

Ifpen formerly *Ifp* as Institut FranÃ§ais du PÃ©trole, adding now new and renewable energy sources in its working area

IFPRI International Food Policy Research Institute (based in Wahington)

IIASA International Institute for Applied Systems Analysis (Austria)

IMPACT International Model for Policy Analysis of Agricultural Commodities and Trade (model conceived by *IFPRI*)

INRA Institut National de la Recherche Agronomique, becomes *INRAE* in 2020

INRAE Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (see *INRA*)

INSEA "INtegrated Sink Enhancement Assessment", a shared-cost action financed through the 6th Framework Program of the European Commission.

IPCC International Panel for Climate Change

JRC Joint Research Center (Ispra, Italy; Seville, Spain), related to the European Commission

LAI leaf area index

Linux *OS* supporting the *AROPAj* model use and development (from 2014)

LP linear programming, or linear program

LS least squares

LS livestock

LU (or *LSU*) livestock Unit: used for animal accounting based on *FADN* equivalence factors (the parameter `uii(i)` in *AROPAj*)

LUCAS refers to point land use data (Land Use/Cover Area frame Statistical Survey), developed by the *JRC*

MACC marginal abatement cost curve

MARKAL stands for "MARKet ALlocation computer model generator"

MARS unit of the *JRC* in charge of climate data

ME maximum entropy

MIP mixed integer programming

MNL refers to the "multinomial logit" statistical model

MODCOU hdrogeological model devoted to soil-aquifer transfers and related to the Seine river bassin

MP mathematical programming

MS Member States (of the *EU*)

N nitrogen

N_2O nitrous oxide

nexBtL a renewable diesel fuel production commercial process, see also *BtL*

NH_3 ammonia

NLP non linear programming

NO_3 nitrate

NUTS Nomenclature of Units for Territorial Statistics, refers to the European territorial classification

O_3 ozone

objective function refers to the function to be maximized (or minimized) in mathematical programming

OLS ordinary least squares

OM organic matter

OS operating system (*AROPAj* was developed under *Unix* until 2013, then *Linux*)

OTEX refers to the French "orientation technico-Ã¯Â¿Â½conomique de l'exploitation" (see also *FT*)

OURSE the energy demand model developed by *Ifpen*

OURSEurope the energy demand model developed for Europe and derived from *OURSE*

PAM refers to the "Partitioning Around Medoids" algorithm

PaSim process-based grassland biogeochemical model (https://www1.clermont.inra.fr/urep/modeles/pasim.htm)

PDIE proteins digestible by the small intestine due to the energy provided by feed

PDIN proteins digestible by the small intestine due to the nitrogen provided by feed

PET potential evapo-transpiration

PMP positive mathematical programming

POSTGreSQL software for data management

POSTGIS allows for interfacing *POSTgres* and geographical information system procedures

primal variable refers to an explicit variable of a mathematical programming model

psql query langage interface for working with *POSTGreSQL*

R a *GNU* programming language and software environment for statistical computing

RCA3 one of the Regional Climate Models (*RCM*)

RCM regional climate models

RCP representative concentration pathway, built by the *IPCC* for its 5th assessment report

RHS right-hand side (stands for *LP*)

Roth-C model to simulate change in organic matter stock on the long term (explanation about *RothC* and *Carbo-Pro*)

ROW stands for "Rest of the World"

RUMINANT a dynamic model simulating the process of ruminant digestion

sed see *awk*

SFP single farm payment

shadow price refers to a mathematical programming constraint and to the optimal value of the limited resource related to this constraint (also shadow cost, implicit price or cost, dual value, Lagrange multiplier)

SimU simulation unit (refers to spatial modeling tools)

SRES special report on emissions scenarios, shared into 4 scenario families (i.e. *A-B* $\times$ *1-2*) developed by the *IPCC* for its 4th report

SSP stands for Shared Socio-economic Pathway (related to the *IPCC*' *AR5*)

STICS crop model, access *STICS*

TIMES stands for "The Integrated *MARKAL-EFOM* System"

UNFCCC United Nations Framework Convention on Climate Change: here access the *UNFCCC* site (provides numerous data on *GHG* emissions per country per year per sector)

UAA utilized agricultural area

UFL unit of energy content in fodder for dairy herds

UFV unit of energy content in fodder for non dairy herds

Unix *OS* under which *AROPAj* was formerly developed

UOR urban organic residues

USS urban sewage sludge

UZ unsaturated zone (hydrogeology)

WLS weighted least squares

Communication

Different elements are provided. In addition to scientific papers in peer-reviewed journals (see a part of references below for scientific papers related to *AROPAj*), communicative papers have also been developed. They are of different types, as developed below:

1. Gérard, M. & Jayet, P. (2023). European farmers' response to crop residue prices and implications for bioenergy policies. *Energy Policy.* https://doi.org/10.1016/j.enpol.2023.113561
2. Devaraju, N., Prudhomme, R., Lungarska, A., Wang, X., Yin, Z., de Noblet-Ducoudré, N., Chakir, R., Jayet, P.-A., Brunelle, T., Viovy, N., De Palma, A., Gonzalez, R., & Ciais, P. (2022). Quantifying the benefits of reducing synthetic nitrogen application policy on ecosystem carbon sequestration and biodiversity. *Scientific Reports 2022 12:1*, 12(1), 1–15
3. Prudhomme, R., Chakir, R., Lungarska, A., Brunelle, T., Devaraju, N., de Noblet, N., Jayet, P.-A., De Cara, S., & Bureau, J.-C. (2022). Food, climate and biodiversity: a trilemma of mineral nitrogen use in European agriculture. *Review of Agricultural, Food and Environmental Studies*
4. Gossiaux, E. & Jayet, P.-A. (2021). Exploring the potential for calories and bioenergy in france. *BE-Sustainable Magazine*, 12, 11–17. https://issuu.com/besustainablemagazine/docs/be-sustainable_magazine_issue_12_-_april_2021/s/12150135
5. Bamière, L., Jayet, P.-A., Kahindo, S., & Martin, E. (2021). Carbon sequestration in french agricultural soils: a spatial economic evaluation. *Agricultural Economics.* doi: https://doi.org/10.1111/agec.12619
6. Isbasoiu, A., Jayet, P.-A., & De Cara, S. (2020). Increasing food production and mitigating agricultural greenhouse gas emissions in the european union: impacts of carbon pricing and calorie production targeting. *Environmental Economics and Policy Studies*, (pp. 32p). doi: https://doi.org/10.1007/s10018-020-00293-4
7. Barberis, D., Humblot, P., Chiadmi, I., Jayet, P.-A., Lungarska, A., & Ollier, M. (2020). Climate change and irrigation water: Should the north/south hierarchy of impacts on agricultural systems be reconsidered? *Environmental Modeling and Assessment.* doi: https://doi.org/10.1007/s10666-020-09724-8
8. Jayet, P.-A., Isbasoiu, A., & De Cara, S. (2020). Slaughter cattle to secure food calories and reduce agricultural greenhouse gas emissions? some prospective estimates for france. *Review of Agricultural, Food and Environmental Studies*, (pp. 24p). doi: https://doi.org/10.1007/s41130-020-00117-9
9. Ben Fradj, N., Jayet, P.-A., Rozakis, S., Georganta, E., & Jedrejek, A. (2020). Contribution of agricultural systems to the bioeconomy in poland: Integration of willow in the context of a stylised cap diversification. *Land Use Policy*, 99, 24p. doi: https://doi.org/10.1016/j.landusepol.2020.104797
10. Bayramoglu, B., Chakir, R., & Lungarska, A. (2020). Impacts of land use and climate change on freshwater ecosystems in france. *Environmental Modeling and Assessment*, 25, 147–172. doi: https://doi.org/10.1007/s10666-019-09673-x
11. Lungarska, A., Chakir, R., Ben Fradj, N., Jayet, P.-A., Bamière, L., Isbasoiu, A., Ollier, M., Gossiaux, E., Chiadmi, I., De Cara, S., & Kahindo, S. (2020). Approche intégrée des productions animales et végétales dans un modèle économique pour l'analyse des interactions agriculture – environnement. *Innovations Agronomiques*, 80, 69–85
12. Mosnier, C., Britz, W., Jullière, T., De Cara, S., Jayet, P.-A., Havlik, P., Franck, S., & Mosnier, A. (2019). Greenhouse gas abatement strategies and costs in dairy production – a comparison across bio-economic models. *Journal of Cleaner Production*, (236), 17589. doi: 10.1016/j.jclepro.2019.07.064
13. Jayet, P.-A., Barberis, D., Humblot, P., & Lungarska, A. (2018). Spatialisation de la demande en eau d'irrigation estimée par un modèle bioéconomique. *Revue internationale de géomatique*, 4(28), 485–503. doi: 10.3166/rig.2018.00064
14. Ben Fradj, N. & Jayet, P.-A. (2018). Optimal management of perennial energy crops by farming systems in france: A supply-side economic analysis. *Biomass and Bioenergy*, (116), 113–121. doi: 10.1016/j.biombioe.2018.06.003
15. De Cara, S., Henry, L., & Jayet, P.-A. (2018). Optimal coverage of an emission tax in the presence of monitoring, reporting, and verification costs. *Journal of Environmental Economics and Management*, 89, 1–23. doi: 10.1016/j.jeem.2018.03.001
16. Lungarska, A. & Chakir, R. (2018). Climate-induced Land Use Change in France: Impacts of Agricultural Adaptation and Climate Change Mitigation. *Ecological Economics*, 147, 134–154

