MIND STEP

MODELLING INDIVIDUAL DECISIONS TO SUPPORT THE EUROPEAN POLICIES RELATED TO AGRICULTURE

Deliverable 3.2: An overarching IDM model structure

Interfaces within the MIND STEP model toolbox

AUTHORS	Marc Müller (WR), David Schäfer (UBO), Wolfgang Britz (UBO), John Helming (WR)
APPROVED BY WP MANAGER:	John Helming (WR)
DATE OF APPROVAL:	27-02-2023
APPROVED BY PROJECT COORDINATOR:	Hans van Meijl (WR)
DATE OF APPROVAL:	28-02-2023
CALL H2020-RUR-2018-2	Rural Renaissance
WORK PROGRAMME Topic RUR-04-2018	Analytical tools and models to support policies related to agriculture and food - RIA Research and Innovation action
PROJECT WEB SITE:	https://mind-step.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement N° 817566.

EXECUTIVE SUMMARY

Following up on the previous **MIND STEP** deliverable 3.1 on the "Specification of model requirements: Protocols for code and data", this deliverable 3.2 provides an outline on how a modular approach to model integration was realized in practice. The focus is on the interactions of methods and results developed in tasks within the **MIND STEP** work package 3 (WP3), titled "Development of modular and customisable suit of models focussing on the IDM farming unit". These tasks apply rather heterogeneous methods, including micro-econometric analyses of crop-management choices, behavioural aspects of technology adoption, or risk-preferences, as well as the usage of farm-level simulation models for ex-ante impact assessment of policy and technology options. Integrating these approaches requires a conceptual structure that defines the interfaces between them and exploits the exchanges of data and methods as much as possible. An important conceptual decision was to select a farm-level simulation model as the integrative core of the overarching framework and to develop interfaces to empirical and methodological works accordingly. Chapter 2 starts with a discussion why the simulation model FarmDyn (Britz et al. 2016) was selected as the core model in this case. An in-depth survey of four applied farm-level models (Britz et al. 2021) showed that FarmDyn already follows modular design principles, which facilitates its extension based on the results from empirical and conceptual work in **MIND STEP**. The detailed representation of farming activities, like the wide range for feeding options, the endogenously selectable intensity levels for crops, the explicit representation of machinery and buildings, and the detailed representation of variable inputs also provide direct links to the different tasks in **MIND STEP**, like the exploration of GHG mitigation options or the inclusion of risk preferences. In the subsequent chapter, conceptual linkages of FarmDyn to the different tasks within WP3 are explored and the concept of modularity is highlighted.

Chapter 3 then summarizes how the proposed concepts are implemented within a modular framework. The preparation of data for the Dutch version of FarmDyn is explained, which is then used to analyse technical and behavioural aspects of the adoption of GHG mitigation options. This application includes additional information based on a survey of Dutch dairy farms and newly introduced GHG mitigation measures, which was part of Task 3.3. Possibilities to connect the micro-econometrically estimated crop-management and cost-allocation models developed in task 3.4 are then also discussed, followed by an outline of the work on risk preferences at farm level in Task 3.5. In this context, a new risk module for FarmDyn was developed, which permits the direct inclusion of empirical findings on risk preferences and risk mitigation measures in model-based simulations.

MIND STEP, together with the sister projects AGRICORE and BESTMAP, form the AgriModels Cluster, which aims at better understanding the impact of individual decision making in agriculture on the uptake of policies and technologies. Since inception, the three project have interacted closely with regard to data acquisition and dissemination of results. As an outcome of jointly organized sessions during past seminars, for instance a simulation model from the AGRICORE project became part of the combined modelling activities in **MIND STEP**. Based on this experience and the work related to establishing an overarching framework for model integration in **MIND STEP** WP3, this report concludes with an elaboration on the importance of a network of model developers and users for the continued development of a generic and flexible overarching structure with a farm-level simulation model at its core.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

1. INTRODUCTION

Building an overarching structure for individual decision-making models (IDM) at farm level is a declared objective of the **MIND STEP** project. Conceptual work on the design of a modular system that permits flexible integration of empirical work and newly developed features of simulation models has started right at the beginning of the project. The most important aspects of modularity in the context of farm-level modelling have been discussed in Britz at al. 2021, including the importance to select a core model, to which new modules may be linked. This also implies that a clear definition of obligatory inputs and outputs (interfaces) is necessary to ensure that the equations in the module can be executed, e.g. by providing default values for all parameters. The technical documentation of core model and modules, and the development of protocols for contributor should receive particular attention from the very beginning if model development and maintenance is to be distributed across multiple teams. Definition of inputs and outputs, protocols for module development, and general aspects of quality management have been put forward in **MIND STEP** Deliverable 3.1 on “Specification of model requirements: Protocols for code and data” (Müller at al., 2021). Building on the conceptual work by Britz at al. (2021) and Müller et al. (2021), the present report outlines how an overarching structure for the integration of the farm-level simulation model FarmDyn (Britz et al., 2016) with new modules and empirical work within **MIND STEP** has been realized.

Overall, the idea of an overarching model is here neither strictly interpreted as one single operational stand-alone IDM model, i.e. a one model fits all approach, nor solely as a conceptual idea without concrete implementation steps. Rather, a modular structure in which different IDM models present in **MIND STEP** become more flexible, modular and general is proposed, facilitating two-way loose links between the different IDM models, including FarmDyn, covering the possibility to apply IDM models more easily to different farm samples.

Chapter 2 first provides a discussion on the selection of FarmDyn as a core model, followed by a description of relevant aspects in FarmDyn regarding the connection with other models as well as the elements identified in the initial proposal and the extent by which FarmDyn could integrate findings and developments from different work packages envisaged in **MIND STEP.** This moves FarmDyn closer to a more generic, modular and flexible IDM (Britz et al. 2021) and thus towards a kind of an overarching model within the **MIND STEP** toolbox. Next, possibilities to let FarmDyn inform other IDM models are explored. It will then provide further insights in specific aspects of the model, presented in the section behavioural model, technology choice, interface to policies. In addition, more technical aspects of FarmDyn are discussed in the calibration and technical implementation section. Chapter 2 concludes with a discussion on how concepts like generic, modular, and flexible have to be defined in the context of an IDM model.

Chapter 3 then gives an overview on the actual implementation of this overarching structure in the **MIND STEP** project. It is structured along the different tasks within work package 3, namely tasks 3.3 concerning the implementation and application of FarmDyn for GHG mitigation options in the Netherlands, task 3.4 concerning the integration of crop-choice models, and 3.5. on risk preferences and -management.

As the **MIND STEP** project is part of the AgriModels cluster, it is desirable that the work on an overarching structure for IDM modelling permits also the integration of models developed in the AgriCore and BESTMAP projects. This is addressed in chapter 4.

The conclusions on chapter 5 focus on the experiences made across the different modelling teams working towards the realization of the overarching modelling structure. It highlights the importance of a network of model developers and users, because of the increased complexity and reduced transparency and tractability of models that comes with the increased flexibility.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

2. TOWARDS A MODULAR MODEL STRUCTURE

2.1. Selection of a Core Model

The modular design principles put forward by Britz et al. (2021) and Müller et al. (2021) distinguish between core model and contributed modules. The core model should provide a minimum set of equations and the required data for parameterization, as well as possible interfaces to contributed modules. This implies that the core model should already follow modular design principles and have features, to which additional functionality can be added or which can be replaced by new developments. Two IDMs are currently available within the **MIND-STEP** consortium: IFM-CAP (Louhichi et al., 2017) and FarmDyn (Britz et al., 2016). Both were included in the detailed model review by Britz et al. (2021), where certain aspects of model implementation and content features are compared. Both are mathematical programming models, where an objective function is optimized subject to a set of constraints. IFM-CAP features a quadratic objective function and linear constraints, whereas FarmDyn is completely linear, but includes integer variables. Important content features of both models are shown in Table 1. With regard to regional coverage, IFM-CAP is designed to run for all individual farms included in the EU-wide FADN dataset, thus providing a wide regional coverage. In contrast, FarmDyn provides richer detail for the representation of farm technology, policy measures, factors of production like labour, machinery, and buildings, as well as feed and fertilizer options. This level of detail requires substantially more data than is available from EU-FADN and applications of FarmDyn are therefore restricted to regions, where the needed information is available. Still, it is possible to construct case-study farms for a larger number of regions of the EU for FarmDyn by using EU-FADN to identify typical or representative farms and relying on other databases for the parameterization of the technology-related parts of the model. From a **MIND STEP** perspective, three blocks of functionalities are particularly important: First, the nutrient and GHG accounting, in combination with the possibility to select alternative intensity levels for crop production and flexible representation of animal feeding. This is important for the analysis of policy and management options aiming at the reduction of GHG and nutrient emissions from agriculture, as addressed in Tasks 3.3. FarmDyn distinguishes between alternative input-levels for each crop and permits to include maximum and minimum rotation shares. This allows for the usage of empirical results regarding choice of input regimes and crop allocation in Task 3.4 because of the very detailed representation of such crop management choices in FarmDyn. Secondly, as many GHG mitigation options also require long-term investments in machinery and buildings, more detail for types of machinery, their attributes and costs are desirable. In addition, permitting endogenous investment decisions, ideally in a multi-periodic setting, are useful features for policy and technology-related scenarios, e.g. to identify under which conditions farmers would switch to less emission-intensive but more costly farm equipment. As can be seen from Table 1, FarmDyn provides such features already and is therefore a good candidate for the core model from this perspective. A third important feature is the representation of attitudes towards risk at farm level and risk-mitigation options, which is the main topic of Task 3.5 in **MIND STEP**. As IFM-CAP has already a quadratic objective function, a typical expected-value – variance approach could be implemented without changing the model structure. FarmDyn permits the usage of several linearized decision rules for risk utility, most recently including cumulative prospect theory (Britz 2022). This is well aligned with the work carried out by the teams involved in Task 3.5.

