Bio-economic farm modelling for integrated assessment of agricultural and environmental policies:

Towards re-usability and improved empirical validity

Argyris P. Kanellopoulos

Thesis committee

Thesis supervisor

Prof. dr. ir. A.G.J.M. Oude Lansink
Professor of Business Economics
Wageningen University

Thesis co-supervisors

Dr. ir. M.K. van Ittersum
Associate professor, Plant Production Systems
Wageningen University

Dr. ir. P.B.M. Berentsen
Assistant professor, Business Economics
Wageningen University

Other members

Prof. F. Arfini, University of Parma, Italy
Prof. dr. H.B. Meinke, Wageningen University
Dr. ir. W.A.H. Rossing, Wageningen University
Prof. dr. E.C. van Ierland, Wageningen University

This research was conducted under the auspices of the Mansholt Graduate School

Bio-economic farm modelling for integrated assessment of agricultural and environmental policies:

Towards re-usability and improved empirical validity

Argyris P. Kanellopoulos

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 27 September 2010
at 11 a.m. in the Aula.

Argyris P. Kanellopoulos

Bio-economic farm modelling for integrated assessment of agricultural and environmental policies: Towards re-usability and improved empirical validity

PhD thesis, Wageningen University, Wageningen, NL (2010).
With references, with summaries in English and Dutch

ISBN 978-90-8585-702-0

Abstract

The main objective of this PhD thesis was to develop and evaluate a generic bio-economic farm model that can be used under different biophysical and socio-economic conditions for integrated assessment of a variety of agricultural and environmental policies. The functionality of the generic bio-economic farm model developed in this thesis was illustrated with an analysis of the impacts of the 2003 reform of the Common Agricultural Policy in the European Union for arable and livestock farms in a context of market liberalisation.

In bio-economic studies, estimation of model parameters related to increasing costs because of limited machinery and managerial capacity, decreasing yields because of land heterogeneity and risk aversion is often not possible because of lack of data. Not including or misspecifying such parameters can have negative consequences on the forecasting performance of the model. In this thesis, methodologies based on Positive Mathematical Programming and Maximum Entropy estimation were proposed and implemented to recover unknown parameters underlying the actual decision making of farmers and to improve the forecasting performance of the model. The proposed methods relax a number of arbitrary assumptions of existing calibration methods and enhance representation of the actual decision making. The forecasting capacity of the models calibrated with the proposed methods was tested in ex-post experiments in which the models were calibrated with historical data of a particular base year and used to forecast policies and price changes of the following historical years. Results of these ex-post experiments showed that the proposed calibration methods improve the forecasting capacity of the model.

For meaningful assessment of future policies using bio-economic models, a comprehensive set of alternative activities must be identified. Combinatorial procedures and filtering rules have been used in the literature to generate a set of activities that can be evaluated in bio-economic models. One very important limitation of combinatorial procedures is that the number of generated activities can easily explode. However, many of these activities are inferior with respect to their input-output relationships and they will never be part of the solution of the bio-economic farm model. In this thesis, a method based on Data Envelopment Analysis was proposed to identify and select alternative agricultural activities, representative for specific policy questions that can be used in bio-economic models. The Data Envelopment Analysis method reduced the number of alternative agricultural activities generated by existing combinatorial procedures by 95%, arriving at a number that can easily be applied in bio-economic farm models. The proposed method was applied to a problem of alternative nutrient management in Flevoland (the Netherlands).

Keywords: integrated assessment; environmental policy; agricultural policy; market liberalization; bio-economic model; farming systems; mathematical programming; maximum entropy estimation; data envelopment analysis; agricultural activity; land use; future studies.

Preface

The accomplishment of this PhD thesis would not be possible without the valuable contribution of a number of people that have supported me during all these years. I feel that the least I can do is to try and acknowledge them in the next few lines. However, words are only words and thus I hope my gratitude has also been expressed through actions during all these previous years.

First of all, I would like to take the opportunity and thank my supervisors, Prof. dr. Alfons Oude Lansink, Dr. Martin van Ittersum and Dr. Paul Berentsen for their contribution in finalizing this thesis. Obviously, without them this thesis would never have come to an end. Alfons, thank you for your critical comments and your methodological guidelines on the agricultural economics side of this thesis. Your multi-disciplinary background and the broad overview of methodological advances in various fields of science brought in this PhD thesis a lot of creative criticism. Martin, you have great management skills and an excellent way of dealing with people. Thank you for the thorough review of my articles even in busy or very busy periods and thank you for making my PhD a continuous learning process. Without your support, there would be a great risk and many practical reasons that I would not be able to finalize this thesis. Paul, your door was always open for me and I never felt alone in taking important decisions. You have contributed the most in improving my scientific writing skills while your knowledge on mathematical programming techniques and your experience with modelling livestock farming systems have been vital for this thesis.