17. Bounaffaa, M., Florio, A., Le Roux, X., & Jayet, P.-A. (2018). Economic and environmental analysis of maize inoculation by plant growth promoting rhizobacteria in the french rhône-alpes region. *Ecological Economics*, 146, 334–346. doi: 10.1016/j.ecolecon.2017.11.009
18. Jayet, P. A., Barberis, D., & Humblot, P. (2017b). The economic impact of climate change on agricultural water demand in france. In D. Kocasevic (Ed.), *Book of proceedings - VIII International Scientific Agriculture Symposium* (pp. 2032–2038).: AgroSym 2017
19. Chakir, R. & Lungarska, A. (2017b). Agricultural rent in land-use models: comparison of frequently used proxies. *Spatial Economic Analysis*, 12, 379–303. doi: 10.1080/17421772.2017.1273542
20. Humblot, P., Jayet, P.-A., & Petsakos, A. (2017). Farm-level bio-economic modeling of water and nitrogen use: Calibrating yield response functions with limited data. *Agricultural Systems*, 151, 47–60. doi: 10.1016/j.agsy.2016.11.006
21. Lungarska, A. & Jayet, P. A. (2016). Impact of spatial differentiation of nitrogen taxes on french farms' compliance costs. *Environmental and Resource Economics*. doi: 10.1007/s10640-016-0064-9
22. Assaiante, C., Jayet, P.-A., & Lantz, F. (2016). From raw material to biofuel demand: the agraf model. In A. R. E. Sabine Sauvage, José-Miguel Sánchez-Pérez (Ed.), *8th International Congress on Environmental Modelling and Software*. http://scholarsarchive.byu.edu/iemssconference/2016/Stream-A/15
23. Jayet, P. A., Aghajanzadeh-Darzi, P., & Petsakos, A. (2017a). Improvement of a bio-economic mathematical programming model in the case of on-farm source inputs and outputs. *Journal of Quantitative Economics*, 15, 489–508. doi: 10.1007/s40953-016-0058-z
24. Aghajanzadeh-Darzi, P., Martin, R., Laperche, S., & Jayet, P.-A. (2016). Climate change impacts on european agriculture revisited - adding the economic dimension of grasslands. *Regional Environmental Change*. doi: 10.1007/s10113-016-1018-z
25. Bourgeois, C., Habets, F., Jayet, P.-A., & Viennot, P. (2016). Estimating the marginal social value of agriculturally-driven nitrate concentrations in an aquifer: a combined theoretical-applied approach. *Water Economics and Policy*. doi: 10.1142/S2382624X16500211
26. Ben Fradj, N., Jayet, P. A., & Aghajanzadeh-Darzi, P. (2016). Competition between food, feed, and (bio)fuel: a supply-side model based assessment at the european scale. *Land Use Policy*, 52, 195–205. 10.1016/j.landusepol.2015.12.027
27. Jayet, P. A. & Petel, E. (2015). Economic valuation of the nitrogen content of urban organic residue by the agricultural sector. *Ecological Economics*, 120, 272–281
28. Bourgeois, C., Ben Fradj, N., & Jayet, P.-A. (2014). How cost-effective is a mixed policy targeting the management of three agricultural n-pollutants? *Environmental Modeling and Assessment*, 19(5), 389–405
29. Humblot, P., Leconte-Demarsy, D., Clerino, P., Szopa, S., Castell, J.-F., & Jayet, P.-A. (2013). Assessment of ozone impacts on farming systems: A bio-economic modeling approach applied to the widely diverse french case. *Ecological Economics*, 85, 50–58
30. Leclère, D., Jayet, P.-A., & de Noblet Ducoudré, N. (2013). Farm-level autonomous adaptation of european agricultural supply to climate change. *Ecological Economics*, 87, 1–14
31. Aghajanzadeh-Darzi, P. & Jayet, P. A. (2013). Impact économique du changement climatique sur l'alimentation animale et l'élevage dans l'union européenne. *Fourrages*, (215), 81–90
32. Ciaian, P., Espinosa, M., Gomez y Paloma, S., Heckelei, T., Langrell, S., Louichi, K., Sckokao, P., Thomas, A., & Vard, T. (2013). *Farm level modelling of CAP: a methodological overview*. Technical report, JRC Scientific and policy reports. doi:10.2791/86415
33. Jayet, P.-A. & Petsakos, A. (2013). Evaluating the efficiency of a uniform n-input tax under different policy scenarios at different scales. *Environmental Modelling and Assessment*, 18, 57–72
34. Cantelaube, P., Jayet, P.-A., Carré, F., Zakharov, P., & Bamps, C. (2012). Geographical downscaling of outputs provided by an economic farm model calibrated at the regional level. *Land Use Policy*, 29, 35–44
35. De Cara, S. & Jayet, P.-A. (2011). Marginal abatement costs of greenhouse gas emissions from european agriculture, cost effectiveness, and the eu non-ets burden sharing agreement. *Ecological Economics*, 70, 1680–1690
36. Galko, E. & Jayet, P.-A. (2011). Economic and environmental effects of decoupled agricultural support in the eu. *Agricultural Economics*, 42, 605–618
37. Durandeau, S., Gabrielle, B., Godard, C., Jayet, P.-A., & Le Bas, C. (2010). Coupling biophysical and microeconomic models to assess the effect of mitigation measures on greenhouse gas emissions from agriculture. *Climatic Change*, 98, 51–73

38. De Cara, S. & Thomas, A. (2008). *Projections d'émissions/absorptions de gaz à effet de serre dans les secteurs forêt et agriculture aux horizons 2010 et 2020.* Technical report. 192 pp
39. Godard, C., Roger-Estrade, J., Jayet, P.-A., Brisson, N., & Le Bas, C. (2008). Use of available information at a european level to construct crop nitrogen response curves for the regions of the eu. *Agricultural Systems*, 97, 68–82
40. De Cara, S. & Jayet, P.-A. (2006). Mitigation of greenhouse gas emissions in EU agriculture: An assessment of the costs of reducing agricultural emissions and enhancing carbon sinks in agricultural soils. In *INSEA Report SSP1-CT-2003-503614-Final, European Commission, INSEA, IIASA, Laxenburg, Austria*
41. Jayet, P.-A. (2006). *Deliverable 4.* Technical report, GENEDEC project document
42. Bamière, L., Godard, C., Debove, E., De Cara, S., Jayet, P.-A., & Niang, B. (2005). Interface between agriculture and the environment: integrating yield response functions in an economic model of EU agriculture. In *Modelling agricultural policies: State of the Art and New Challenges, Proceedings of the 89th European Association of Agricultural Economists, Parma, Italy, Feb. 3-5, 2005*
43. Jayet, P.-A. & Labonne, J. (2005). Impact d'une réforme de la politique agricole commune par le découplage. *Economie et Prévision*, 167(1), 101–116
44. De Cara, S., Houzé, M., & Jayet, P.-A. (2005). Methane and nitrous oxide emissions from agriculture in the EU: a spatial assessment of sources and abatement costs. *Environmental and Resource Economics*, 32(4), 551–583
45. Jayet, P. A. (2002). Mathematical programming for the computation of multi-market equilibria: an attempt in modelling the french markets of cereals and livestocks feeds. In S. Rozakis & J. Sourie (Eds.), *Comprehensive economic and spatial bio-energy modelling*, number 48 in Série A. Séminaires Méditerranéens (pp. 35–46).: CIHEAM
46. De Cara, S. & Jayet, P.-A. (2000a). Emissions of greenhouse gases from agriculture : the heterogeneity of abatement costs in france. *European Review of Agricultural Economics*, 75(4), 597–623
47. De Cara, S. & Jayet, P.-A. (2000b). Régulation de l'effet de serre d'origine agricole : Puits de carbone et instruments de second rang. *Economie et Prévision*, 143/144(2-3), 37–46
48. De Cara, S. & Jayet, P.-A. (1999). Evaluation et régulation de l'effet de serre d'origine agricole. *L'Actualité Economique*, 75(4), 597–623
49. Jayet, P.-A., Birfet, A., & Hofstetter, A. (1998). Forêt paysanne et politique agricole commune : une évaluation des impacts d'une incitation au reboisement. *Cahiers d'Economie et Sociologie Rurales*, 48, 6–35

Manual A description of part of the model building and file architecture was previously developed. It was partly open access, and will soon be entirely open access. These materials are mainly dedicated to *AROPAj* users.