Apart from content-related model features, Britz et al. (2021) also review the extent to which the models already follow a modular set-up. All of the compared models separate model code from data generation. The modularization of equation blocks that are used in the model statement is particularly observable in FarmDyn, which is structured along functional units of code for farming activities (e.g. equations specific for dairy farming) and permits the inclusion of alternative policy modules, e.g. for fertilization regimes.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

Based on these considerations, it appears that FarmDyn is well suited as a core model of the overarching structure, to which newly developed features within **MIND STEP** can be added. The following sections will provide an overview on the structure of FarmDyn in more detail and some conceptual considerations for the use of FarmDyn within the **MIND STEP** overarching structure.

Table 1: Comparison of models with respect to content features

Model	IFM-CAP	FARMDYN
Regional- and product coverage; technology representation		
Regional coverage in the EU	covering EU	5 EU countries + Switzerland
Coverage of farm population	almost full	case studies
Individual farms	yes (≈80,000 farms)	yes
Representative farms	yes	
Max. no. of crop & animal activities	35/16	case specific
Max. no. of crop & animal commodities	30/7	case specific
No. of animal activities per commodity	single	multiple
Catch or cover crops	yes	
Grassland management	two type of grass	grazing only or also for cuts, different management options*
Herd flow representation Herds (by age, sex, year, months)	yes	yes also breeds, feed regime
Manure types	no	several
Management and technology options for activities	no	Tillage options, intensity levels**
Crop rotation	yes	possible
Temporal resolution	year	Year, month
Plot representation (land quality)	no	yes
Policy representation		
Direct payments and common organisation of the markets in agricultural products (CAP Pillar 1) (+ set-aside and quota, EU's greening reform)	Voluntary coupled support & ceiling of EU budget endogenous	yes
	diversification as binary decision	
Nitrate and Water framework directive	no	in high detail
GHG policies (CO2 pricing, ceilings)	no	yes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

Trade and market policies, Tariff Rate Quotas, Tariff cut (e.g., liberalization, WTO G20 proposal)	prototype link to CAPRI	
Link to other types of models	market model	link to crop growth models possible to generate I/O coefficients
Factors		
Covered types of land endowments (arable/grassland/permanents)	all	no perm.
Labour constraints	no	block labour for management of farm and branches, available field working days, off-farm labour (fractional or in integers)
Temporal resolution of simulation steps towards baseline scenario	instantaneous	yearly, monthly
Machinery items considered	no	multiple
Buildings considered	no	multiple
Land markets	no	Lease and buying options
Emission accounting & Indicator calculation		
Climate change adaptation	No	Scenario dependent
Environmental indicators	Intensification / extensification, pesticide risk, soil erosion, soil organic matter, crop diversity	N- and C-emissions, N- and P balances, GWP**, link to LCA with many indicators
Feed and fertilizer representation		
Feed activities	tradeable and non-tradeable fodder activities	tradeable and non-tradeable fodder activities, including grazing; feeding regimes distinguished by herds, breeds, feed regime, feed, year, months
Feed constraints	energy, protein, min/max shares	
Attributes of fertilizing activities	constant intensities	crops, plot, tillage, intensity, fertilizer, year, month
Temporal resolution & Investments		
Multi-period optimization	no	yes
Endogenous investment decisions		
Financial constraints		
Economic behavioural assumptions and calibration		
Objective function maximization	farm utility Including risk component	maximization of discounted profit withdrawals plus returns from off-farm labour, after taxes. Stochastic setting: Several decision rules for risk utility possible

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

Model type	QP+ & MIP++	MIP
Calibration approach	PMP++++	Bi-level
* grazing, silage, bales, hay; by month; ** GWP: Global Warming Potential; + Quadratic Programming; ++ Mixed Integer Programming;+++ Positive Mathematical Programming		

Source: Derived from Britz et al. 2021

2.2. FarmDyn structure and the connection to identified IDM elements

As mentioned in the previous chapter, FarmDyn comes already equipped with a modular and flexible template structure, where modules are turned on and off depending on the needs of the user. The user has the possibility to adjust specific details of model with respect to different farming systems (dairy/mother cows, beef fattening, pig fattening, piglet production, arable farming, biogas plant). Further, it comprises multiple options with regard to time dynamics (including fully dynamic, comparative static or short run) and offers an annual and sub-annual temporal resolution depending on the decision variable. The farmer's behaviour is based on an optimization approach, either maximizing the NPV of returns to farm assets (including working off-farm) in the deterministic approach or using stochastic dynamic programming to maximizes the expected NPV. The latter approach can be extended to cover risk behaviour based on different options (value at risk, MOTAD…). In addition, FarmDyn comprises a wider range of environmental indicators including N and P balances, GHG emissions, a nitrate leaching indicator, different protein and calories indicators, partly also considering emissions from up-stream industries in bought inputs. The modular structure is shown in following figure.

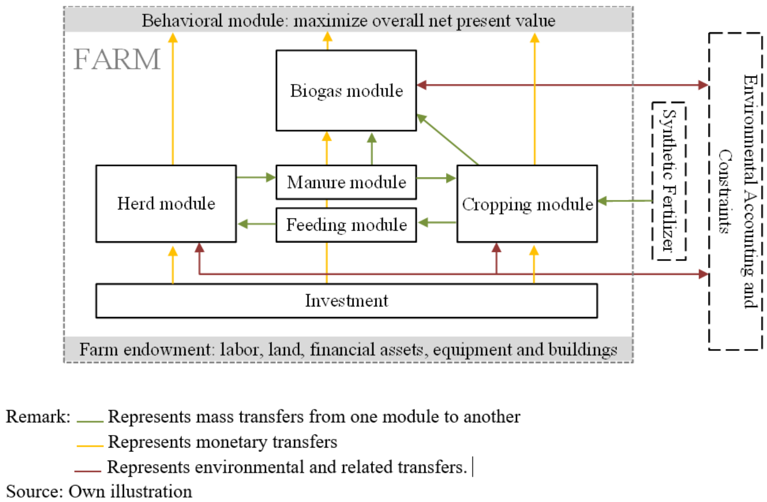


Figure 1 Schematic FarmDyn Structure

Source: Britz et al. (2016)

The implemented technology is detail rich and comprises, inter alia, farm labour needs, machinery, stable and manure silo use.

In the **MIND STEP** proposal, important IDM elements for WP 3 are already identified and shown in the "honey comb" figure and will be most likely expanded and well defined in the working of the WP 3.2. To get an idea on the possible linkages between the initially defined IDM elements and the modular

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

structure of *FarmDyn*, we linked each of the IDM elements to existing modules in *FarmDyn* which are clustered in three groups in the following table.

Table 2: IDM-elements and links to MIND STEP tasks

IDM elements	FarmDyn thematic groups			Potential link to IDM models in WP3 Tasks
	Behavioural module	Farm economic module	Technology modules	
Risk management	X			Task 3.5
Profitability and viability		X		
Investment/Financing		X		
Cost accounting		X		Task 3.4
Feed			X	
Yield			X	Task 3.4
Rotation / Land use			X	Task 3.4
Crop management			X	Task 3.4
Greenhouse Gasses			X	Task 3.3
Labour use		X	X	

Table 2 provides thematical FarmDyn groups and linkages to the IDM models in WP3 and the included tasks. Here, FarmDyn as a partly overarching model for other IDM models can be used in two ways. One the one hand, as a source of input data by generating huge data sets (experience can be drawn from previous work presented in the next section and will be directed to the work in Task 4.5 for machine learning). On the other hand, generated output of the Tasks 3.3-3.5 can be used to improve FarmDyn in different ways.

2.3. FarmDyn as a data source of meta-models

FarmDyn has been used in several applications as a model to generate data which is further used in a meta-model. In the applications, a meta-model was defined as a model which quantifies major input/output relationships embedded in the structure of the complex model by using standard statistical techniques on these results. The standard statistical techniques can here refer also to more complex econometric exercises such as in MIND STEP Tasks 3.3-3.5. Examples of the applications can be found in Lengers et al. (2014), Kuhn et al. (2019a), Kuhn et al. (2019b).

2.4. Potential improvements of FarmDyn through other IDM models

2.4.1. Behavioural model

MIND STEP considers multiple behavioural models which encompass risk (Task 3.5) or even multi-criteria utility functions (Task 3.3). FarmDyn is based on an optimization approach which clearly limits depicting non- rational behaviour. The only currently discussed option to move away from strict

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

rationality is to use a weighted sum of household income (or a risk behavioural model) and absolute deviation from a given past farm program to consider cautious behaviour which consequently would no longer be fully technical and allocative efficient. While that is relatively simple to build in, a realistic parameterization might be challenging. Results from **MIND STEP** Task 3.3 may provide the needed insights.

If additional arguments beyond household income (after income taxes) are to be represented in a utility function in FarmDyn, the limitation to linear constraints and a quadratic strictly convex objective function restricts possible extensions. It is however relatively easy to optimize one attribute of the utility function such as income under minima for others such as different indicators of environmental quality. The resulting sample of points allows to meta-model the multi-dimensional frontier of the decision space with regard to the indicators considered. That in turn can be used in combination with a more complex utility function in another IDM, for instance based on the work on GHG mitigation preferences as described in Helming et al. (2023) and Wang et al. (2022). Here, the recent implementation of many different environmental indicators into FarmDyn provides a good starting point if environmental considerations shall become part of a more flexible utility function.

Different risk behavioural models are already built into FarmDyn, currently all linked to a stochastic dynamic programming approach which is extremely demanding from a computing perspective. It is somewhat doubtful if these approaches can be used for larger farm samples in MINDSTEP. One could integrate instead a linearized version of the classic mean-variance model in FarmDyn or another linearized form of risk utility, like Tversky-Kahnemann utility (Tversky and Kahnemann, 1992) utility as developed in Task 3.5 and integrated in FarmDyn by Britz (2022).