I would also like to thank all co-authors that have been contributed in different chapters of this thesis. Special thanks to Prof. Thomas Heckelei for the discussions and his valuable input on calibration methods and Positive Mathematical Programming. Many thanks to Sander Janssen for being a person that I could count on at work but also in my everyday life here in the Netherlands. I also feel that I should acknowledge my colleagues from the Business Economics group and the Plant Production Systems group for creating a friendly and pleasant working environment during all these years.

My stay in the multicultural environment of Wageningen brought me close to a number of people that offered me very beautiful moments. Christo, Chrysa, Greg, Gianni, Hans, Igor, Irina, Lan, Lina, Lucine, Maki, Maraki, Marie, Natasha, Nico, Pano, Sander, Spyro, Stefane, Tania, Tarek, Theo, Thodore, Thomas, Venetia thank you for the tequila parties, the bier afternoons, the dinners, the different points of view, the basketball games, the cycling tours and all beautiful moments. Greg, thank you for the long discussions and stimulating conflicts for science, politics and women over a bier or even during working hours. Thank you for tolerating me as a roommate all these years. Xenofonta, thank you for reminding me that finishing a PhD is definitely not only about contributing in science but also developing as a human and contributing to society. Thank you for not giving up in periods that all for me was work and thank you for not missing an opportunity reminding me with your unique way how ignorant I still am. Demostheni (ξάδελφε), they say “you choose your friends but not your family”. I am glad that you are and you will always be member of my family but I would also like to thank you for being one of the few people that I can call “a close friend”. Of course I can’t forget those people who most of the time are far away but still they have their way to be so close. Christina, Gianni, and Maria your friendship means a lot to me. Thank you for believing in me and for having a beautiful way to create a safe environment that always enhance my confidence. Giorgo, thank you so much for your friendship the laughter and the beautiful moments that will never be erased from my mind. Stavro, thank you for being a person who I know I can count on during all difficult moments of my life, thank you for the discussions and all these beautiful memories.

Those that I feel I have to thank the most are my family for supporting me in all possible different ways during all these years and experiencing with me all good and not so good moments of my life. Πατέρα και μητέρα πώς να σας ανταποδώσω τόση αγάπη και πώς να σας ευχαριστήσω για ότι έχετε κάνει για εμένα όλα αυτά τα χρόνια; Είμαι πεπεισμένος ότι δεν θα μπορέσω ποτέ να τα καταφέρω! Αισθάνομαι ότι τα λόγια εδώ είναι πολύ λίγα. Εύχομαι όμως κάποια μέρα να μπορέσω να ανταποδώσω μόνο ένα μικρό μέρος της ανιδιοτελούς σας αγάπης. Ματινάκι μου, σε ευχαριστώ για την αγάπη σου, τις ωραίες συζητήσεις και τα γέλια που έχουμε ρίξει. Θα είμαι πάντα δίπλα σου! Τζανέτο, Τασία και Αργύρη σας ευχαριστώ για την υποστήριξη, τα ψαρέματα, τις συζητήσεις, τα σχέδια, τις γιορτές και τις χαρούμενες στιγμές.

My dear Foteini (Φωτεινή μου), I really hope that those few lines of text are not necessary to show you my appreciation for being there for me all these years. You have a magic way of bringing light and color when all seems to be so dark. There is no doubt that without your love, care, and laughter I wouldn’t have managed to finalize this thesis.

Argyris Kanellopoulos (Wageningen, July 2010)

Table of contents

Chapter 1

1. General Introduction

Argyris Kanellopoulos[1,2]

[1] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*

[2] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*

1.1. Bio-economic farm models for integrated assessment

Agricultural systems in Europe are confronted with critical issues such as trade liberalization, globalization and changes in the political, social and physical environment. Adaptation to the new conditions through redesign of farming systems and adoption of alternative production techniques are required to contribute to sustainable development. Effective policy decisions are necessary at global, national, regional and even farm level to promote sustainable development and enable quick diffusion of alternative technologies. To ensure the efficiency and effectiveness of agricultural and environmental policies, it is necessary to evaluate and analyze them before their application (ex-ante assessment). The European Commission has formalized this through a mandatory ex-ante impact assessment of new agricultural and environmental policies (EC, 2005). The System for Environmental and Agricultural Modelling: Linking European Science and Society (SEAMLESS) (Van Ittersum *et al.*, 2008) was one of the projects funded by the EU to develop scientific methods to support ex-ante assessment of agricultural and environmental policies.