Electronic reports Programs supported by the *AROPAj* model or which support it are available for *AROPAj* users. They reflect a major part of the development and problem solving over the years. Any research program is likely to supply model development opportunities. A series of "cahiers" / workbooks aggregates users' and developers' contributions over the years. In addition, technical reports may be provided for scientific partners, stakeholders and sponsors. This is the case for instance for the *GENEDEC* program from which strong *AROPAj* developments originate: Source: Synthetic material around *AROPAj* from the *GENEDEC* program

Circular slides A series of interlinked slides highlights different technical and informational aspects referenced in this paper. This series offers practical information about *AROPAj* use and its potential.

Side events Seminars dedicated to *AROPAj* developments and applications and to associated theoretical aspects in applied mathematics and economics are regularly organized. Three four-days seminars have been organized in Alps in 2010, 2013 and 2016 (consequently called "ski-minars").

How to create this document The present document originates from an assembly of text files, figures and charts. Text files are of *LaTeX* or *ORG* format, and edited through *Emacs*. Most of figures and charts originate from published papers and reports, and a few emanate from additional *R*, *Mathematica* or *CorelDraw* applications. The detailed block structure (see **??**) is provided by a *CShell* script applied to the last version of the kernel (see the "inception date" mentioned at the end of **??** introduction).

The master document is compiled through *Emacs-ORG*.

How to cite this document as a citation and a reference included in the bibliography,
as plain reference (updated 2020, June 30):

Jayet P.-A., Petsakos A., Chakir R., Lungarska A., De Cara S., Petel E., Humblot P., Godard C., Leclère D., Cantelaube P., Bourgeois C., Clodic M., Bamière L., Ben Fradj N., Aghajanzadeh-Darzi P., Dumollard G., Isbăşoiu A., Adrian J., Pilchak G., Bounaffaa M., Barberis D., Assaiante C., Ollier M., Henry L., Florio A., Chiadmi I., Gossiaux E., Ramirez R., Gérard M., and Reineix J., (2020). **The European agro-economic AROPAj model**. INRA - Economie Publique, 207p.
https://www6.versailles-grignon.inra.fr/economie_publique_eng/Research-work.

and encoded in *JabRef* or *Bibtex* as follows:

```
@book{Jayet2018,
author = {Jayet, Pierre-Alain
and Petsakos, Athanasios
and Chakir, Raja
and Lungarska, Anna
and {De Cara}, St{\'{e}}phane
and Petel, Elvire
and Humblot, Pierre
and Godard, Caroline
and Lecl{\`{e}}re, David
and Cantelaube, Pierre
and Bourgeois, Cyril
and Clodic, M\'elissa
and Bami{\`{e}}re, Laure
and {Ben Fradj}, Nosra
and Aghajanzadeh-Darzi, Parisa
and Dumollard, Gaspard
and Isb\u{a}\c{s}oiu, Ancuta
and Adrian, Juliette
and Pilchak, Gregory
and Bounaffaa, Myriam
and Barberis, Delphine
and Assaiante, Coline
and Ollier, Maxime
and Henry, Lo\"ic
and Florio, Alessandro
and Chidami, Ines
and Gossiaux, Eva
and Ramirez, Erica
and G{\'{e}}rard, Maxence
```

and Reineix, Julie
},
pages = {1--209},
publisher = {INRA, UMR Economie Publique},
address = {Paris Saclay Applied Economics, Université Paris Sacrin,
\url{https://www6.versailles-grignon.inrae.fr/psae/Media/fichiers/ArticlAROPAj}},
title = {{The European agro-economic AROPAj model}},
year = {2022}
}

References

Aghajanzadeh-Darzi, P. & Jayet, P. A. (2013). Impact économique du changement climatique sur l'alimentation animale et l'élevage dans l'union européenne. *Fourrages*, (215), 81–90.

Aghajanzadeh-Darzi, P., Martin, R., Laperche, S., & Jayet, P.-A. (2016). Climate change impacts on european agriculture revisited - adding the economic dimension of grasslands. *Regional Environmental Change*. doi: 10.1007/s10113-016-1018-z.

Arfini, F. (2001). *Mathematical Programming Models Employed in the Analysis of the Common Agriculture Policy*. INEA Working Papers 9.

Assaiante, C., Jayet, P.-A., & Lantz, F. (2016). From raw material to biofuel demand: the agraf model. In A. R. E. Sabine Sauvage, José-Miguel Sánchez-Pérez (Ed.), *8th International Congress on Environmental Modelling and Software*. http://scholarsarchive.byu.edu/iemssconference/2016/Stream-A/15.

Bamière, L., Godard, C., Debove, E., De Cara, S., Jayet, P.-A., & Niang, B. (2005). Interface between agriculture and the environment: integrating yield response functions in an economic model of EU agriculture. In *Modelling agricultural policies: State of the Art and New Challenges, Proceedings of the 89th European Association of Agricultural Economists, Parma, Italy, Feb. 3-5, 2005*.

Bamière, L., Jayet, P.-A., Kahindo, S., & Martin, E. (2021). Carbon sequestration in french agricultural soils: a spatial economic evaluation. *Agricultural Economics*. doi: https://doi.org/10.1111/agec.12619.

Barberis, D., Humblot, P., Chiadmi, I., Jayet, P.-A., Lungarska, A., & Ollier, M. (2020). Climate change and irrigation water: Should the north/south hierarchy of impacts on agricultural systems be reconsidered? *Environmental Modeling and Assessment*. doi: https://doi.org/10.1007/s10666-020-09724-8.

Bayramoglu, B., Chakir, R., & Lungarska, A. (2020). Impacts of land use and climate change on freshwater ecosystems in france. *Environmental Modeling and Assessment*, 25, 147–172. doi: https://doi.org/10.1007/s10666-019-09673-x.

Ben Fradj, N. & Jayet, P.-A. (2018). Optimal management of perennial energy crops by farming systems in france: A supply-side economic analysis. *Biomass and Bioenergy*, (116), 113–121. doi: 10.1016/j.biombioe.2018.06.003.

Ben Fradj, N., Jayet, P. A., & Aghajanzadeh-Darzi, P. (2016). Competition between food, feed, and (bio)fuel: a supply-side model based assessment at the european scale. *Land Use Policy*, 52, 195–205. 10.1016/j.landusepol.2015.12.027.

Ben Fradj, N., Jayet, P.-A., Rozakis, S., Georganta, E., & Jedrejek, A. (2020). Contribution of agricultural systems to the bioeconomy in poland: Integration of willow in the context of a stylised cap diversification. *Land Use Policy*, 99, 24p. doi: https://doi.org/10.1016/j.landusepol.2020.104797.

Bounaffaa, M., Florio, A., Le Roux, X., & Jayet, P.-A. (2018). Economic and environmental analysis of maize inoculation by plant growth promoting rhizobacteria in the french rhône-alpes region. *Ecological Economics*, 146, 334–346. doi: 10.1016/j.ecolecon.2017.11.009.

Bourgeois, C., Ben Fradj, N., & Jayet, P.-A. (2014). How cost-effective is a mixed policy targeting the management of three agricultural n-pollutants? *Environmental Modeling and Assessment*, 19(5), 389–405.

Bourgeois, C., Habets, F., Jayet, P.-A., & Viennot, P. (2016). Estimating the marginal social value of agriculturally-driven nitrate concentrations in an aquifer: a combined theoretical-applied approach. *Water Economics and Policy.* doi: 10.1142/S2382624X16500211.

Bourgeois, C. & Jayet, P. A. (2016). Regulation of relationships between heterogeneous farmers and an aquifer accounting for lag effects. *Australian Journal of Agricultural and Resource Economics*, 60(1), 39–59. doi: 10.1111/1467-8489.12102.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y., Cellier, P., Debaeke, P., Gaudillère, J., Hénault, C., Maraux, F., Seguin, B., & Sinoquet, H. (2003). An overview of the crop model stics. *European Journal of Agronomy*, 18(3-4), 309–332. Modelling Cropping Systems: Science, Software and Applications.

Brisson, N. & Levrault, F. (2010). *Changement climatique, agriculture et for en France: simulation d'impacts sur les principales esp?es. In: Livre vert de projet CLIMATOR (2007-2010).* ADEME.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M. H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.-M., Meynard, J.-M., & Delécolle, R. (1998). Stics: a generic model for the simulation of crops and their water and nitrogen balances. i. theory and parameterization applied to wheat and corn. *Agronomie*, 18(5-6), 311–346.