2.4.2. Technology choice

Technology choice is driven by the behavioural model and reflects the representation of the choice set in an IDM. Based on linear programming, FarmDyn requires distinct Leontief combinations of netputs to represent the overall technology. The netputs partly refer to quasi-fixed factors. FarmDyn avoids the notion of “unobserved” costs often found in PMP based models and uses a typically quite detailed technology representation linked to a rich constraint set. For instance, the field calendar for crops depicted specific crop operations is present in a two-weekly resolution during the growth season and linked to machinery, intermediate input (such as diesel, plant protection products …) and labour needs and available field working days. For each crop, multiple tillage options (plough, reduced till, no till) in combination with different intensity levels can be introduced. Nutrient needs of the crops can be covered by different types of mineral and organic fertilizer for which different application technologies are depicted. Feeding is depicted at a monthly resolution and linked to a detailed representation of requirements for the different animals, reflecting for cows the lactation phase. A flexible herd dynamics model for cattle captures endogenous choice of calving month, cross-breeding and sexing, breaking growth processes in flexibly defined multiple growth periods. Competing options can be defined as well. Multiple grassland management strategies which capture the use of biomass for grazing, grass silage or hay in each month can be introduced. There is a distinction between variable labour needs linked to field and herd operations and blocked size management labour related to certain farm branches and the farm operation in total. Technology choice in FarmDyn is also linked to a detailed presentation of investments into machines and structures. Parameterization of these aspects implies data needs well beyond what can be found e.g. in FADN.

The development of detailed GHG mitigation options not yet captured by FarmDyn in **MIND STEP** Task 3.3 fits well to the concept of FarmDyn. Estimation approaches to represent production functions such as used in Task 3.4 could be used to derive representative points for input-output combinations, which could be integrated into FarmDyn using the option to endogenously select intensity levels for crops. The technology choice set actually covered by an IDM model is clearly linked and can be informed by

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

policy measures to analyse, such as, for instance, the national implementation of the Nitrates directives or measures from the second pillar, a point which is touched upon in the next section.

2.5. Interfaces to policies

An important reason to move from an aggregate agent paradigm to IDMs is to better represent policy impacts emerging from policy measures which depend on single farm attributes, be it (quasi-fixed) factor endowments or details of farm management. Examples provide the greening measures under the first pillar of the CAP, opt-in measures from the second pillar or agri-environmental command-and-control measures as often found under the implementation of the Nitrates or Water Framework Directives. In many cases, integer variables are required to correctly depict these policies, due to various thresholds and if-else conditions. FarmDyn so far covers the greening condition of the current CAP and measures related to the German implementation of the Nitrates and Water Framework Directives, and a rudimentary implementation of support to organic agriculture. Coupled support to specific crop and cattle processes can also be introduced. Its detailed technology representation eases and allows analysing such measures in quite some detail (see e.g. Kuhn et al. 2019a, Kuhn et al. 2019b, Schäfer et al. 2017).

Introducing agri-environmental opt-in measures from the second pillar is quite data demanding. It will typically require introducing specific crop and herd management options fitting to selected measures.

2.6. Calibration

All the points above need also be considered in model calibration which encompasses both benchmarking of the model to one observation of given farm data, for instance a record from FADN, and but also to observed allocative changes. The ease of benchmarking is a core reason for the popularity of the PMP approach, for instance used in IFM-CAP. Drawing on ideas borrowed from PMP, FarmDyn now comprises an algorithm based on bi-level programming which can automatically adjust different parameters in the model such as expected yields and prices, labour need or feed requirements to calibrate an instance of the model for a single farm against given data without introducing unobserved costs or revenues. The approach was so far successfully tested in FarmDyn for the comparative-static case where just one observation must be recovered. Simultaneous calibration to multiple points clearly is a more recommendable approach as it would allow tracking allocative behaviour, for instance, by calibrating parameters related to risk behaviour or cautious decision taking. However, this requires probably a change in the current objective function and is therefore linked to discussion around alternative behavioural models.

2.7. Technical aspects

The points touched upon above are also closely linked to the code implementation of an IDM model. The application to different farm samples, potentially sourced from different data sets, requires a flexible data interface which feeds information on the behavioural model, on the considered technology set and, potentially, on calibration targets into the IDM model. FarmDyn has now separated data quite strictly from processing code.

FarmDyn follows already modular design principles as discussed in Britz et al. (2021) and also in the FarmDyn documentation: different farm branches (arable, dairy, mother cows, beef, fattners, biogas) can be switched on or off for individual instances of the model, governed by a data-file containing individual farm settings. Recent coding efforts aim at a more data driven representation of the technology such that list of available crops, definition of field operations for these crops under different tillage systems and related machinery and time needs etc. can be flexibly changed without the need re-program model equations. First steps have been taken to also allow to exchange code

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

blocks in the model, for instance, to represent policy measures governing nutrient management related to Nitrate and Water Framework Directive.

2.8. Modular Design

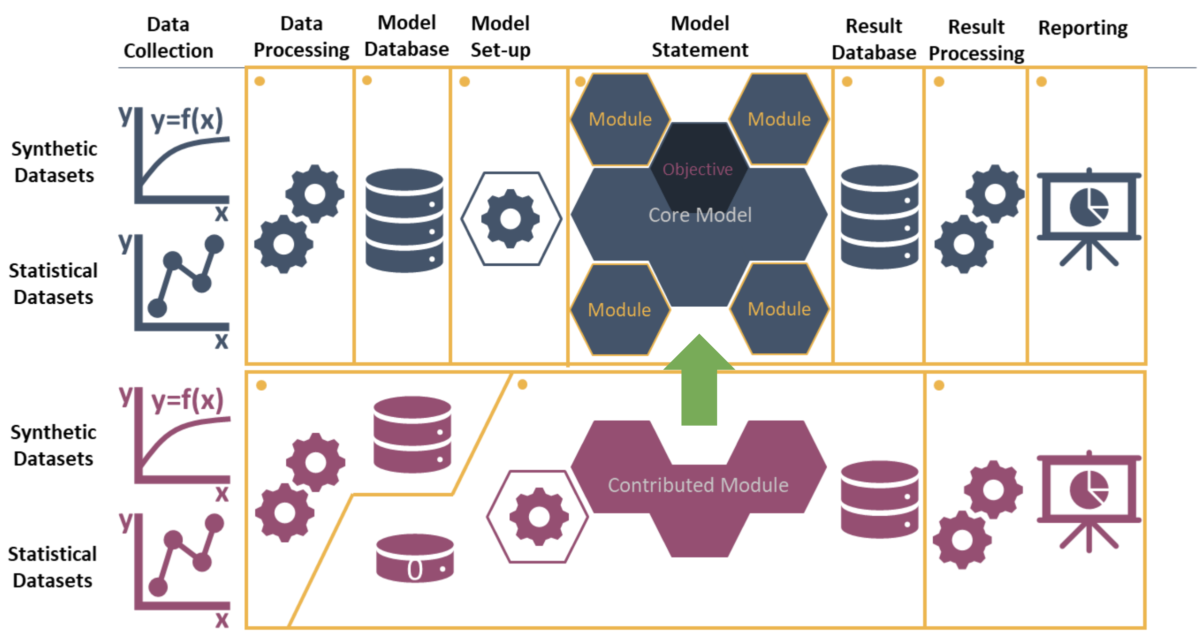
The aspects of FarmDyn mentioned above are also addressed by Britz et al. (2021) Müller et al. (2021). In there, the typical workflow when setting up a simulation model like FarmDyn is depicted as in *Source: Based on Britz et al. (2021)*

Figure 2, namely processing of statistical data or data from other sources into a model database, the model parameterization, the combination of blocks of equations into a model instance, and reporting of results. Modularity here means that certain units of this workflow can be replaced by the user without changing the structure of the model, provided that the new features comply with certain rules. A distinction between native and contributed modules is useful here (*Source: Based on Britz et al. (2021)*

Figure 2). By definition, a native module (the hexagons labelled: “Module” in *Source: Based on Britz et al. (2021)*

Figure 2) can be always fully parameterized from the general model database, while a contributed equation module offering additional functionalities (the purple block of hexagons in *Source: Based on Britz et al. (2021)*

Figure 2) might require additional data which it must provide by own code for data preparation. The same holds for the reporting step.



Source: Based on Britz et al. (2021)

Figure 2 Modular Setup

The most crucial aspects for design and integration of modules in such a setting are the clear definition of obligatory inputs and outputs (interfaces) and ensuring that the equations in the module can be executed by providing default values for all parameters. This also implies that the technical

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

documentation of core model and modules, and the development of protocols for contributor should receive particular attention from the very beginning if model development and maintenance is to be distributed across multiple teams with high staff turnover rates.

Modularity also needs to reflect user-model interaction. Three models reviewed by Britz et al. (2021) (CAPRI-FT, IFM-CAP, FarmDyn) feature a GUI, all realized in GGIG (Britz 2014), to facilitate, for instance, choosing the included modules or the data base to use. An important question is to which extent a specific model configuration (farm branches, activities covered, specific policy implementation etc.) is driven by the data base or defined by user interactions. Second, to what extent should the user be able to provide (or overwrite) via the GUI data otherwise read from the model data-base, such as, e.g. run specific prices, yields or values of policy measures. Third, should the GUI also cover such functionalities for contributed modules? If yes, how is this technically achieved and institutionally organized?