Successful ex-ante evaluation of agricultural and environmental policies can be achieved by integrated assessment which was defined by Rotmans *et al.* (1996) as “an interdisciplinary and participatory process combining, interpreting and communicating knowledge from diverse scientific disciplines to allow a better understanding of complex phenomena”. Integrated assessment can be facilitated by interdisciplinary and quantitative tools that are able to systematically analyze the consequences of policies to the farm household and reveal the effects of the aggregate demand and supply to the regional market conditions (Van Ittersum *et al.* 2008). Such tools for integrated assessment are bio-economic models, which are model formulations of farmer’s resource management decisions linked to biophysical models that describe production processes and the conditions of natural resources (Janssen and Van Ittersum, 2007; Bardier & Carpentier 2000; Barbier & Bergeron, 1999).

In existing bio-economic studies, the farm household is the key decision making unit (Ruben *et al.* 1998). The agro-ecological environment and the farm endowments define feasible production activities, while the socio-economic environment influences the decision making of the farm household by offering incentives and disincentives for selecting or declining the available production activities. The socio-economic environment is affected by policies regarding e.g. technology, infrastructure and environment, while aggregated demand and supply influence the regional market conditions.

1.2. Model requirements for bio-economic farm models used in integrated assessment

Integrated assessment of agricultural and environmental policies requires analysis at field, farm, regional, national, continental or even global scale and it involves scientific methods used in various disciplines. For that reason bio-economic farm models which are created for integrated assessment must meet a number of important requirements (Janssen *et al.*, 2010):

- The model must be integrated with other models operating for different scales, sectors of the industry and/or scientific disciplines in a comprehensive and integrated framework, where outputs of one model can be easily translated to inputs for other models. The integration must be streamlined in terms of methodology (e.g. temporal and spatial scales), concepts, scenarios and software.
- The conceptual design of bio-economic farm models for integrated assessment must be generic and easy to modify for assessing different policies under various socio-economic and bio-physical conditions (e.g. different farm types and different regions) to minimize development time and resources needed to re-use the integrated framework for new questions and applications.
- Production activities and available technology must be described in an explicit and transparent way to improve the explanatory power of the model.
- The data needs should preferably be limited to those data available, minimizing the resource demanding process of data collection. The models must be robust enough to function with data like those from Farm Accounting Data Network (FADN) collected at European Union (EU) level. Moreover, the model must be capable to exploit more detailed data available at regional level or data at EU level that is not currently available but might become available in the future.

1.3. Calibration and validation of bio-economic farm models

Fully specified bio-economic farm models remain an ambitious undertaking. This is mainly due to complexity and lack of data which necessitates simplifications and assumptions with respect to the human decision making (Berger, 2001), the specification of currently used or alternative production activities, the dynamic nature of many processes and (dis)aggregation issues. Depending on the purpose of the analysis a particular model specification might be more appropriate than others. Poorly specified bio-economic models result in unrealistic model simulations which do not (and should not) convince policy makers and decision makers in terms of quality of the analysis.

Given certain assumptions and simplifications the decision making of the farmer can be modelled in many different ways with different levels of detail. The results of the model will generally differ substantially between different approaches and modelling techniques. More detailed specified models are expected to produce more accurate results. However, often, detail increases complexity. The desire for accuracy and detail must be balanced against computational requirements and modelling purposes (King *et al.*, 1993). A researcher (model developer) is challenged to develop a model that is conceptually as simple as possible, is not so data intensive, is computable with existing technology and produces acceptable results for a specific purpose. The required level of detail of different bio-economic analyses must be determined in an iterative process, where model development is followed by model evaluation which results in new insights for model improvement (and so on). Despite the importance of an evaluation procedure, little attention has been paid to this issue in existing bio-economic modelling literature (Janssen and Van Ittersum, 2007). This leads to either very complicated models with enormous data requirements (not always available or of poor quality) or to very simple models that do not capture a satisfactory part of reality. In both cases for different reasons this causes lower confidence in the quality of the results.

In many existing bio-economic studies, Linear Programming (LP) models are used to simulate the behaviour of farmers and forecast future decision making. The advantage of LP models is the simplicity of the method and the limited data requirements. However, the solution of LP models suffers from overspecialization[1] while the response of LP models to policy changes is in many cases rough (i.e. “jumpy” behaviour), resulting in poor simulations. A main reason for poor results of LP models is the neglect of non-linearities

---

[1] Linear Programming bio-economic models are known to suffer from overspecialization i.e. the number of selected activities are much lower than the number of activities observed in reality. In general a large number of region and farm specific constraints are needed to ensure a more realistic solution.

(e.g. economy and diseconomy of scale, risk, land heterogeneity, multiple objectives – utility) involved in the farm production process. Ignoring the existence of non-linearities in the farmer's decision making is a common assumption which is made mainly because of lack of data (Heckelei, 2002). Lack of data is a more severe problem in cases where analysis at higher levels or scales are needed such as whole countries or the EU. In such cases, there is not enough information available to enable estimation of a non-linear model using traditional econometric approaches such as ordinary least squares, maximum likelihood, generalized method of moments (Verbeek, 2004).