Cantelaube, P., Jayet, P.-A., Carré, F., Zakharov, P., & Bamps, C. (2012). Geographical downscaling of outputs provided by an economic farm model calibrated at the regional level. *Land Use Policy*, 29, 35–44.

Caputo, M. R. & Paris, Q. (2008). Comparative statics of the generalized maximum entropy estimator of the general linear model. *European Journal of Operational Research*, 185(1), 195–203.

Caurla, S., Delacote, P., Lecocq, F., Barthès, J., & Barkaoui, A. (2013). Combining an intersectoral carbon tax with sectoral mitigation policies: Impacts on the French forest sector. *Journal of Forest Economics*, 19(4), 450–461.

Chakir, R. (2009). Spatial downscaling of agricultural land-use data: An econometric approach using cross entropy. *Land Economics*, 85(2), 238–251.

Chakir, R. & Lungarska, A. (2017a). Agricultural rent in land-use models: comparison of frequently used proxies. *Spatial Economic Analysis*, 12(2-3), 279–303.

Chakir, R. & Lungarska, A. (2017b). Agricultural rent in land-use models: comparison of frequently used proxies. *Spatial Economic Analysis*, 12, 379–303. doi: 10.1080/17421772.2017.1273542.

Ciaian, P., Espinosa, M., Gomez y Paloma, S., Heckelei, T., Langrell, S., Louichi, K., Sckokao, P., Thomas, A., & Vard, T. (2013). *Farm level modelling of CAP: a methodological overview.* Technical report, JRC Scientific and policy reports. doi:10.2791/86415.

Clodic, M. (2013). *Modélisation de lffre de biocarburants à l'echelle européenne, échanges intra- et extra-européens, effets croisés avec les marchés ptroliers et contrainte carbone.* PhD thesis, ABIES-AgroParisTech, dir.: P.-A. Jayet.

Day, R. H. (1963). On aggregating linear programming models of production. *Journal of Farm Economics*, 45(4), 797–813.

De Cara, S., Henry, L., & Jayet, P.-A. (2018). Optimal coverage of an emission tax in the presence of monitoring, reporting, and verification costs. *Journal of Environmental Economics and Management*, 89, 1–23. doi: 10.1016/j.jeem.2018.03.001.

De Cara, S., Houzé, M., & Jayet, P.-A. (2005). Methane and nitrous oxide emissions from agriculture in the EU: a spatial assessment of sources and abatement costs. *Environmental and Resource Economics*, 32(4), 551–583.

De Cara, S. & Jayet, P.-A. (1999). Evaluation et régulation de l'effet de serre d'origine agricole. *L'Actualité Economique*, 75(4), 597–623.

De Cara, S. & Jayet, P.-A. (2000a). Emissions of greenhouse gases from agriculture : the heterogeneity of abatement costs in france. *European Review of Agricultural Economics*, 75(4), 597–623.

De Cara, S. & Jayet, P.-A. (2000b). Régulation de l'effet de serre d'origine agricole : Puits de carbone et instruments de second rang. *Economie et Prévision*, 143/144(2-3), 37–46.

De Cara, S. & Jayet, P.-A. (2006). Mitigation of greenhouse gas emissions in EU agriculture: An assessment of the costs of reducing agricultural emissions and enhancing carbon sinks in agricultural soils. In *INSEA Report SSP1-CT-2003-503614-Final, European Commission, INSEA, IIASA, Laxenburg, Austria.*

De Cara, S. & Jayet, P.-A. (2011). Marginal abatement costs of greenhouse gas emissions from european agriculture, cost effectiveness, and the eu non-ets burden sharing agreement. *Ecological Economics*, 70, 1680–1690.

De Cara, S. & Thomas, A. (2008). *Projections d'émissions/absorptions de gaz à effet de serre dans les secteurs forêt et agriculture aux horizons 2010 et 2020.* Technical report. 192 pp.

Devaraju, N., Prudhomme, R., Lungarska, A., Wang, X., Yin, Z., de Noblet-Ducoudré, N., Chakir, R., Jayet, P.-A., Brunelle, T., Viovy, N., De Palma, A., Gonzalez, R., & Ciais, P. (2022). Quantifying the benefits of reducing synthetic nitrogen application policy on ecosystem carbon sequestration and biodiversity. *Scientific Reports 2022 12:1*, 12(1), 1–15.

Durandeau, S., Gabrielle, B., Godard, C., Jayet, P.-A., & Le Bas, C. (2010). Coupling biophysical and microeconomic models to assess the effect of mitigation measures on greenhouse gas emissions from agriculture. *Climatic Change*, 98, 51–73.

Errington, S. (1989). Estimating Enterprise Input-Output Coefficients From Regional Farm Data. *Journal of Agricultural Economics*, 40(1), 52–56.

Florio, A., Pommier, T., Gervaix, J., Bérard, A., & Le Roux, X. (2018). Soil C and N statuses determine the effect of maize inoculation by plant growth-promoting rhizobacteria on nitrifying and denitrifying communities. *Scientific reports*, 7(1), 8411.

Freund, R. J. (1956). The Introduction of Risk into a Programming Model. *Econometrica*, 24(3), 253–263.

Galko, E. & Jayet, P.-A. (2011). Economic and environmental effects of decoupled agricultural support in the eu. *Agricultural Economics*, 42, 605–618.

Godard, C., Roger-Estrade, J., Jayet, P.-A., Brisson, N., & Le Bas, C. (2008). Use of available information at a european level to construct crop nitrogen response curves for the regions of the eu. *Agricultural Systems*, 97, 68–82.

Golan, A., Judge, G., & Miller, D. (1996). Maximum entropy econometrics: Robust estimation with limited data. *New York, John Wiley eds.*

Golan, A., Judge, G., & Perloff, J. (1997). Estimation and inference with censored and ordered multinomial response data. *Journal of Econometrics*, 79(1), 23–51.

Gomez, E. (2002). *Modélisation intégrée du transfert de nitrate à l'échelle régionale dans un système hydrologique. Application au bassin de la Seine.* PhD thesis, Ecole Nationale Supérieure des Mines de Paris, France.

Gomez, E., Ledoux, E., Viennot, P., Mignolet, C., Benoit, M., Bornerand, C., Schott, C., Mary, B., Billen, G., Ducharne, A., et al. (2003). Un outil de modélisation intégrée du transfert des nitrates sur un système hydrologique: application au bassin de la seine. *La Houille Blanche*, (3), 38–45.

Good, I. J. (1963). Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. *The Annals of Mathematical Statistics*, Vol. 34, No. 3., 911–934.

Gossiaux, E. & Jayet, P.-A. (2021). Exploring the potential for calories and bioenergy in france. *BE-Sustainable Magazine*, 12, 11–17. https://issuu.com/besustainablemagazine/docs/be-sustainable_magazine_issue_12_-_april_2021/s/12150135.

Gérard, M. & Jayet, P. (2023). European farmers' response to crop residue prices and implications for bioenergy policies. *Energy Policy.* https://doi.org/10.1016/j.enpol.2023.113561.

Havlík, P., Schneider, U. A., Schmid, E., Böttcher, H., Fritz, S., Skalskỳ, R., Aoki, K., De Cara, S., Kindermann, G., Kraxner, F., et al. (2011). Global land-use implications of first and second generation biofuel targets. *Energy Policy*, 39(10), 5690–5702.

Hazell, P. B. R. & Norton, R. D. (1986). *Mathematical Programming for Economic Analysis in Agriculture*. New York: MacMillan.

Heckelei, T. & Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy. *European Review of Agricultural Economics*, 30(1), 27–50.

Hénin, S. & Dupuis, M. (1945). *Essai de bilan de la matière organique du sol*, volume 15.

Herrero, M., Thornton, P., Notenbaert, A., Wood, S., Msangi, S., Freeman, H., Bossio, D., Dixon, J., Peters, M., Steeg, J., et al. (2010). Smart investments in sustainable food production: revisiting mixed crop-livestock systems.

Howitt, R. E. (1995). Positive Mathematical Programming. *American Journal of Agricultural Economics*, 77(2), 329–342.

Humblot, P., Jayet, P.-A., & Petsakos, A. (2017). Farm-level bio-economic modeling of water and nitrogen use: Calibrating yield response functions with limited data. *Agricultural Systems*, 151, 47–60. doi: 10.1016/j.agsy.2016.11.006.

Humblot, P., Leconte-Demarsy, D., Clerino, P., Szopa, S., Castell, J.-F., & Jayet, P.-A. (2013). Assessment of ozone impacts on farming systems: A bio-economic modeling approach applied to the widely diverse french case. *Ecological Economics*, 85, 50–58.