An at least partial answer to this question is the concept of wrapper functions pursued in **MIND STEP** (see e.g. Müller et al. 2022 for a concept of the **MIND STEP** wrappers). These wrappers are functions implemented in the R programming language that permit the execution of a suite of models, and the exchange of data, from an R environment. While exploiting specific features of the GUI mentioned above, it also permits the generation and use of additional input data in the model without the user having to work directly with the GUI.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

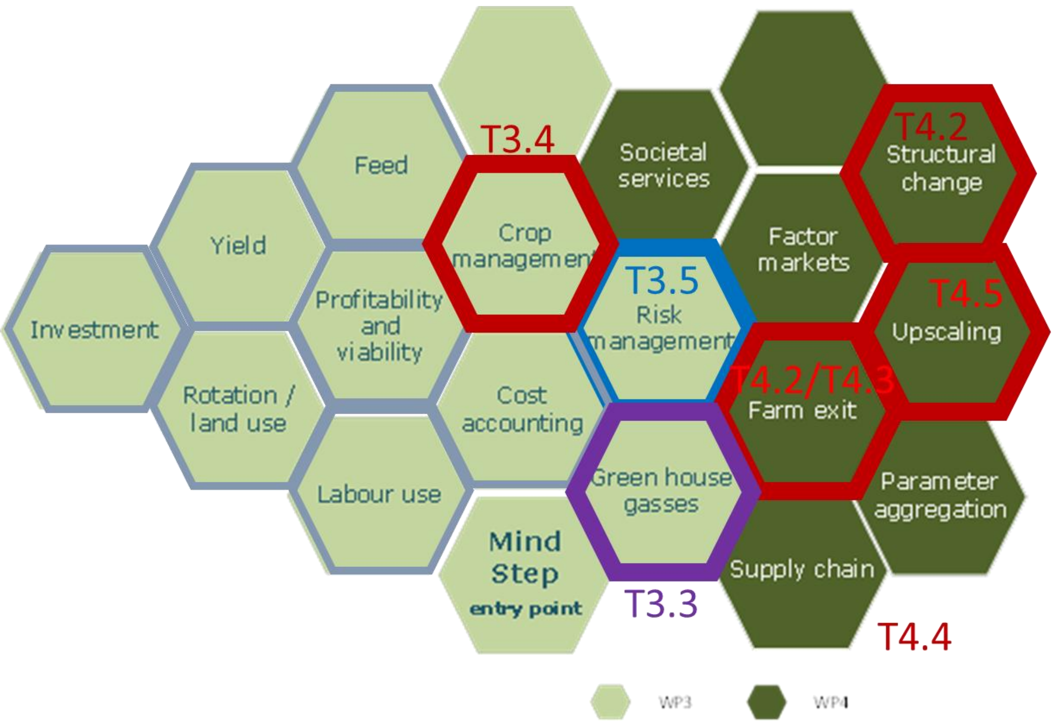
3. COUPLING EMPIRICAL AND SIMULATION MODELS AT FARM LEVEL

3.1. MIND STEP Tasks within the overarching model structure

Within **MIND STEP's** WP3, the role of Task 3.2 reported here is to establish an overarching structure that integrates methods and results from the other involved Tasks. Following the considerations in Britz et al. (2021) and Müller et al. (2021), this is realized by identifying a core model and structures for exchange of information with the other tasks, as outlined in chapter 2. The **MIND STEP** Tasks 3.3, 3.4. and 3.5. are then conceptually linked to this core model and either contribute to it by providing improved input data or modules with additional functionality. Figure 3 summarizes the interactions within WP3. Task 3.3 (purple hexagon in Figure 3) investigated new GHG mitigation technologies and behavioural aspects of their adoption at farm-level, and thus improve the representation of farmer's decision making in the core model. The improved core model was used within Task 3.3 to derive strategies for farm-level decisions to mitigate GHG emissions under a range of policy and technology options (Helming et al., 2023). Task 3.4. adds empirical findings on area allocation and intensity levels for arable crops, which can be used for the core simulation model by creating farm-specific cost allocations to selected crops within a certain crop regime. Finally, Task 3.5 (blue hexagon in Figure 3). investigates risk-management behaviour and risk-management instruments for farmers. While the former refers to an improvement of the objective function by including cumulative prospect theory, the latter increases the number of decision variables by adding the usage of RMIs.

These interactions between the different tasks in WP3 are structured around the concept of core model and modules as depicted in *Source: Based on Britz et al. (2021)*

Figure 2 and also applied in the documentation of FarmDyn (https://farmdyn.github.io/documentation/FarmDynDocumentation/KeyFeatures/modularity/).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

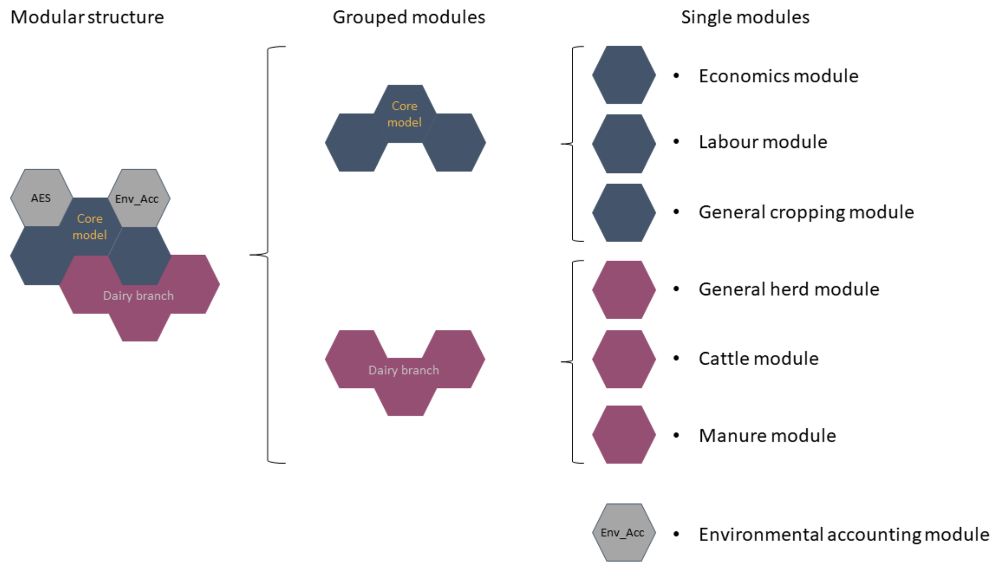
Figure 3 Towards the honeycomb: Integration of Tasks

3.2. Linkages to the core model

The MIND STEP honeycomb implies already a variety of functionalities that should be covered by the core model within the overarching model structure, for example endogenous adjustment of yield, selected variable inputs, animal rations, land and labour use, investment and farm viability (the colour-framed, light green hexagons in Figure 3). The previous sections have shown that the FarmDyn model is capable to address these aspects and is used as core model, although the main concepts put forward in the subsequent chapters will apply also to comparably structured IDM models, for instance IFM-CAP. Following the logic of *Source: Based on Britz et al. (2021)*

Figure 2, modular enhancements of the core model can be achieved by improving and extending the model database without making changes to the core model equations themselves, or by contributing a module that extends the original model functionalities.

In line with this broad definition and the description of preferred generic and modular implementations of bio-economic farm-scale models (Britz et al. 2021), FarmDyn is structured as a modular system where each module comprises a block of equations and variables which can be activated depending on the user case. Practically, each module captures a specific farm management or methodology domain within a file. Further, the definition of parameters is either done in a module if it is only used in that specific file or is organized globally in the set and parameter declaration module, templ_decl.gms. Each module can be either a standalone module or is linked to other modules and grouped together to build a specific aspect of the farm, e.g. the farm branch, or the core model. The structure of the core version of FarmDyn is illustrated in Figure 4.



Source: FarmDyn Documentation

(https://farmdyn.github.io/documentation/FarmDynDocumentation/KeyFeatures/modularity/)

Figure 4 Modular structure of the core model

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

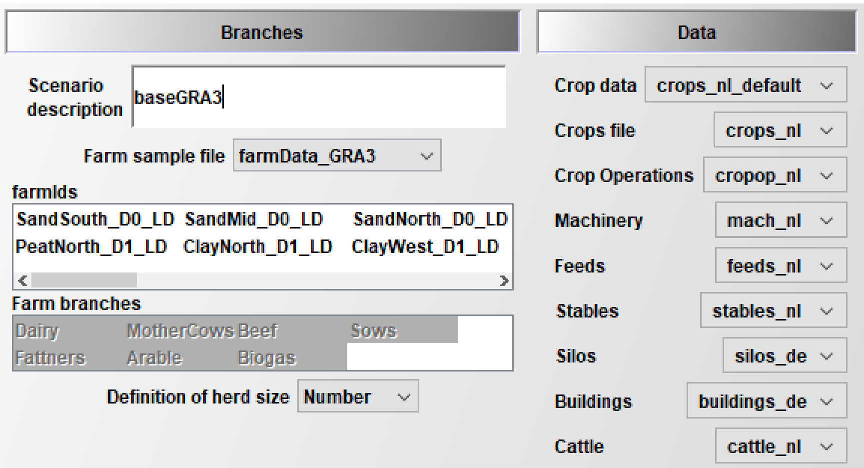
The specification of individual farm-instances in FarmDyn is to a large extent driven by the underlying database and connection between new modules and the core model hinges on the way in which new functionalities can be included in there. It is useful to distinguish between farm-specific datasets and general datasets. This distinction is not a statement about the homogeneity of heterogeneity of certain types of information across farms, but rather refers to the availability of data from typical farm statistics like FADN. Data about acreage, crop yields, herd size, or labour endowments can be usually taken from such statistics . General datasets are those typically not available from such sources, e.g. crop-specific variable inputs or nutrient contents of animal feeds, although they are in reality farm specific. The less farm-specific data is available, the higher is the reliance on regional or national average values.

In general, input data for FarmDyn takes three forms:

- Global variables, which can be text strings or numerical values, are used to govern the model set-up. For instance, if the global variable "herd" is true, then model blocks related to animal production are activated, else not. This greatly helps to keep instances of FarmDyn as small as possible by only activating the required variables and equations.
- Farm endowments: farm assets like number of persons working on the farm, available cropland, or the number of dairy cows
- Data-parameters: numerical values, technology coefficients like crop yields or nutrient demand, and prices for inputs and outputs. They can be specific for each farm or apply to groups of farms.

The distinction between global variables and data-parameters is important here because it influences the way in which the provided data is imported and processed by FarmDyn, and hence how farm-level statistics need to be arranged to serve as input to FarmDyn.

For larger samples of farms, FarmDyn permits loading all farm-specific data and settings from a file with default name "farmData_XXX", where "XXX" is a placeholder for any file-specific suffix (see "Farm sample file" in Figure 5 for a specific case). This file has to be provided in the "GAMS data exchange" binary format (GDX) and can contain data parameters as well as global variables. For reasons related to the combined use of the Python and GAMS programming languages in the set-up of FarmDyn, global variables have to be encoded numerically in the data generation process and are translated back to strings when needed.