A number of calibration procedures and econometric approaches have been proposed to deal with recovering non-linearities involved in agricultural production with a limited dataset. Positive Mathematical Programming (PMP) was presented by Howitt (1995) as an elegant calibration procedure that could be used to recover the unknown non-linear parameters of the model's objective function. After the first introduction of PMP in agricultural economic modelling a large number of PMP variants have been developed (Helming *et al.*, 2001; Heckelei, 2002; Röhm and Dabbert, 2003; Buysse *et al.*, 2007). The Maximum Entropy (ME) criterion could be used to exploit available information more efficiently than PMP and to recover the value of the unknown parameters using existing prior information in cases of limited available datasets. Paris and Howitt (1998) demonstrated the applicability of ME in bio-economic modelling of ill-posed problems while Oude Lansink (1999) used ME to estimate farm-specific output-supply and input-demand relationships to capture technological heterogeneity between farms. Heckelei and Wolff (2003) used ME to estimate bio-economic farm models based on the optimality conditions of a sector gross margin maximization problem.

Both existing PMP and ME based methods guarantee a good reproduction of historical data and more realistic simulations compared to LP models. The problem is that the calibration procedures will dominate the simulation process and the calibrated model will reproduce historical data adequately even in poorly specified models. In such cases, the capacity of the model to forecast future changes is limited and the quality of the analysis doubtful. Evaluation of the forecasting performance of the model seems to be absolutely necessary for assessing the quality of the model and subsequently of the whole analysis. Unfortunately, evaluation of the forecasting capacity of models is not a panacea in existing bio-economic literature (Janssen and Van Ittersum, 2007).

1.4. Alternative agricultural activities and technological innovations in bio-economic farm models

Ex-ante assessment of agricultural and environmental policies using bio-economic models is not complete without exploring alternative activities and technological innovations at farm level. The production opportunities available to a farmer today are not the same as those available in the future because of changes in the social, economic, institutional and bio-physical environment. For meaningful ex-ante assessment of future policies a set of representative activities, which is adequate to satisfy all possible targets of different objectives, is needed. Selecting a representative set of alternative activities and opportunities given a specific policy framework is a challenging procedure because it can involve multiple and conflicting objectives of the different stakeholders but also because the assessed policy regime and the available farm resources can restrict the feasible "window of opportunities" from which farmers can choose to make decisions for the future.

Procedures for the identification and quantification of alternative activities have been proposed by Hengsdijk and Van Ittersum (2003). Existing bio-economic studies have used combinatorial approaches and filtering agronomic rules to identify alternative activities in a uniform and reproducible way (Dogliotti *et al.*, 2003; Janssen, 2009). Crops, livestock, rotation requirements and management options are combined into agricultural activities that have specific input requirements. Outputs and externalities are quantified using bio-physical models and/or expert rules. The filtering rules used in this kind of tools are mainly related to crop frequency, crop sequence and management and they are used to filter out those combinations which are not feasible from an agronomic point of view. The quantified set of activities is then offered to a farm level optimization model to simulate the farmer's behaviour. This approach assures that no feasible option from an agronomic point of view, is excluded a priory and that the set of generated activities includes a wide variety of options that will or may become available to farmers in the future. One important limitation of this approach is that the number of feasible activities can increase exponentially with the number of crops, managements and bio-physical conditions (Wossink *et al.*, 1992; Dogliotti *et al.*, 2003; Janssen, 2009).

Many of the activities generated by combinatorial approaches are inferior with respect to their input-output relationships or irrelevant given a specific policy question. However, the multi-dimensional nature of the input-output relationships of such activities does not allow for straight-forward selection. Offering the full set of generated alternative activities to bio-economic farm models increases computational costs and complicates the analysis of the simulated results of the optimization process.

1.5. Research Objectives

The main objective of this PhD thesis is to develop and evaluate generic bio-economic farm models that can be used for integrated assessment of agricultural and environmental policies at multiple levels (i.e. farm, regional, national, EU). The specific objectives of this PhD thesis are:

1. To develop a generic bio-economic farm model that can be applied to assess ex-ante a wide variety of policy questions under different biophysical and socioeconomic conditions.
2. To propose and test methodology that overcomes limitations of existing calibration and estimation procedures that use limited data sets to recover unknown parameters underlying the actual decision making of farmers.
3. To propose and test methodology for identifying and selecting a set of representative alternative agricultural activities for policy assessment and future-oriented land use studies.

1.6. The SEAMLESS Integrated Framework

The Integrated Framework, System for Environmental and Agricultural Modelling; Linking European Science and Society (SEAMLESS-IF) is a framework of models that aims to ex-ante evaluate agricultural and environmental policies at multiple levels (Van Ittersum *et al.*, 2008).