Isbasoiu, A., Jayet, P.-A., & De Cara, S. (2020). Increasing food production and mitigating agricultural greenhouse gas emissions in the european union: impacts of carbon pricing and calorie production targeting. *Environmental Economics and Policy Studies*, (pp. 32p). doi: https://doi.org/10.1007/s10018-020-00293-4.

Izaurralde, R., Williams, J., McGill, W., Rosenberg, N., & Quiroga Jakas, M. (2006). Simulating soil c dynamics with epic: Model description and testing against long-term data. *Ecological Modelling*, 192(3), 362–384.

Jayet, P. A. (2002). Mathematical programming for the computation of multi-market equilibria: an attempt in modelling the french markets of cereals and livestocks feeds. In S. Rozakis & J. Sourie (Eds.), *Comprehensive economic and spatial bio-energy modelling*, number 48 in Série A. Séminaires Méditerranéens (pp. 35–46).: CIHEAM.

Jayet, P.-A. (2006). *Deliverable 4*. Technical report, GENEDEC project document.

Jayet, P. A., Aghajanzadeh-Darzi, P., & Petsakos, A. (2017a). Improvement of a bio-economic mathematical programming model in the case of on-farm source inputs and outputs. *Journal of Quantitative Economics*, 15, 489–508. doi: 10.1007/s40953-016-0058-z.

Jayet, P. A., Barberis, D., & Humblot, P. (2017b). The economic impact of climate change on agricultural water demand in france. In D. Kocasevic (Ed.), *Book of proceedings - VIII International Scientific Agriculture Symposium* (pp. 2032–2038).: AgroSym 2017.

Jayet, P.-A., Barberis, D., Humblot, P., & Lungarska, A. (2018). Spatialisation de la demande en eau d'irrigation estimée par un modèle bioéconomique. *Revue internationale de géomatique*, 4(28), 485–503. doi: 10.3166/rig.2018.00064.

Jayet, P.-A., Birfet, A., & Hofstetter, A. (1998). Forêt paysanne et politique agricole commune : une évaluation des impacts d'une incitation au reboisement. *Cahiers d'Economie et Sociologie Rurales*, 48, 6–35.

Jayet, P.-A., Isbasoiu, A., & De Cara, S. (2020). Slaughter cattle to secure food calories and reduce agricultural greenhouse gas emissions? some prospective estimates for france. *Review of Agricultural, Food and Environmental Studies*, (pp. 24p). doi: https://doi.org/10.1007/s41130-020-00117-9.

Jayet, P.-A. & Labonne, J. (2005). Impact d'une réforme de la politique agricole commune par le découplage. *Economie et Prévision*, 167(1), 101–116.

Jayet, P. A. & Lapierre, O. (1998). *La dynamique des systèms*, chapter Modélisation des marchés de l'alimentation animale, (pp. 73–93). Hermès.

Jayet, P. A. & Petel, E. (2015). Economic valuation of the nitrogen content of urban organic residue by the agricultural sector. *Ecological Economics*, 120, 272–281.

Jayet, P.-A. & Petsakos, A. (2013). Evaluating the efficiency of a uniform n-input tax under different policy scenarios at different scales. *Environmental Modelling and Assessment*, 18, 57–72.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physics Review*, 106, 620–630.

Kampas, A., Petsakos, A., & Rozakis, S. (2012). Price induced irrigation water saving: Unraveling conflicts and synergies between European agricultural and water policies for a Greek Water District. *Agricultural Systems*, 113, 28–38.

Kennedy, J. O. (1975). Using regression analysis to reduce aggregation bias in linear programming supply models. *The Australian Journal of Agricultural Economics*, 19(1), 11p.

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G., & Ullerstig, A. (2011). 21st century changes in the european climate: uncertainties derived from an ensemble of regional climate model simulations. *Tellus A*, 63(1), 24–40.

Leclère, D., de Noblet-Ducoudré, N., & Jayet, P.-A. (2014). Understanding the farm-scale response of european agriculture to climate change: evaluation of a coupled economic-crop model. *submittted*.

Leclère, D., Jayet, P.-A., & de Noblet Ducoudré, N. (2013). Farm-level autonomous adaptation of european agricultural supply to climate change. *Ecological Economics*, 87, 1–14.

Ledoux, E. (1980). *Modélisation intégrée des écoulements de surface et des écoulements souterrains sur un bassin hydrologique*. PhD thesis, Ecole des Mines de Paris.

Ledoux, E., Girard, G., de Marsily, G., Villeneuve, J., & Deschenes, J. (1989). Spatially distributed modelling: conceptual approach, coupling surface water and groundwater, 1989. *Unsaturated flow in hydrologic modelling–theory and practice. NATO ASI ser. CNorwell, Massachussett: Kluwer academic*, (pp. 435–454).

Ledoux, E., Girard, G., & Villeneuve, J. (1984). Proposition d'un modèle couplé pour la simulation conjointe des écoulements de surface et des écoulements souterrains sur un bassin hydrologique. *La houille blanche*, (1-2), 101–120.

Ledoux, E., Gomez, E., Monget, J., Viavattene, C., Viennot, P., Ducharne, A., Benoit, M., Mignolet, C., Schott, C., & Mary, B. (2007). Agriculture and groundwater nitrate contamination in the seine basin. the stics-modcou modelling chain. *Science of the Total Environment*, 375(1-3), 33–47.

Leip, A., Weiss, F., Wassenaar, T., Perez, I., Fellmann, T., Loudjani, P., Tubiello, F., Grandgirard, D., Monni, S., & Biala, K. (2010). *Evaluation of the livestock sector's contribution to the EU greenhouse gas emissions (GGELS).* Technical report, European Commission, Joint Research Centre (JRC).

Léon, Y., Peeters, L., Quinqu, M., & Surry, Y. (1999). The Use of Maximum Entropy to Estimate Input-Output Coefficients From Regional Farm Accounting Data. *Journal of Agricultural Economics*, 50(3), 425–439.

Lobianco, A., Delacote, P., Caurla, S., & Barkaoui, A. (2016). Accounting for active management and risk attitude in forest sector models. *Environmental Modeling & Assessment*, 21, 391–405.

Louhichi, K., Jacquet, F., & Butault, J. P. (2012). Estimating input allocation from heterogeneous data sources: A comparison of alternative estimation approaches. *Agricultural Economics Review*, 13(2), 83–102.

Lungarska, A. & Chakir, R. (2018). Climate-induced Land Use Change in France: Impacts of Agricultural Adaptation and Climate Change Mitigation. *Ecological Economics*, 147, 134–154.

Lungarska, A., Chakir, R., Ben Fradj, N., Jayet, P.-A., Bamière, L., Isbasoiu, A., Ollier, M., Gossiaux, E., Chiadmi, I., De Cara, S., & Kahindo, S. (2020). Approche intégrée des productions animales et végétales dans un modèle économique pour l'analyse des interactions agriculture – environnement. *Innovations Agronomiques*, 80, 69–85.

Lungarska, A. & Jayet, P. A. (2016). Impact of spatial differentiation of nitrogen taxes on french farms' compliance costs. *Environmental and Resource Economics*. doi: 10.1007/s10640-016-0064-9.

Makowski, D., Wallach, D., & Meynard, J. (1999). Models of yield, grain protein, and residual mineral nitrogen responses to applied nitrogen for winter wheat. *Agronomy Journal*, 91(3), 377–385.

Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77–91.

McCarl, B. A. & Spreen, T. H. (1980). Price Endogenous Mathematical Programming as a Tool for Sector Analysis. *American Journal of Agricultural Economics*, 62(1), 87–102.

Mills, G., Buse, A., Gimeno, B., Bermejo, V., Holland, M., Emberson, L., & Pleijel, H. (2007). A synthesis of aot40-based response functions and critical levels of ozone for agricultural and horticultural crops. *Atmospheric Environment*, 41(41), 2630–2643.

Mosnier, C., Britz, W., Jullière, T., De Cara, S., Jayet, P.-A., Havlik, P., Franck, S., & Mosnier, A. (2019). Greenhouse gas abatement strategies and costs in dairy production – a comparison across bio-economic models. *Journal of Cleaner Production*, (236), 17589. doi: 10.1016/j.jclepro.2019.07.064.

Olesen, J. & Bindi, M. (2002). Consequences of climate change for european agricultural productivity, land use and policy. *European Journal of Agronomy*, 16(4), 239–262.

Peltre, C., Christensen, B., Dragon, S., Icard, C., Kätterer, T., & Houot, S. (2012). RothC simulation of carbon accumulation in soil after repeated application of widely different organic amendments. *Soil Biology and Biochemistry*, 52(2012), 49–60.

Penman, J., D, K., Galbally, I., Hiraishi, T., Nyenzy, B., Emmanul, S., Buendia, L., Hoppaus, R., Martinsen, T., Meijer, J., Miwa, K., & Tanabe, K., Eds. (2001). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Intergovernmental Panel on Climate Change (IPCC). Japan: Institute for Global Environmental Strategies (IGES).