Source: FarmDyn GUI

Figure 5 Data selection for farm samples in FarmDyn

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

The FarmDyn sample file “farmData_XXX” has to include at least the farm identifiers (“farmIds”) and a GAMS-parameter “p_farmData” containing global variables that govern the model compilation and execution, e.g. by indicating whether the farm has a dairy branch. Further information passed to the model in this instance are general boundary conditions, like the total available acreage of arable land. In addition to such global settings, data on crop yields has also be provided, not only to indicate the productivity of the arable area, but also to restrict the selection of cropping activities for a certain farm model instance. Following the logic of the schematic modular setup depicted in *Source: Based on Britz et al. (2021)*

Figure 2, this “farmData_XXX.GDX” file is an example of a “model database”. For this reason, the following sections will assume that changes to the model database will refer to changes of this particular file in the case of FarmDyn.

3.3. GHG mitigation options and farmers‘ choices

Task 3.3 in WP3 illustrates the extendibility of the **MIND STEP** model structure by applying FarmDyn to Dutch farms and by adding modules to analyse GHG mitigation strategies and behavioural aspects of their adoption. Deliverable 3.3 “Report on modelling greenhouse gas emission including adoption behaviour of farmers regarding mitigation strategies and interfaces to the **MIND STEP** model toolbox” (Helming et al., 2023) discusses the parameterization of FarmDyn for Dutch dairy and arable farms, the GHG accounting, new GHG mitigation technologies and the determinants of their adoption. To increase the number of options for on-farm feed production, a more comprehensive module for grassland management is also added.

This chapter outlines how the new developments for the FarmDyn model integrate into the overarching model structure. It first gives an overview on the construction of the database for the Dutch version of FarmDyn, followed by a description, how results from other activities within Task 3.3 are connected.

3.3.1. Database Construction Modules for the Dutch version of FarmDyn

The Dutch version of FADN, the Bedrijveninformatienet (BIN), and the Kwantitatieve Informatie voor de Akkerbouw (KWIN) databases are the most relevant sources of information to parameterize FarmDyn for Dutch circumstances. BIN is an extensive database for farm statistics that is maintained at Wageningen Economic Research. It includes a panel of around 1.500 agricultural and horticultural enterprises. The established workflow to make BIN data usable for the FarmDyn model is depicted in Figure 6: Dedicated GAMS routines (“readBINData.gms”) import the different tables and merge them into a GAMS data parameter, stored in a gdx container (here: “processedData.gdx”). This file contains data for all available farms and years. To test the outcome of this first step, several R files, usually for single-purposes, were generated. The intermediate datafile is then further processed by a GAMS file (here: “build_scenarioData_XXX.gms”) which performs largely aggregation (e.g. averaging several years), sub-setting (in case only specific farms should be included in a scenario), and so on. An important step is the generation of global variables for FarmDyn and to encode certain information which FarmDyn expects as string variables at a later stage. The result of this step is the data container “farmData_XXX” with the farmIds for which FarmDyn should be executed and the associated data.

This project has received funding from the European Union’s Horizon 2020
research and innovation programme under grant agreement N° 817566.

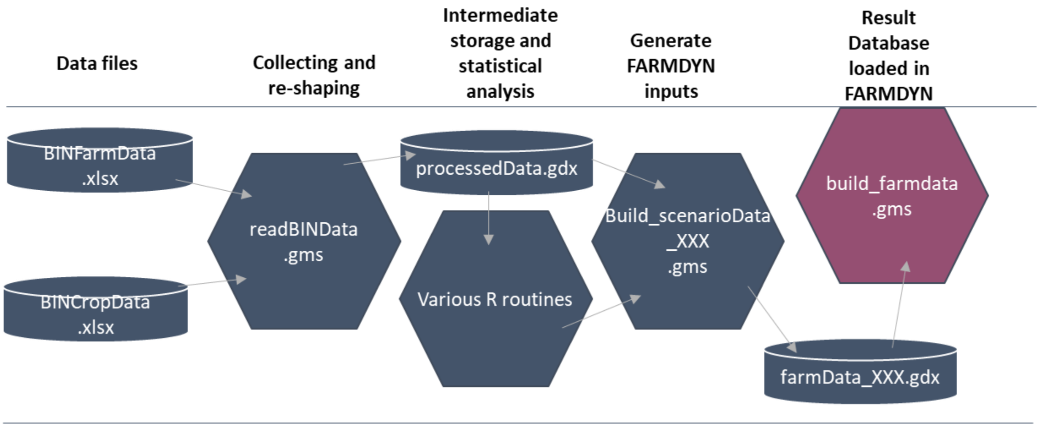


Figure 6 BIN and LMM Data processing

A typical head of the resulting GAMS parameter p_farmData is depicted below. The global variables derived from BIN contain numerical values that are passed to GAMS parameters as well as structural information ("farmBranchDairy"), which is encoded as negative numerical values (here - 1 for "TRUE"). The FarmDyn file "build_farmdata.gms" (highlighted in purple in Figure 6) assign values to FarmDyn parameters, builds the needed set-elements and global variables that govern the model execution.

Based on this workflow, it is possible to customize the complete set-up of FarmDyn based on farm-specific data from BIN.

A crucial set of data for FarmDyn are the machinery requirements for field operations and the associated labour requirements and costs. This information is not available in BIN, at least not in the needed detail. Data on field operations was therefore collected from the KWIN database and stored in gdx-format and the file cropop_nl was created for inclusion in FarmDyn. The workflow is much simpler than in the case of BIN data because field operation data are assumed to be equal for all farms, so no aggregation or sub-setting is required.

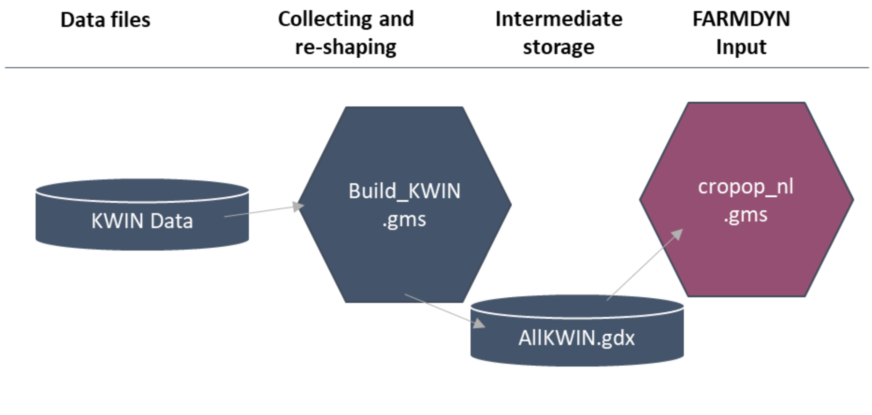


Figure 7 KWIN Data Processing

Typical data-parameters included in AllKWIN.gdx are for instance machinery attributes in the GAMS parameter "op_attr":

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

Head of op_attr

	Diesel	varCost	labTime	fixCost	nPers
Herb	1	3	0	4	1
Combine	21		2		2
sowMachine	5	11	1	9	1
rotaryHarrow	9	22	1	8	1
singleSeeder	4	18	1	28	1
directSowMachine	7	23	1	23	1
seedBedCombi	6	12	1	8	1
springTineHarrow	7	14	1	7	1
mulcher	8	21	1	15	1
chopper		410			

The described steps to combine BIN and KWIN datasets in Figure 6 and Figure 7 into a typical database expected by the core model as illustrated in *Source: Based on Britz et al. (2021)*

Figure 2 can be summarized as shown in Figure 8:

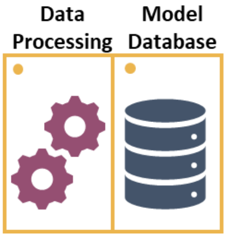


Figure 8 Example for a contributed data processing module

Processing routines (in purple) have been developed to generate the default database required by the model (in blue), using a statistical data source from a country to which the core model was not yet applied.

Applying FarmDyn to a new country also requires adjusting model equations if the default settings are not in line with e.g,. national regulations on fertilizer application levels or accounting standards for nutrient flows. This is illustrated in Figure 9, which shows the work steps included in the Dutch version of the nutrients and emission accounting modules. This module comprises of several GAMS files related to the construction of the module's database and the set-up of the included equations.

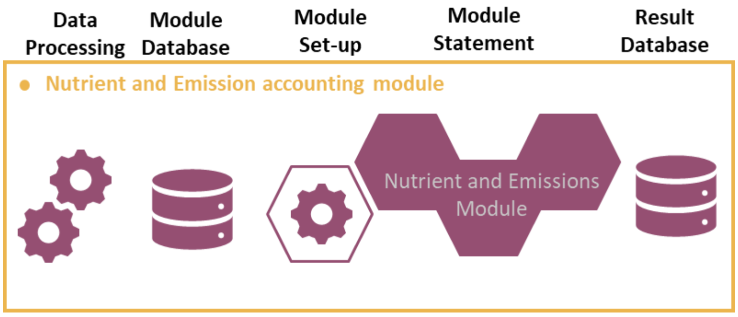


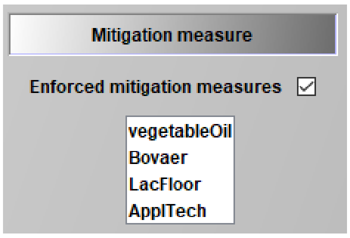
Figure 9 Elements of the Dutch nutrient intake and emission modules

3.3.2. GHG mitigation options for dairy farms

Several GHG mitigation options can be selected endogenously by FarmDyn, for instance by replacing emission-intensive feeds in the animal rations in the case of substituting soybean meal with on-farm

This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement N° 817566.

produced feeds, which can reduce upstream emissions. The more detailed representation of grassland management options as described in chapter 4 of Helming et al. (2023) falls into this group of endogenous mitigation measures. Other mitigation options have to be selected exogenously as new settings in the GUI (screenshot below). These mitigation options are discussed in chapter 2 of Helming et al. (2023). The newly introduced options are the feed additive Bovaer (R), which reduces emissions from enteric fermentation in dairy farming by up to 30% if provided in a minimum quantity. Extending the number of lactation periods per cow is a further option, which can be selected in the GUI ("LacFloor") in the screenshot below.