To enable analysis and policy assessment at multiple levels, a number of different models were integrated. On the field level, a survey was designed to identify and quantify (in terms of inputs, outputs and externalities) current agricultural activities across the EU (Borkowski *et al.*, 2007; Zander *et al.*, 2009). Combinatorial approaches and biophysical models were used to generate and quantify alternative activities (i.e. activities that are not currently used but might become interesting in the future) (Janssen, 2009).

At the farm level, an optimization model was used to allocate agricultural activities optimally to the available agricultural area and calculate a number of socio-economic and environmental indicators for the farm types of a number of representative regions (Chapter 2).

Advanced econometric procedures were use to extrapolate farm level results to other not-simulated regions and calculate price-supply relationships for all currently existing farm types in EU (Pérez Dominquez *et al.*, 2009). A partial equilibrium model (Britz *et al.*, 2007) was used to calculate the equilibrium of price and supply of the agricultural sector and generate a set of future prices used at farm level for scenario testing.

The most important challenge of SEAMLESS was integration of all these components in one modelling framework because it involves interconnection of many disciplinary models and communication of a large number of scientists from different locations in Europe of different disciplines and cultures (Janssen, 2009).

The bio-economic farm model used in SEAMLESS-IF is presented in this PhD thesis. The farm model is used to reveal the limitations of existing calibration and estimation methods, which are currently used to recover unknown parameters in ill-posed problems. The farm model is also used to assess the proposed alternative methodologies for recovering the value of the unknown parameters underlying the actual farm’s behaviour. The survey of current agricultural activities (Borkowski *et al.*, 2007; Zander *et al.*, 2009) and the set of activities generated by combinatorial approaches and filtering rules (Janssen, 2009) of SEAMLESS-IF were used to assess the proposed methodology for identifying and selecting a representative set of alternative agricultural activities (objective 3).

1.7. Outline of the thesis

In Chapter 2, a brief overview of the SEAMLESS Integrated Framework (SEAMLESS-IF) is presented and the modelling requirements of the farm model are revealed. The main components of the proposed bio-economic farm model for integrated assessment are presented. The capacity of the model to simulate different farming systems across Europe is demonstrated in an application of arable and dairy farms of Flevoland (The Netherlands) and Midi-Pyrenees (France).

In Chapter 3, some important limitations of the standard PMP approach (Howitt, 1995) are identified and an alternative PMP variant is proposed for calibration of the farm model. An ex-post experiment for the arable farming systems of Flevoland (the Netherlands) and Midi-Pyrenees (France) is designed to compare the forecasting performance of the model calibrated with the two PMP methods.

In Chapter 4, an estimation procedure based on Maximum Entropy is proposed to exploit information available in EU level databases, recover a risk aversion coefficient and improve the forecasting performance of the bio-economic farm model. Ex-post experiments are also used to evaluate the forecasting performance of the proposed ME method.

Finally, in Chapter 5, a method for selecting superior alternative agricultural activities based on Data Envelopment Analysis (DEA) is presented. An experiment related to fertilization options for arable farming in Flevoland (the Netherlands) has been set up to demonstrate the method. Chapter 6 discusses the findings of this thesis and concludes.

1.8. References

Barbier, B. and Bergeron, G. (1999). Impact of policy intervention on land management in Honduras: results of a bio-economic model. Agricultural Systems, 60: 1-16.

Bardier, B. and Carpentier, C. (2000). The conditions for sustainability of tropical agriculture. Bioeconomic models applied to five contrasting faming systems. In mini-symposium on modelling at the international association of agricultural economists Berlin.

Berger, T. (2001). Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics, 25: 245-260.

Borkowski, N., Hecker, J.M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G., Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 1: 33-34.

Britz, W., Pérez, I., Zimmermann, A., Heckelei, T. (2007). Definition of the CAPRI Core Modelling System and Interfaces with other Components of SEAMLESS-IF. SEAMLESS Report No. 26, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, pp. 114 (www.SEAMLESS-IP.org).

Buysse, J., Fernagut, B., Harmignie, O., Henry de Frahan, B., Lauwers, L., Polomé, P., Van Huylenbroeck, G., and Van Meensel J. (2007). Farm-based modelling of the EU sugar reform: impact on Belgian sugar beet suppliers. European Review of Agricultural Economics, 34: 21 - 52.

Dogliotti, S., Rossing, W.A.H., and Van Ittersum, M.K., (2003). ROTAT, a tool for systematically generating crop rotations. European Journal of Agronomy, 19: 239-250.

EC. Impact Assessment Guidelines. European Commission, Brussels, pp 99.

Howitt, R. E. (1995). ‘Positive Mathematical Programming’, American Journal of Agricultural Economics, 77: 329-342.

Heckelei, T. (2002). Calibration and Estimation of Programming Models for Agricultural Supply Analysis. Habilitation Thesis, University of Bonn, Germany (http://www.ilr1.uni- bonn.de/agpo/ staff/heckelei/heckelei_hab.pdf).