Prudhomme, R., Chakir, R., Lungarska, A., Brunelle, T., Devaraju, N., de Noblet, N., Jayet, P.-A., De Cara, S., & Bureau, J.-C. (2022). Food, climate and biodiversity: a trilemma of mineral nitrogen use in European agriculture. *Review of Agricultural, Food and Environmental Studies*.

Reidsma, P., Ewert, F., Oude Lansink, A., & Leemans, R. (2009). Vulnerability and adaptation of european farmers: a multi-level analysis of yield and income responses to climate variability. *Regional Environmental Change*.

Riedo, M., Grub, A., Rosset, M., & Fuhrer, J. (1998). A pasture simulation model for dry matter production, and fluxes of carbon, nitrogen, water and energy. *Ecological Modelling*, 105, 141–183.

Rötter, R. & van de Geijn, S. (1999). Climate change effects on plant growth, crop yield and livestock. *Climatic Change*, 43(4), 651–681.

Schmid, M., Neftel, A., Riedo, M., & Fuhrer, J. (2001). Process-based modelling of nitrous oxide emissions from different nitrogen sources in mown grassland. *Nutrient Cycling in Agroecosystems*, 60, 177–187.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.

Soussana, J. & Lüscher, A. (2007). Temperate grasslands and global atmospheric change: a review. *Grass and Forage Science*, 62(2), 127–134.

Thornley, J. (1998). *Grassland dynamics - An ecosystem simulation model*. CAB International.

Vacheron, J., Desbrosses, G., Bouffaud, M.-L., Touraine, B., Moënne-Loccoz, Y., Muller, D., Legendre, L., Wisniewski-Dyé, F., & Prigent-Combaret, C. (2013). Plant growth-promoting rhizobacteria and root system functioning. *Frontiers in Plant Science*, 4(September), 1–19.

Viennot, P., Monget, J., Ledoux, E., & Schott, C. (2007). *Modélisation de la pollution nitrique des aquifères du bassin de la Seine : intégration des bases de données actualisées des pratiques agricoles, validation des simulations sur la période 1970-2004, simulations prospectives de mesures agro-environnementales.* rapport r070209pvie, 50 p., Ecole des mines de Paris, Centre de Géosciences.

Vuichard, N., Ciais, P., Viovy, N., & Soussana, J. (2007a). Estimating the greenhouse gas fluxes of european grasslands with a process-based model: 2. simulations at the continental level. *Global Biogeochemical Cycles*, 21.

Vuichard, N., Soussana, J., Ciais, P., Viovy, N., Ammann, C., Calanca, P., Clifton-Brown, J., Fuhrer, J., Jones, M., & Martin, C. (2007b). Estimating the greenhouse gas fluxes of european grasslands with a process-based model: 1. model evaluation from in situ measurements. *Global Biogeochemical Cycles*, 21.

Appendix

A User's call for the model

The user is supposed to have an account on the workstation dedicated to the model. Some slight preliminary operations are required before the model's use, properly speaking, regarding "login" and "path" command files and internal access rights. The user should be also preferably familiar with basic *Linux*-shell commands.

A few screen capture schemes detail how the user accedes to the *AROPAj* based shell leading to simulations and outputs exploitation. The following figures illustrate how it runs through screen copies from the user's point of view.

In order to initiate a new simulation:

```
          ##   #  ####  ####    ##          ####
         ### ##   ##   ##  #   ###      ##  ##
        ## ###   ##   #####   ## #      #  ###
       ##   #   ##   ##  #   #####   # #  ##
      ##   ##  ###  ##  ##  ##   ## ###  ####

     Projet dans lequel s'inscrit votre activité:
    ---------------------------------------------
    | choix de l'étude (indiquer le nom) et/ou RET  |
    ---------------------------------------------
 (x pour STOP)    -> ▌
```

1. *After the execution of the "miraj" command, select a project (ask the administrator for the project list).*

```
    ---------------------------------------------
    |  i  Information générale (cahier memento ...) |
    |  o  Organigramme relatif (executable CSH e )  |
    |  a  Automate de simulation (s au préalable)   |
    |     (choix optionnel d'un répertoire travail) |
    |  f  Extraction résultats avec séparateur Xcel |
    |     (après étape "s" complète ou étape "a"    |
    |  g  Test "groupe" d'une simulation réalisée   |
    |     (suppose l'"initialisation pour "belg")   |
    |  n  Destruction complète d'une simulation     |
    |     (choix optionnel d'un répertoire travail) |
    |  s  Simulation (préparation CSH au préalable) |
    |     (choix optionnel d'un répertoire travail) |
    |  h  Spatialisation des résultats - sortie dbf |
    |     (suppose choix d'un projet nom prédéfini) |
    |  x  Sortie du menu                            |
    ---------------------------------------------
    choix (i/s/a/... + nom éventuel/x -> ▌
```

2. *Select "s" for initiating the simulation process.*

```
-----------------------------------------------------------------
|               choix de la version "modèle+aropaj"               |
|-----------------------------------------------------------------|
|  Opération en cours : indiquer identifiant 12 chiffres (tableau) |
|  Sinon choix parmi :                                            |
|      600 : UE_15  (février 2001)                                |
|      700 : UE-15  (juin 2005)                                   |
|      800 : UE-24  (aout 2007)                                   |
|     20?? : UE-27  (novembre 2014 -3 annees RICA 2007-8-9)       |
|            (indiquer un nombre < 109 si identifiant oublie)     |
-----------------------------------------------------------------
 choix --> (?00/../2009/sinon x) ▌
```

3. *Select a version of the AROPAj model ("2009" for the current version).*

```
-----------------------------------------------------------------
|          intégration d un problème pour tous les pays           |
|-----------------------------------------------------------------|
| Choix de la commande générique : (extension "csh")              |
| Cette commande doit se trouver dans le répertoire de travail    |
| Elle doit être conçue avec le "pays" comme paramètre            |
| Elle doit créer les fichiers PVFGI.(par var gar f ini)          |
-----------------------------------------------------------------
 choix --> (.../n/r/sinon x) ▌
```

4. *Select the name of one of your csh shell simulation script (without ".csh"). If you don't remember the name of the script, type any letter and Miraj will propose the available scripts.*

...

Launching simulations asked by some complex scenario design may lead to run times of several hours. At the moment (2015), one hundred simulations for a complete European cover (e.g. 27 Member States and more than 1800 unit *LP*s) lasts around half an hour when the workstation is

...

```
-----------------------------------------------------------------
| Sélection des options PAC sous la forme "IP1+IP2+IP3..." ou "0"  |
| où IPi est un indice de type novpac (1 2 ... 13 ...) (ex: 13+1)  |
-----------------------------------------------------------------
 choix --> (.../sinon x/r) 0
```

5. *Fill with "0" as a choice for CAP options.*

```
-----------------------------------------------------------------
| Incorporation dans le fichier des paramètres (mgg.par) de :      |
| modpac (1) refscop (10) lisscal (100) externe (1000)             |
| Les fichiers modpac.par (voir opération "3") et externe.par sont |
| supposés pré-exister (pour les version et pays demandés)         |
| Le format d entrée : somme des puissances de 10 (ex: 1001)       |
-----------------------------------------------------------------
 choix --> (.../sinon 0/x/r) 1000
```

6. *Fill with "1000" to chose the parameters file. Note the simulation identifier (12 digits). Output file names will refer to this identifier and to farm group index.*

```
-----------------------------------------------------------------
| détruire les fichiers créés lors d'une simulation précédente     |
| ... si non, conserve les exécutables pour gain de temps compil.  |
-----------------------------------------------------------------
 choix --> (y/n/x/r) n
           ''''''
Identifiant ensemble de simulations : 150522145458
           ''''''
>>>> Opération sur tableau.150522145458
```

7. *Usually fill with "n". Fill with "y" to delete files from previous simulation. Note the simulation identifier (12 digits).*

```
-----------------------------------------------------------------
| Envoi en boucle sur tous pays de l opération à venir ?           |
-----------------------------------------------------------------
 choix --> (y/n/x) n
```

8. *Fill with "n" (fill with "y" to start a loop).*

```
-----------------------------------------------------------------
| Choix du pays (4 lettres) parmi :
| belg bgra cypr czek dani deu1 deu2 deu3 ella esp1 esp2 esto fra1
 ita3 ltua luxe lvat mlta nede osto pol1 pol2 pol3 pol4 port rou1
|
-----------------------------------------------------------------
 choix --> (a4/sinon x) ▌
```

9. *Chose the AROPAj country or sub-country, identified by a 4-character sequence.*

```
-----------------------------------------------------------------
| Choix opération (i1) parmi :                                     |
|     p  opération sur commande propre avec "pays" comme paramètre |
|     0  (base)                                                    |
|     3  (modpac)                                                  |
|     4  (validation)                                              |
|     5  (initialisation)                                          |
|     6  (simulation série)                                        |
|     7  (agrégation résultats)                                    |
|     8  (tabulation résultats)                                    |
|     n  ("nettoyage fichiers")                                    |
-----------------------------------------------------------------
 choix --> (./r/x) ▌
```

10. *Initiate the simulation process by calling successively "p", "0", "3", "4", and "5" for a country. The process can be stopped anytime to check the first LPs solving (fill with "x" as many times as necessary to leave the Miraj menu).*
You may complete the steps on the order "p", "0", "3", "4", "5", ("6" in case of iterated simulations), "7", and "8".