These measures are a contributed module in the sense of *Source: Based on Britz et al. (2021)*

Figure 2: Default data for cost and GHG reduction potential of these measures are provided (Figure 10, the purple database symbol with a "0"), a set of processing steps that convert the default data into model parameters, the changes to model equations, and dedicated reporting.

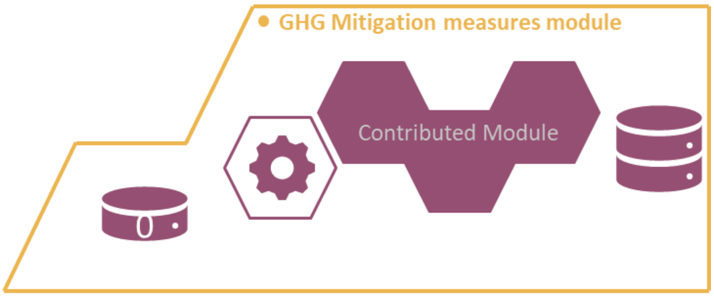


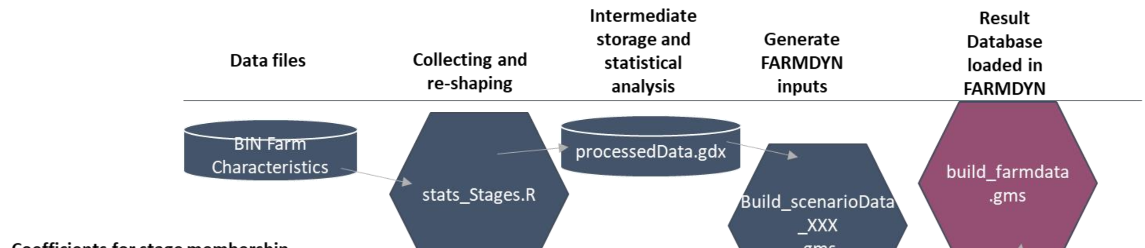
Figure 10 Contributed GHG mitigation measures module

3.3.3. Selection of GHG mitigation options

The exogenously set mitigation options outlined above and in chapter 2 of Helming et al. (2023) can be implemented as experiments by extending the farm sample data with the respective settings for each farm or farm-type, depending on farm -characteristics (e.g. permitting the extension of lactation periods only for highly intensive farms). The survey among Dutch dairy farms carried out in Task 3.3 of **MIND STEP** and described in Helming et al. (2021) (see also Wang et al., 2022) contains relevant information regarding technology adoption behaviour of farmers. It can therefore be used to differentiate farmers' behaviour within a sample. In line with relative low marginal abatement costs of the selected technologies, the survey conducted found that especially intensive dairy farms are willing to adopt GHG mitigation technologies first. This information was used to assume different adoption patterns per group of dairy farms per GHG mitigation technology, as described for instance in the scenario construction in Helming et al. (2023). As an alternative to the stratification of farms based solely on livestock density, it was also explored to which extent socio-demographic farm-characteristics like age or educational attainment, which are available from BIN, could be used to group farms according to their likely stage membership. To align with the analyses by Wang et al.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

(2022), additional data was collected from BIN, permitting to the coefficient estimates for socio-demographic factors that contribute to a certain stage membership (Figure 11):



	Dependent variable: Stage membership			
	Coefficients (SE)	P-value	95% CIs	Odds ratio
Socio-demographical factors				
Age	0.107** (0.041)	0.010	0.025, 0.188	1.113**
Age squared	-0.001** (0.0005)	0.019	-0.002, -0.000	0.999**
Basic agricultural education	0.379*** (0.025)	0.000	0.331, 0.428	1.461***
Full agricultural education	-0.183*** (0.051)	0.000	-0.283, -0.082	0.833***
Yearly family farm income	-0.004 (0.013)	0.722	-0.029, 0.020	0.996
Livestock density	0.337* (0.175)	0.054	-0.005, 0.680	1.401*

Source: Own compilation, table screenshot from Helming et al. (2023), Table 12; Wang et al. (2022)

Figure 11 Socio-demographic factors for determination of stage memberships

If more farm specific information would be available for all farms in the sample, especially cognitive and behavioural factors as used by Wang et al. (2022), they can be included as well to better determine the likely stage membership of a farm and which mitigation measures are likely to be taken up.

The described inclusion of additional farm-level characteristics to identify likely stage membership and, based on that, determine the exogenously set GHG mitigation measures for groups of farms can be seen as a contributed data-generation module within the overarching structure (Figure 12).

Figure 12 Module: Construction of model database from GHG survey results

3.4. Crop management choices

The work by Féménia et al. (2023) in **MIND STEP** Task 3.4 involves the estimation of micro-econometric multi-crop (MEMC) and cost-allocation models (MECA). Both model branches are trained and tested on detailed French data, but can be used to estimate crop-yields, the corresponding input demands, and their costs, as well as the impact on area-allocation at farm level using EU-FADN data. Both models are made available as function libraries ("packages") in the R programming language. The MEMC

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

model with “endogeneous regime switch” is included in the package “RPMulticrop” while the MECA model is available from the package “WInputAll”. Both packages will be made available either through the main distribution homepage of R packages (CRAN) or through the MIND STEP server at IIASA. Also, both models are currently applied to new regions in the EU as described in Appel et al. (2023) The following chapters describe how these MEMC and MECA models can be integrated in the overarching model structure proposed in MIND STEP.

3.4.1. Multi-crop model with endogenous regime switch (ERS-MECM)

Crop-area allocation at farm level depends on two decision: First , which yield per area (intensive margin) is economically reasonable given the cost for inputs and the output price, and second, how the available area should be allocated to the different crops (extensive margin). In typical MEMC models, farmers are assumed to allocate their cropland to the crops of a given crop set in order to maximize their expected profit or the expected utility of their profit. This ensures the economic consistency of the resulting models. However, currently available MEMC models ignore or poorly describe an important decision of crop producers: their choice to specialize on a subset of potentially producible crops, which may not have been part of the farm’s production plan before (Féménia et al., 2023). Such endogenous regime switches (ERS) are of particular importance for the FarmDyn model, because the selection of crops that can be included in the model’s cropping plan is set exogenously, currently depending on past observations at farm level: If a certain crop was not part of the cropping plan during the reference period, it will not be included unless a specific scenario setting demands it. When creating scenarios with projected prices for future periods, for instance from market-level models like GLOBIOM or MAGNET, the set of cropping options for a certain farm should not only depend on past observations. For the creation of such alternative cropping plans, results from the ERS-MEMC provide an empirically grounded tool to identify a range of production choices available to a FarmDyn instance.

The allocation of areas to a certain crop is an endogenous variable in FarmDyn, constrained by maximum and minimum rotation shares. In addition to regime switches, i.e. which crops are likely to be produced on a specific farm type, ERS-MEMC area allocation results can also provide information on plausible minimum and maximum ranges of crop rotation shares combined with information on changes in marginal costs per crop when share in crop rotation changes. In the perfect situation additional Information is available to distribute the costs changes over the different costs components in FarmDyn.

3.4.2. Cost allocation model (MECA)

Information about production costs for each crop at the farm level is very important when analyzing multi-crop farms’ behaviours. It is indeed very useful to investigate variable input uses decisions of farmers for policy purpose. Production costs per crop can also be used as explanatory variables in more complex models of production choice (Letort and Carpentier, 2010). However, information these cost per crop is generally not provided in accountancy datasets, such as the EU-FADN data available to agricultural economists. The information on variable input uses in these data only concerns aggregate expenditure at the farm level, and adequate statistical and\or economic modelling are necessary to allocate this aggregate information among the different crops produced by the farms.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

3.4.3. Integrating MECM and MECA models into the overarching structure

The empirical MECM and MECA models have been integrated in the R-packages “RPMulticrop” and “WInputAll”. Both packages are currently tested for regions in Italy, Romania, Spain and Hungary (Appel et al., 2023) within the **MIND STEP** project. These applications provide information on cost structure and cropping decisions in regions, for which no data is directly available from FADN, but which is a crucial input for the application of FarmDyn to regions in the EU not covered so far. In particular the input cost allocation model permits the parameterization of variable input cost in FarmDyn with default as well as with farm-specific estimates. Adjustments of crop yields in response to price changes and switches between cropping regimes, e.g. from root-and-tuber to cereals farming can be used to update the sample-farm parameters and settings, e.g. by including crops in the farm plan that were not observed before.

3.5. Risk management models

The aim of Task 3.5 is to improve the capacity to model CAP risk management policies by developing an IDM model which captures the acceptance and risk-reducing impacts of different risk management instruments (RMI). The empirical application focusses on weather risk in crop farms. The analysis is based on a large data set combining new survey data with existing data from detailed regional FADN and biophysical data (weather, soils, yields) according to the procedures developed in WP2. The survey explores risk preferences and attitudes to use RMI of farmers in Germany and Italy. The output of this Task is a module to analyse the acceptance and risk-reducing impacts of different RMI. The module allows to analyse the propensity to adopt RMI for a range of behavioural theories and farm and farmers characteristics (e.g. household; off-farm income; wealth; personal traits), and to analyse the impact of RMI that reduce income volatility .

3.5.1. A new risk module for FarmDyn

The empirical work on risk attitudes in Germany and Italy is based on the concept of Cumulative Prospect Theory (CPT, Tversky and Kahnemann, 1992). In this context, risk attitudes are represented by a utility function that puts higher weights on potential losses than on potential gains, which is an improvement over previous approaches that only accounted for the overall variability of farm outcomes, like expected-value-variance (EV) risk utility functions. In addition, the TK utility function can be augmented by providing subjective probability weights to address the fact that farmers may perceive certain negative outcomes as more likely than suggested by available data. The FarmDyn model was set up to incorporate certain types of risk modelling approaches, like MOTAD, but a TK utility function was not included until recently (Britz, 2022). A particular challenge was that the TK utility function is not linear and cannot directly be used to replace the current objective function without converting the model into a non-linear mixed-integer optimization problem (MINLP), which are difficult to solve, because they combine the combinatorial nature of mixed integer programs (MIP) and the challenges in solving nonlinear programs (NLP). They also require different numerical algorithms than the FarmDyn default solver CPLEX. To address these computational challenges, Britz (2022) included a linear approximation of the TK utility function.