Heckelei, T. and Wolf, H. (2003). ‘Estimation of constraint optimization models for agricultural supply analysis based on generalised maximum entropy’ European Review of Agricultural Economics, 30: 27-50.

Helming, J.F.M., Peeters, L., Veendendaal P.J.J. (2001). Assessing the consequences of environmental policy scenarios in Flemish agriculture. In: Heckelei, T., Witzke, H.P., and Henrichsmeyer, W. (Ebs): Agricultural sector Modelling and policy information systems. Proceedings of the 65th EAAE seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel, pp. 237-245.

Hengsdijk, H., van Ittersum M. K., (2003). Formalizing agro-ecological engineering for future-oriented land use studies. European Journal of Agronomy, 19: 549-562

Janssen, S. (2009). Managing the Hydra in integration: developing an integrated assessment tool for agricultural systems. PhD dissertation, Chapter 1: Introduction to integration in Integrated Assessment and Modelling and bio-economic farm models. Pg. 21-22 (http://edepot.wur.nl/10616).

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Meuter, E., Andersen, E., Belhouchette, H., Blanco, M., Borkowski, N., Heckelei, T., Hecker, M., Li, H., Oude Lansink, A., Stokstad, G., Thorne, P., Keulen, H. and Van Ittersum, M. (2010). A generic bio-economic farm model for environmental and economic assessment of Agricultural systems. Submitted paper under review.

Janssen, S. and Van Ittersum, M. (2007). Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems, 94: 622-636.

King, P. R. Lybecker, W. D. Regmi, A. and Swinton, M. S. (1993). Bioeconomic models of crop production systems: Design, developement, and use. Review of Agricultural Economics, 15: 389-401.

Oude Lansink, A. (1999). Generalized Maximum Entropy Estimation and Heterogeneous Technologies. European Review of Agricultural Economics, 26:101–115.

Paris, Q. and Howitt, R. E. (1998): An Analysis of Ill-posed Production Problems Using Maximum Entropy, American Journal of Agricultural Economics, 80: 124-138.

Pérez Domínguez, I., Bezlepkina, I., Heckelei, T., Romstad, E., Lansink, A.O., Kanellopoulos, A, 2009. Capturing market impacts of farm level policies: a statistical extrapolation approach using biophysical characteristics and farm resources. Environmental Science & Policy, 12: 588-600.

Röhm, O. and Dabbert, S. (2003). Integrating agri-environmental programs into regional production models: an extension of Positive Mathematical Programming. American Journal of Agricultural Economics, 85: 254-265.

Rotmans, J. and van Asselt, M.B.A. Integrated assessment: growing child on its way to maturity (1996). An editorial essay, Climatic Change, 34:327–336.

Ruben, R. Moll, H. & Kuyvenhoven, A. (1998). Integrating agricultural research and policy analysis: analytical framework and policy applications for bio-economic modeling. Agricultural Systems, 58: 331-349.

Van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Olsson, J. A., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A. E., van der Wal, T., Wien, J. E., and Wolf, J. (2008). 'Integrated Assessment of agricultural systems: a component-based framework for the European Union (SEAMLESS)', Agricultural Systems, 96: 150-165.

Verbeek, M. (2004). A guide to modern econometrics, 2nd edition. Wiley, pp. 429.

Wossink, G. A. A., de Koeijer, T. J. and Renkema, J. A. (1992). Environmental-economic policy assessment: A farm economic approach. Agricultural Systems, 39: 421-438.

Zander, P., Borkowski, N., Hecker, J. M., Uthes, S., Stokstad, G., Rørstad P., and Bellocchi, G. (2009). Conceptual Approach to Identify and Assess Current Activities. P.D 3.3.9. SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2 (www. SEAMLESS-IP.org).

Chapter 2

2. FSSIM, a Bio-Economic Farm Model for Simulating the Response of EU Farming Systems to Agricultural and Environmental Policies

Kamel Louhichi[1,*], Argyris Kanellopoulos[2,3,*], Sander Janssen[2,3], Guillermo Flichman[4], Maria Blanco[5], Huib Hengsdijk[6], Thomas Heckelei[7], Paul Berentsen[2], Alfons Oude Lansink[2] and Martin Van Ittersum[3]

[1] *INRA- UMR Economie Publique, Avenue Lucien Brétignières, 78850 Thiverval-Grignon, France*
[2] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*
[3] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*
[4] *IAMM-CIHEAM, 3191 Route de Mende, 34093 Cedex 5, Montpellier, France*
[5] *European Commission Joint Research Centre (JRC), IPTS, Edificio Expo, Inca Garcilaso 3, 41092 Sevilla, Spain.*
[6] *Plant Research International, Wageningen University, P.O. Box 616, 6700 AK Wageningen, The Netherlands*
[7] *University of Bonn, Economic and Agricultural policy, Nussallee 21, 53115 Bonn, Germany.*

Agricultural Systems, accepted for publication.