Figure 55. Launching the call for *AROPAj* (1).

In order to complete a simulation that has already been initiated :

```
-----------------------------------------------------------------
|               choix de la version "modèle+aropaj"               |
|-----------------------------------------------------------------
|  Opération en cours : indiquer identifiant 12 chiffres (tableau)
|  Sinon choix parmi :
|      600 : UE_15  (février 2001)
|      700 : UE-15  (juin 2005)
|      800 : UE-24  (aout 2007)
|     20?? : UE-27  (novembre 2014 -3 annees RICA 2007-8-9)
|            (indiquer un nombre < 109 si identifiant oublie)
-----------------------------------------------------------------
  choix --> (?00/../2009/sinon x) 00000
##################################################
 >>> ATTENTION : choix de 12 chiffres parmi: <<<
150319161521
150407152450
150407162456
150522145458
##################################################
```

*Repeat steps **1.** and **2.**. Then, select the simulation identifier (usually the last one). If you don't remember the identifier, type any number set and Miraj will propose the available identifiers.*
*Go directly to step **8.**, **9.** and **10.**, to continue the simulation process.*

Once out of the Miraj menu (fill with "x" as many times as necessary), you can follow the simulation process through the table identified by the 12-digit identifier. In this table, a "country" (or a "sub-country" when country splitting) refers to a 8-digit identifier.

```
    Tableau des operations cree selon l'indexe date : 150522145458  - indicatif version : 2009
                                                                    - commande integree : simuGES
                                                                    - selection modpac  : 0
                                                                    - selection fichier : 1000
```

PAYS	nbgrp	base (0)	modpac (3)	comm. ex. (p)	selection(4)	initialisat. (5)	identif.	simulation (6)	agréga res (7)	tabula res (8)	nettoyage
belg	29										
bgra	42										
cypr	7										
czek	37	0 052214	3 052214	p 052214	4 052214	5 052215	22150059				
dani	26										

*Experimented users may execute all simulations through the automatic simulation process by selecting "a" from the **2.** menu (see Figure **??**).*

Figure 56. Launching the call for *AROPAj* (2).

moderately loaded.

In case of computation problems, the user is kindly advised to check first her/his own files (writing format, inadequate blank lines in `scen.(par var gar)` files - see section **??**, `FORTRAN` code in `scen.f` or *Cshell* code in `scen.csh`). Problems may be identified thanks to information tracking provided on files of type name `sortie.*`. Secondly, some *LP*s may be "infeasible" due to inadequate values of parameters (e.g. in case of too restrictive quotas). Thirdly, as a researchers' model surely not sufficiently developed in line with the rules of the computation art, *AROPAj* encounters many problems that often PhD students are able to invent.

B Contributors to *AROPAj* and to related publications

The list mentions contributors who used at least partially the model and who directly or indirectly contributed to its development, as *PhD* or *MSc* students and engineers or researchers hired on permanent positions (*PP*) or on short or middle term contracts (*EHC*).

- Annie Hofstetter (*PP*; 1990-1997)
- Philippe Bontems (*MSc*; 1990, 1992)
- Nathalie Taverdet-Popiolek (*PhD*; 1990-1993)
- Abdel Aloualiti (*EHC*; 1991)
- Frédérique Lefaudeux (*MSc, EHC*; 1991-1992)
- Vincent Thuillier (*MSc*; 1991)
- Jérome Adda (*MSc*; 1991)
- Sophie Chavanne (*MSc*; 1992)
- Joël Mathurin (*MSc, PhD*; 1992, 1994-1997)
- Alessandra Schiavina (*MSc*; 1993)
- Franck Leclerq-Blondel (*EHC*; 1994)
- Gilles Le Moguédec (*MSc, EHC*; 1995-1996)
- Carolina Chaya (*PhD*; 1995-1997)
- Roberta Silvagni (*MSc*; 1996)
- Nicolas Godfroy (*MSc*; 1996)
- Alain Birfet (*EHC*; 1996-1997)
- Stéphane De Cara (*MSc, PhD, PP*; 1997-2000, occasional, 2015-)
- Solène Maugars (*MSc*; 1998)
- Eleonora Marzocchi (*MSc, EHC*; 1998-2000)
- Kais Abbes (*MSc*; 1999)
- Michele Donati (*MSc, EHC*; 2000)
- Guillaume Gruère (*MSc*; 2000)
- Caroline Godard (*PhD*; 2002-2005)
- Laure Bamière (*MsC, PP*; 2002-2005, occasional, 2017-)
- Julien Labonne (*EHC*; 2003)
- Martin Houzé (*MSc*; 2003-2004)
- Bineta Niang (*MSc, EHC*; 2003)

- Mathieu Godet (*MSc*; 2004)
- Elodie Debove-Galko (*PhD*; 2004-2007)
- Edouard Baranger (*MSc, EHC*; 2005-2006)
- Raja Chakir (*PP*; 2005-2006, occasional, 2016-)
- Nathalie Novello (*EHC*; 2006-2008)
- Fanny Guillet (*MSc*; 2006)
- Sophie Durandeau (*EHC*; 2006-2007)
- Marta Piñies (*MSc*; 2006)
- Paul Zakharov (*EHC*; 2006-2010)
- Agathe Rouaix (*MSc*; 2007)
- Bernd Kuepker (*MSc*; 2007)
- Melissa Clodic (*MSc, PhD*; 2007-2013)
- Cloé Garnache (*MSc*; 2007)
- Basile Grosdidier (*EHC*; 2007)
- Marie Génin (*MSc*; 2008)
- Pierre Cantelaube (*EHC*; 2008)
- Audrey Pollard (*MSc*; 2008)
- Cyril Bourgeois (*PhD*; 2008-2011)
- David Leclère (*MSc, PhD*; 2008-2013)
- Athanasios Petsakos (*MSc, EHC*; 2009-2010, 2012-2014)
- Nosra Ben Fradj (*MSc, PhD, EHC*; 2009-2013)
- Delphine Leconte-Demarsy (*EHC*; 2010)
- Parisa Aghajanzadeh-Darzi (*MSc, PhD*; 2010-2013, 2015-2016)
- Paola Clerino (*EHC*; 2011)
- Joao-Pedro Domingues (*MSc*; 2011)
- Pierre Humblot (*EHC, PhD, own company*; 2011-2016, 2017-2018)
- Myrsini Papadaki (*MSc*; 2012)
- Anna Lungarska (*MSc, PhD, EHC*; 2012-2016, 2017-)
- Sara Yavari (*MSc*; 2013)
- Elvire Petel (*EHC*; 2013-2015)
- Gaspard Dumollard (*PhD*; 2014-2016)
- Gregory Pilchak (*MSc*; 2015)
- Ancuta Isbasoiu (*EHC, PhD*; 2015-2019)
- Juliette Adrian (*EHC*; 2015-2016)
- Clément Servant (*EHC*, supervised by S. De Cara; 2015)
- Myriam Bounaffaa (*EHC*; 2015-2017)
- Gabrielle Despres (*EHC*; 2015-2016)
- Coline Assaiante (*EHC*; 2015-2016)
- Delphine Barberis (*EHC*; 2016-2017)
- Loïc Henry (*MSc, PhD*, supervised by S. De Cara; 2016-2020)
- Maxime Ollier (*MSc, EHC, PhD*; 2017-2018, 2018-2021)

- Monia El-Akkari (*PhD*; 2018-2020, supervised by B. Gabrielle and J.L. Drouet)
- Salomé Kahindo (*MSc*, supervised by L. Bamière, 2018; 2019-)
- Jeffrey Norville (*EHC*, supervised by V. Martinet; 2018-2019)
- Eleni Georganta (*MSc* ERASMUS, 2019)
- Panagiota Arampatzi (*MSc* ERASMUS, 2019)
- Amal Azizi (*PostDoc*, 2019)
- Inès Chiadmi (*EHC*, *PhD* co-supervised by S. Van Passel; 2019-2021, 2021-2024)
- Eva Gossiaux (*MSc*, 2019-2020)
- Adrien Coiffard (*MSc*, 2020)
- Erica Ramirez (*MSc*, *EHC*, 2020-2021)
- Anaïs Kanellos (*MSc*, *EHC*, 2021)
- Maxence Gérard (*MSc*, *EHC*, 2021)
- Julie Reineix (*MSc*, *EHC*, *PhD* co-supervised by A. Ducharne; 2021, 2021-2024)
- Théodore Fontenaille (*MSc*, supervised by Ministère de la Transition Ecologique, 2022)
- Lisa Baldi (*visiting PhD* Univ. Parma, 2023)
- Olga Zuravel (*EHC*, 2023-2024)
- Mingzhu Weng (*MSc*, 2023)

C Funded projects and supports benefiting *AROPAj*

Projects that follow contributed to the development of *AROPAj*. Obviously fundings that were obtained were generally only partly devoted to the model itself.