In general, the risk behaviour of an individual model instance in FarmDyn can be parameterized within the GUI (Figure 13, with parameterization from the original paper by Tversky and Kahnemann, 1992)

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

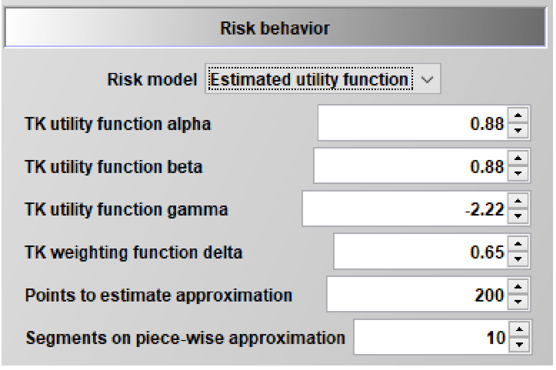


Figure 13 Default parameters for the TK risk utility function

The CPT risk module is meanwhile (Britz, 2022) included in the main distribution of FarmDyn and it is a is a good example for a contributed module in the sense of *Source: Based on Britz et al. (2021)*

Figure 2 as it provides default values, processing steps, additional model equations and dedicated reporting as shown in Figure 14:

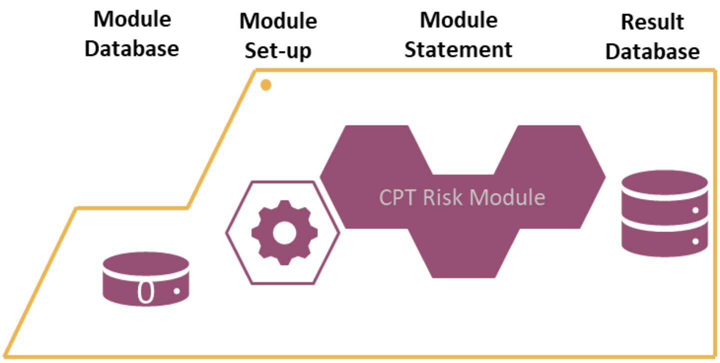
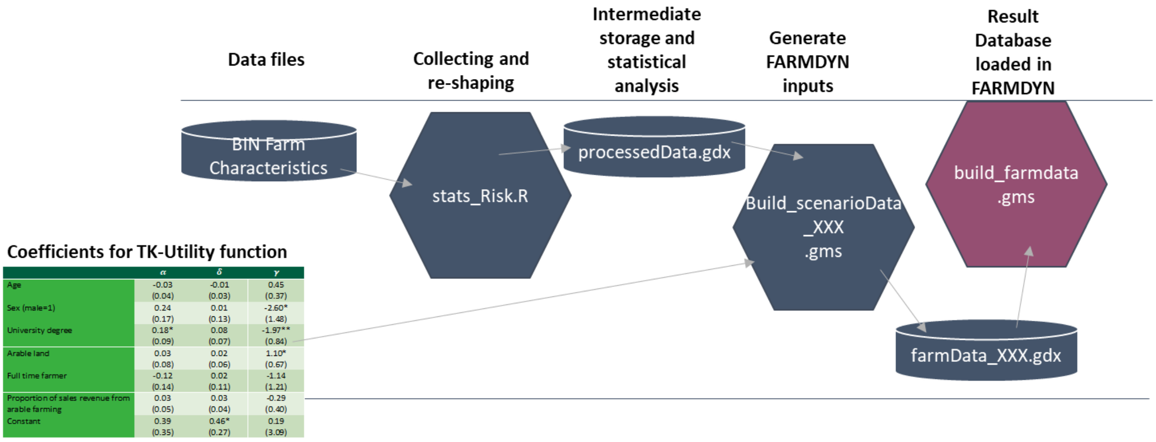


Figure 14 CPT Module in FarmDyn

3.5.2. Farm-specific risk preferences

The shape of the used utility TK function depends on a range of farm characteristics, for instance age and university degree of the farm manager as well as acreage and economic size. Duden et al. (2023) estimate the relation of such farm characteristics with the shape parameters of the TK utility function as required by FarmDyn. Similar to the integration of empirical results for technology adoption preferences as described in chapter 3.3.3, the additionally required data were extracted from the FarmDyn database and combined with the estimated coefficients by Duden et al. (2023) to generate farm-specific versions of the TK function (Figure 15).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.



Source: Own compilation, table screenshot from Duden et al. (2023)

Figure 15 Integration of estimated TK-Parameters in the FarmDyn database construction

The described workflow represents a contributed module to the overall FarmDyn data generation process as it adds new information (farm characteristics) to the database, which is then used to calculate farm-specific parameters for the TK risk-utility function. In the sense of the modular concept shown in Figure 2, the integration of these risk preferences can be depicted as in Figure 16. More details on the integration of TK-function parameters in the sample-farm construction process can be found in Müller et al. (2022) in the context of establishing a workflow based on a wrapper function.

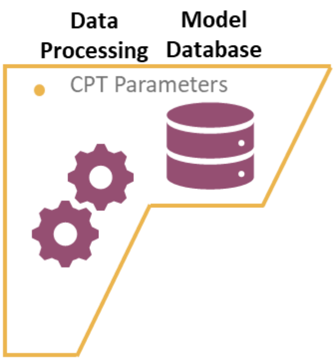


Figure 16 Parameterization of the CPT risk module

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

4. INTERACTION WITH AGRIMODEL CLUSTER MODELS

From the very beginning, representatives of the three projects BESTMAP, AGRICORE, and MIND STEP, which form the AgriModels Cluster (https://agrimodels-cluster.eu/), have met regularly to identify potential overlaps, synergies, and project results that are of common interest. The application for EU-wide FADN data for the respective project consortia was a major topic initially as it showed the challenges arising from the recently introduced General Data Protection Regulation (GDPR) and the data needs of research projects focussing on individual decision making at farm level. While the AgriModels Cluster projects found their specific solutions to this challenge, the general problem that data and results at farm level can only be obtained, exchanged, and published subject to severe privacy and confidentiality restrictions continued to be a topic of discussion because of the requirement that research data and results should be Findable, Accessible, Interoperable, and Reusable (FAIR), which has become a widely applied and required principle for data management. This topic was addressed by a presentation by AgriCore during a jointly organized session of the AgriModels Cluster at the 16th EAAE Congress in July 2021, titled “Using synthetic populations to produce representative and anonymous distributions of farm characteristics of the real farmers’ population of interest from different data sources”. The proposed concept to generate synthetic populations that replicate observed farm-level data in a stochastic sense but avoid the risk of re-identification of individual farms was recognized as useful for future projects with the objective to model individual decision making processes.

Following the positive experience with jointly organized workshops and sessions, the projects involved in the AgriModels Cluster continued in this manner, e.g. at the workshop on the linking of ABM & CGE organized by the BESTMAP consortium (12 and 13th May, 2022) and at the 181th EAAE organized pre-seminar session on the state of play regarding modelling of individual decision making. The presentation by the AGriCore modelling team on a Positive Mathematical Programming (PMP) agent-based model for ex-ante assessment of regional agri-environmental schemes was further explored and led to the participation of the farm level modelling team from AgriCore in data and model collaboration in Task 6.4 of **MIND STEP.** This PMP model provides interactions between farms to model long term dynamics on land markets. Challenges are the use of biophysical data to calculate emissions and include agronomic restrictions, the modelling of alternative technologies and EU coverage, because AgriCore is focussing its modelling activities on three different regions in Europe. To foster cooperation, circumvent double work, and enrich the **MIND STEP** modelling toolbox, the AgriCore modelling team is now working together with the **MIND STEP** modelling teams on mandatory input reduction and GHG emission reduction scenarios. The overarching farm model in **MIND STEP** could be used as coefficient generator for the AgriCore model, comparable to the combined use of FarmDyn and GLOBIOM.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 817566.

5. CONCLUSIONS

The present deliverable 3.2 provides an overview on conceptual considerations and implementations related to an overarching model structure at farm level in the **MIND STEP** project. It builds on the in-depth review of four applied models by Britz at al. (2021) and their proposed design principles for a generic, flexible, and modular bio-economic farm-level model, that can be extended and co-developed by a network of researchers. The **MIND STEP** deliverable 3.1 "Specification of model requirements: Protocols for code and data" (Müller et al., 2021) operationalized these design principles, in particular the concept that the different tasks within work package 3 should be aligned around a core simulation model, to which additional functionality and empirical grounding can be added in a flexible and consistent way. Among the models reviewed by Britz et al. (2021) were IFM-CAP and FarmDyn, which are hosted by partners in the **MIND STEP** consortium and are both potential candidates for a core model. The work carried out in Tasks 3.3, 3.4, and 3.5 and the resulting requirements for the needed additional functionalities of the core model made it clear that FarmDyn is well suited for this role. FarmDyn was subsequently enhanced to meet the requirements within the **MIND STEP** project, in particular with regard to the representation of GHG mitigation options, the integration of farmers' preferences for their adoption, the inclusion of a wider range of grassland management options, or the development of a risk module based on cumulative prospect theory (Britz, 2022), to name a few. The new model features came along with additional data requirements, e.g. for grassland management options, fertilizer restrictions, or for general farm characteristics needed to derive farmers' preferences. These new data processing steps are also integrated in the workflow around the FarmDyn model by adding interfaces to new databases and updating existing ones. Due to the immense data requirements of FarmDyn, it was usually applied in case studies for typical or representative farms. Applications at individual farm level are in principle possible and the databases are constructed to permit this, but in particular for testing and extensive simulations, it appeared more practical to execute the model for a smaller number of such typical farms. In addition, privacy and confidentiality regulations restrict the exchange of data among modelling teams, which can be circumvented by the creation of typical farms based on groups of sufficient size. The flexible construction of farm samples and the creation of meaningful farm groups became also part of the modelling workflow. Chapter 3 of this report gives an overview on the implementation of these new modules in the overarching FarmDyn structure.