* These authors contributed equally to this article

Abstract

The disciplinary nature of most existing farm models as well as the issue specific orientation of most of the studies in agricultural systems research are main reasons for the limited use and re-use of bio-economic modelling for the ex-ante integrated assessment of policy decisions. The objective of this article is to present a bio-economic farm model that is generic and re-usable for different biophysical and socio-economic contexts, facilitating the linking of micro and macro analysis or to provide detailed analysis of farming systems in a specific region. Model use is illustrated in this paper with an analysis of the impacts of the CAP reform of 2003 for arable and livestock farms in a context of market liberalization. Results from the application of the model to representative farms in Flevoland (the Netherlands) and Midi-Pyrenees (France) shows that CAP reform 2003 under market liberalization will cause substantial substitution of root crops and durum wheat by vegetables and oilseed crops. Much of the set-aside area will be put into production intensifying the existing farming systems. Abolishment of the milk quota system will cause an increase of the average herd size. The average total gross margin of farm types in Flevoland decreases while the average total gross margin of farms in Midi-Pyrenees increases. The results show that the model can simulate arable and livestock farm types of two regions different from a bio-physical and socio-economic point of view and it can deal with a variety of policy instruments. The examples show that the model can be (re)-used as a basis for future research and as a comprehensive tool for future policy analysis.

Keywords: bio-economic model; integrated assessment; environmental policy; market liberalization.

2.1. Introduction

Governments and policy agencies attempt to assess consequences of new policies before their introduction. The European Commission has formalized this through a mandatory ex-ante impact assessment of its new agricultural and environmental policies (EC, 2005). Science can contribute to these governmental demands for impact assessment by developing tools that can, in a transparent, rigorous and repeatable fashion, make impact assessments of agricultural and environmental policies better informed. Bio-economic farm models have been proposed for such ex-ante assessments (Flichman and Jacquet, 2003; Janssen and van Ittersum, 2007) and many recent applications (Donaldson *et al.*, 1995; Flichman, 1996; Judez *et al.*, 2001;Berentsen, 2003; Veysset *et al.*, 2005; Onate *et al.*, 2006; Riesgo and Gomez-Limon, 2006; Semaan *et al.*, 2007) assess the impacts of policy changes on economic, environmental and social indicators of agricultural systems. If a bio-economic farm model is to be used as a basis for such ex-ante assessments of agricultural and environmental policies at European level, some requirements must be fulfilled, i.e. it must be possible to upscale the model’s results (e.g. product supply) to higher system levels (e.g. country or market); data with respect to farm types, their locations and production activities must be readily available throughout various regions; the model must be applicable to different farm types including mixed farm types; the application and calibration of the model should not require many ad hoc steps or unjustified strict calibration constraints, and finally it must be possible to assess many different policy instruments. In short, it must be possible to use and apply the same bio-economic farm model in a consistent way across the European Union (EU).

A literature review showed that a generic model meeting the above requirements does not exist (Janssen and Van Ittersum, 2007). Some models focus on simulating specific farm types without providing much opportunities to expand their application beyond the original target domain (e.g. Donaldson *et al.*, 1995; Veysset *et al.*, 2005), while other models require extensive data collection limiting a rapid operationalization (e.g. Riesgo and Gomez-Limon, 2006). Various model applications address very specific EU policy issues and do not allow the assessment of a range of interrelated policy questions that EU decision-makers face (Topp and Mitchell, 2003; Onate *et al.*, 2006).

Each of these models (Donaldson *et al.*, 1995; Topp and Mitchell, 2003; Veysset *et al.*, 2005; Riesgo and Gomez-Limon, 2006) has strengths that made them suitable to be used for specific data-sets and applications. In trying to extend their use to other policies questions and locations, this specificity causes problems. With the limitations of existing approaches in mind, this article has the following two objectives. The first objective is to present the Farm System SIMulator (FSSIM) which aims to be a generic bio-economic farm model that can be applied in combination with higher level models to assess, ex-ante, a variety of policy questions under different bio-physical and socio-economic conditions. The second objective is to demonstrate the applicability of the model as a stand alone tool to assess farm level impact of future policy scenarios for different farm types in different regions. FSSIM has been developed as part of the integrated modelling framework of the System for Environmental and Agricultural Modelling: Linking European Science and Society (SEAMLESS) (Van Ittersum *et al.*, 2008) which targets to integrated assessment of agricultural systems in the EU of 27 member states (EU27). This implies that FSSIM can be and has been linked to other models for multi-scale analyses (Pérez Domínguez *et al.*, 2009).