1. Société Louis-Dreyfus, 1990-1992 (initiates *AROPAj* - France & UK)
2. Unigrains, 1990-1993, (PhD - theory-oriented)
3. AIP Inra "modélisation du secteur agricole", 1991-1993 (coll. Y. Le Roux, A. Hofstetter)
4. Inra Support, 1991-1993 (PhD co-fin. Inra-Edf N. Taverdet)
5. Commissariat Général du Plan & Direction de la Prévision, 1991-1992, convention d'étude CGP n°11/1191, prestation DP 910127, (coll. J.C. Bureau, Y. Le Roux)
6. AIP Inra "Politique Agricole Commune", 1992-1994 (*MS* strategies, coll. E. Giraud-Héraud, Y. Le Roux, A. Schiavina; extensification of agricultural productions, coll. J.P. Butault, A. Barkaoui - ESR Nancy, J. Mathurin)
7. Commissariat Général du Plan & Direction de la Prévision, 1992-1993, convention d'étude CGP n°6/1992, prestation DP 920060
8. Bologna University, 1993-1996 (*AROPAj* Italy, coll. A. Schiavina, 3 alloc. COMETT)
9. Confédération Paysanne, 1993-1995 (funding for internships, coll. G. Le Moguédec)
10. AIP Inra "Nouvelles fonctions de l'agriculture", 1994-1996 (Interactions between agriculural policies and aforestation incentives, coll. ESR Nancy, A. Birfet, A. Hofstetter)
11. Demeter, 1995
12. Funding from MinAgri, 1994-1997 (PhD Ing. Agron. J. Mathurin)
13. AIP Inra "ALIMAN", 1995-1997 (modeling of cereals and feed markets, coupling *AROPAj* - industrial feed formulation, coll. Y. Dronne ESR Rennes, O. Lapierre INAPG et CEREOPA)

14. UP Madrid, 1995-1997, (PhD co-supervising with Université de Bourgogne, C. Chaya)
15. Contrat cadre Inra-Inia (Spain), 1996, support of the previous project
16. Sido, 1996-1998 (coll. Y. Dronne, O. Lapierre - module "feed" and partial equilibrium approach, *MODANI-AROPAj*)
17. Mission Interministérielle sur l'Effet de Serre, 1996-1997 (initiating *GHG* emission studies), lettre de commande n° 14/96.
18. Bologna University, 1999, grant Leonardo
19. Partenaire (initiateur) programme EUROTOOLS, FAIR5 CT97-3403, 1998-2000 (*AROPAj-UE*-12)
20. GICC (MinEnvir & MIES), 1999-2001 (S. De Cara, coll. Inra Avignon related to the *STICS* model)
21. GICC (MinEnvir), 2001-2003 (initiating of spatialisation, coll. Soil Siences Inra Orléans)
22. DG AGRI (Com. Europ.), 2001-2002 (*AROPAj EU*-15, reforming the milk quota system viewed through in *AROPAj*)
23. GICC (MinEnvir), 2002-2004 (initiating the coupling *AROPAj-STICS*, coll. Inra CSE-Avignon & Infosol-Orléans)
24. Ademe, 2002-2005 (PhD grant, C. Godard, coupling *STICS-AROPAj*)
25. Direction de la Prévision, 2002-2003 (*AROPAj* and decoupling in the *CAP*)
26. STREP FP5 GENEDEC-502184, 2004-2007 (coordinating 10 teams in 6 *MS*, support and exchange PhD, coll. *JRC* and Universities, Parma, Madrid, Reading, Rethymno, Verona, Dublin, FAL)
27. Inra-MinAgri, 2004-2007 (funding of a PhD, E. Galko)
28. Strep FP5 INSEA, 2004-2006 (partner, Iiasa coordinating)
29. GICC-2 (ministère de l'environnement), 2004-2006 (*GHG* emissions, climate change, regulation & mitigation, initiating analysis of adaptation, coll. LSCE, N. de Noblet)
30. ADD-DST, 2005-2008 (impact of soil compaction, coll. F. Guillet, J. Roger-Estrade, G. Richard, Inra-Orléans)
31. PIREN-Seine, 2007-2011 (PIREN phase V, coll. ParisVI-Sisyphe, Mines-Géosciences, Inra Mirecourt; initiating water issues analysis with *AROPAj*, water and nitrogen pollutants, coupling *STICS-MODCOU-AROPAj*)
32. Ifp, 2007-2011 (coll. Ifp, PhD grant, M. Clodic, *AROPAj EU*-25 & bio-fuels)
33. Ademe-CEA, 2008-2011 (Climate change, European agriculture and mitigation - adaptation, PhD grant, D. Leclère)
34. R2DS (Ile-de-France region), 2009-2012 (PhD grant, C. Bourgeois, aquifer dynamics & joined pollutions)
35. "Futurol" (INRA, Ifp), 2009-2012 (PhD grant, N. Ben Fradj, *AROPAj* & perennial crops dedicated to bio-fuels & environmental impacts)
36. Eau-dyssée (CNRS-"ec2co"), 2009,2010, (grant dedicated to model coupling, coll. F. Habets Sisyphe-MinesParisTech)
37. CC-TAME FP7-KKBE-2010-4 (partner, coll. S. De Cara, *AROPAj* & *GHG-MAC*s, initiating the *AROPAj* version *EU*-27)
38. ANR-VALIDATE, 2008-2011 (grasslands, livestock & climate change)

39. ANR-VULNOZ, 2008-2012 (ozone impacts on agricultural, coll. D. Leconte, P. Clerino, P. Humblot, coll. Inra-EGC, LSCE)
40. ANIMALCHANGE FP7, 2011-2014 (coord. J.F. Soussana, PhD grant co-funded by ADEME, P. Aghajanzadeh-Darzi, climate change & livestock)
41. PIREN-Seine phase VI, 2011-2015 (coll. Sisyphe Univ. ParisVI & MinesParisTech, policy analysis & directives dedicated to the environment)
42. ANR-ORACLE, 2011-2016 (land use, water, agriculture, grasslands and forest, coll. INRA-LEF, LSCE, Mines-Géosciences, PhD funding for A. Lungarska)
43. Ademe "Pro-Extern", 2011-2014 (coll. Veolia, Ineris, Inra-EGC, analysis of *UOR* spreading and nitrogen valuation)
44. MESR, ENS PhD funding (P. Humblot, 2012-2016)
45. ANR-AZODURE, 2013-2017 (coord. Univ. Lyon 1, EHC funding / M. Bounaffaa)
46. FoodSecure FP7, (Inra partner, 2013-2016)
47. ABIES, PhD funding (G. Dumollard, 2013-2016)
48. BASC (Labex, coll. Univ Paris-sud Orsay, LSCE, 2014-2016, 2017)
49. "Futurol" (INRA), 2015-2016 (raw matter for bio-fuels in France, N. Ben Fradj, coll. UMR EcoSys)
50. PIREN-Seine phase VII, 2016-2019 (coll. Metis Univ. ParisVI & MinesParisTech)
51. ASTREA (Ile-de-France region), 2016-2018 (PhD grant, A. Isbasoiu)
52. LabEx BASC-STIMUL (2016-2018, supervised by R. Chakir)
53. Institut de Convergence CLAND (2018-2023)
54. LabEx BASC-APISMAL, 2018-2019, supervised by V. Martinet)
55. Meta-programme ACCAF & Chaire Economie du Climat (2018-2021)
56. Ademe - APR GRAINE, 2018-2022 (supervised by D. Lorne, Ifpen)
57. ANR-Diet+, 2018-2021 (supervised by S. Marette)
58. Tender JRC-Seville (2019)
59. ANR-PPR 2020 (FAST project supervised by J. Subervie) (2021-2023)
60. Belmont project BLUEGEM (supervised by A. Ducharne) (2021-2024)
61. Ministère de la Transition Ecologique (conj. Min Agric., 2022)
62. TotalEnergies (2023)