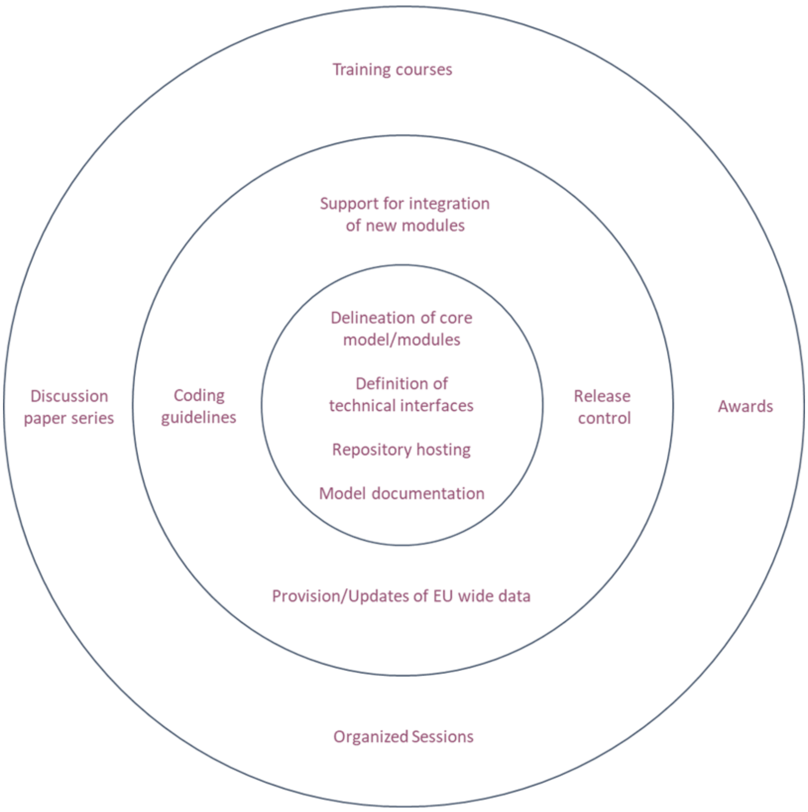
It has to be noted that the described uses of FarmDyn within **MIND STEP** do not yet fully exploit all possible model features. The new CPT risk module (Britz, 2022) has been applied by the teams at Thuenen and UCSC based on surveyed farms in Germany and Italy, but has not yet been tested for the Netherlands. Also, FarmDyn provides rich dynamic features (as suggested by the model name), which are very important for the analysis of investment decisions or for the decision to stop with farming altogether. These model features have a huge potential to provide enhanced insights regarding the impact of new policy and technology options and their use will be further explored.

The enhancements of FarmDyn led to a number of interesting applications within **MIND STEP** and improved the collaboration among the different teams of developers, but also caused a range of new challenges. Increased flexibility and modularity also means increased complexity, with substantial implications for model co-development, testing, and comparison of results between applications among the working groups. Comparison of model results, for instance, is only possible if all model settings are aligned, but this is already not feasible if, for instance, country-specific fertilizer ordinance modules have to be included. The protocols, guidelines, and measures for quality control put forward in Müller at al. (2021). can solve such problems only partly. Good coding practices, separation of code from data, version control, the usage of test cases, and so on, greatly facilitate the exchange of modules, but their application in a new context could be impaired by conflicting settings in other modules. An example is the extended grassland module, which was developed in combination with Dutch dairy farms under Dutch fertilizer restrictions. Applying this for a German case study may cause

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

conflicts with the German fertilizer restrictions unless the provided default data are adjusted. Such content-related problems are difficult to address by pre-defined protocols and can best be solved by direct communication between the involved developers. Ensuring the continued use of the newly developed modules and data processing routines calls for establishing a network of developers and users, that remains in existence beyond the duration of a single project like **MIND STEP.** On this topic, Britz et al. (2021) conclude, that such a network would be essential for ensuring longevity and usability of a modelling framework centred around a bio-economic model like FarmDyn. The composition of the **MIND STEP** consortium and the distribution of the development work has already certain characteristics of a network as proposed by Britz et al. (2021). The developer teams are hosted mainly in public research institutions, i.e., governmental research institutions, European organisations or universities. FarmDyn is developed at the University of Bonn, where the main repository is located. It was adjusted for Dutch conditions by Wageningen Economic Research (as outlined above and in Helming et al., 2023) in close cooperation with the team in Bonn. New GHG mitigation options were also first implemented by the team in Bonn, then tested and re-parameterized in an iterative process with Wageningen Economic Research. The new CPT risk module was also developed in Bonn (Britz, 2022) and then applied at Thuenen and UCSC as described in Duden et al. (2023). The involved teams form already a core network of developers and users as envisaged by Britz et al. (2021). Following their hierarchy of actions for establishing a network (*Source: Britz et al. (2021)*

Figure 17), the delineation of core model and modules, the definition of technical interfaces was conceptualized in deliverable 3.1 and subsequently implemented as described in this report. The repository for FarmDyn is hosted by the Bonn team, which also provides detailed model documentation. New modules as developed in **MIND STEP** can be shared via the repository hosted at IIASA (see **MIND STEP** deliverable 7.4.)



Source: Britz et al. (2021)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

Figure 17 Hierarchy of actions to build a network of model users and developers

Some general problems remain: research organizations and universities rely on third-part funding and usually have a high turnover of staff, such that individual knowledge may be lost for a network partner. If the network is sufficiently established, this could be compensated by taking over certain development responsibility by other partners and inclusion of new partners into the network. Training courses, organized sessions during conferences, or even discussion paper series can provide incentives for new partners to join the network and provide support. In this sense, the activities in **MIND STEP** can be seen as a first step towards a functioning model network to ensure the longevity of the overarching model structure proposed here.

The outlook is that a version of FarmDyn as the overarching farm model with the high level of detail regarding farm inputs and outputs would already be a valuable addition to the set of models used by the EU commission for policy evaluation and design. As the overarching farm model it could serve as coefficient generator for the market models and it could give insights into the heterogeneous impacts of policies and events on farm level.

6. ACKNOWLEDGEMENTS

This deliverable 3.2 is developed as part of the H2020 MIND STEP project which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.

7. REFERENCES

Articles and reports

Britz W., Lengers, B,. Kuhn, T., Schäfer, D., 2016. A highly detailed template model for dynamic optimization of farms - FarmDyn, University of Bonn, Institute for Food and Resource Economics, http://www.ilr.uni-bonn.de/em/rsrch/farmdyn/farmdyn_docu.pdf.

Britz, W., Ciaian, P., Gocht, A., Kanellopoulos, A., Kremmydas, D., Müller, M., Petsakos, A., Reidsma, P., 2021. A design for a generic and modular bio-economic farm model. Agric. Syst. 191, 103133. https://doi.org/10.1016/j.agsy.2021.103133

Britz, W., 2022. Implementing an approximation of Cumulative Prospect Theory into Mixed Linear Programming – an application to bio-economic modelling at farm-scale considering crop insurance. German Journal of Agricultural Economics, under review.

Kuhn, T., Schäfer, D., Holm-Müller, K., Britz, W., 2019a. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems, 173, 233-243.

Kuhn, T., Grauer, Z., Holm-Müller, K., 2019b. Einzelbetriebliche Modellierung der Auswirkung von novellierter Düngegesetzgebung und agrarpolitischer Entwicklung auf landwirtschaftliche Betriebe in NRW.

Lengers B., Britz, W., Holm-Müller, K., 2014. What drives marginal abatement costs of greenhouse gases on dairy farms? A meta-modelling approach, Journal of Agricultural Economics 65(3): 579–599

Letort, E., Carpentier, A., 2010. Variable Input Allocation: Why Heterogeneity Matters? [Paper presentation] 120th EAAE Seminar "External Cost of Farming Activities: Economic Evaluation, Environmental Repercussions and Regulatory Framework", Chania, Crete, Greece

Louhichi K., Ciaian, P., Espinosa, M. , Perni, A. , and Gomez y Paloma, S. , 2017. Economic impacts of CAP greening: application of an EU-wide individual farm model for CAP analysis (IFM-CAP). European Review of Agricultural Economics, https://doi.org/10.1093/erae/jbx029.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty, 5, 297-323.

Wang, S., Höhler, J., Ang, F., Oude Lansink, A. G. J. M., 2022. Dutch dairy farmers' adoption of climate mitigation measures – the role of socio-psychological and socio-demographical factors [Paper presentation]. The Agricultural Economics Society, Leuven, Belgium.

MIND STEP Deliverables :

Appel, F., Ang, F., Wang, S., Müller, M., Helming, J., Féménia, F., Koutchadé, O.F., Addo, F., Schokai, P., Rosell, M., Balazs, K., 2023. Report on the validation of the b model toolbox approach & proof of concept. **MIND STEP** Deliverable 6.2

Féménia, F., Carpentier, A., Koutchadé, P., Letort, E., Wang, S., Ang, F., Sckokai, P., Varacca, A., 2023. Report on modelling crop management practices and interfaces to the **MIND STEP** model toolbox. **MIND STEP** Deliverable 3.4

Duden, C., Offermann, F., 2023. Report on modelling risk management and interfaces to the **MIND STEP** model toolbox. **MIND STEP** Deliverable 3.5

Helming, J., Blokland, P.W., Müller, M., Wang, S., Roerink, G., Schäfer, D., Ang, F., 2023. Report on modelling greenhouse gas emission including adoption behaviour of farmers regarding mitigation strategies and interfaces to the **MIND STEP** model toolbox. **MIND STEP** Deliverable 3.3

Müller, M., Schäfer, D., Britz, W., Sckokai, P., Carpentier, A., Femenia, F., Offermann, F., Wang, S., Ang, F., Oude-Lansink, A., Pahmeyer, C., 2021. Specification of model requirements: Protocols for code and data. **MIND STEP** Deliverable 3.1

Müller, M., Gocht., A., Scherer, H., 2022. Beta version of a wrapper to allow a standardized communication channel between the models. **MIND STEP** Deliverable 7.2

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566.