In Section 2, the SEAMLESS context and the requirements for a model like FSSIM are presented to justify the modelling choices. In Section 3, FSSIM is described. In Section 4, the model is used to simulate arable and dairy farms of Flevoland (The Netherlands) and Midi-Pyrenees (France). In Section 5, the results of the application of FSSIM are described. Section 6 discusses the results and concludes.

2.2. Model requirements following from the SEAMLESS Integrated Framework

The main objective of the SEAMLESS Integrated Framework (SEAMLESS-IF) is to enable ex-ante evaluation of a broad range of agricultural and environmental policies at multiple decision making levels. This framework consists of models which operate in an iterative way (Figure 1). First, the Common Agricultural Policy Regionalized Impact modelling system (CAPRI) which is an EU agricultural sector model (Britz *et al.*, 2007) is used to estimate a set of initial prices for the agricultural products of all EU27 regions. Second, FSSIM uses the estimated prices and calculates supply responses of farms to price shocks in a selection of EU 27 regions. Third, EXPAMOD (Pérez Domínguez *et al*, 2009) is used to extrapolate results of the sample regions to all EU27 regions by means of econometric approaches. Next, CAPRI is recalibrated with the new supply responses coming from EXPAMOD to generate a set of market clearing prices that are used by FSSIM for the final run.

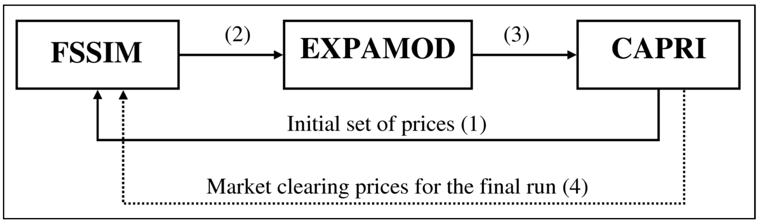

Figure 1: SEAMLESS model chain (Pérez Domínguez *et al*, 2009). See text for explanation.

Modelling all individual farms within EU27 is not possible because of the large number of farms and the existing variation and diversification among farming systems. Therefore, a farm typology was developed associating economic and environmental characteristics of EU farms. This farm typology is based on the existing EU farm typology (EEC, 1985) which classifies farms according to their income and specialization. This farm typology has been enriched with environmental criteria related to the land use and intensity of farming (Andersen *et al.*, 2007).

A spatial allocation procedure was developed to geo-reference farm types allowing the aggregation of model results at farm type level to both natural (territorial) and administrative regional level (Elbersen *et al.*, 2006; Hazeu *et al.*, 2010). FSSIM is used to simulate an "average farm" which is a virtual (not observed in reality) farm derived by averaging data from the Farm Accountancy Data Network (FADN) of farms that are grouped in the same farm type. It is assumed that the "average farm" represents all farms that belong to the same farm type. Structural changes in the sector are related to interactions between farms (land market) and alternative income sources and can only be taken into account by using another model (Zimmermann *et al.*, 2009) of SEAMLESS-IF. However, policy makers can compare the gross margin of an average farm calculated by FSSIM with the estimated income from non-farming activities to draw conclusions on the viability of the particular average farm.

The general context of SEAMLESS and the variety of policy questions that FSSIM should be able to address leads to a number of model requirements. First, FSSIM must be integrated with the other models of SEAMLESS-IF. The integration with components at field and market level must be streamlined in terms of methodology (e.g. temporal and spatial scales), concepts and scenarios being used and software. Second, the conceptual design of FSSIM should be "generic" so that the model can be easily modified and used for assessing different policies under various socio-economic and bio-physical conditions (e.g. different farm types and different regions). Third, production activities and available technology must be described in an explicit and transparent way to improve the explanatory power of the model. Fourth, the data needs of FSSIM should be preferably limited to those data available at EU27 level minimizing the resource demanding process of data collection. The model must be robust enough to function with data like those from FADN. Moreover, the model should be capable to exploit more detailed data that is not currently available but might become available in the future. Finally, FSSIM should be easily adaptable and reusable (modularity). This will allow model users to easily change it to account for different regions, farming systems, and policies.

2.3. Model description

2.3.1. Model specification

FSSIM is an optimization model which maximizes a farm's total gross margin subject to a set of resource and policy constraints. Total gross margin is defined as total revenues including sales from agricultural products and compensatory payments (subsidies) minus total variable costs from crop and animal production. Total variable costs include costs of fertilizers, costs of irrigation water, costs of crop protection, costs of seeds and plant material, costs of animal feed and costs of hired labour. A quadratic objective function is used to account for increasing variable costs per unit of production because of inadequate machinery and management capacity and decreasing yields due to land heterogeneity (Howitt, 1995). The general mathematical formulation of FSSIM is presented below:

$$\text{maximise } Z = w'\,x - x'\,Q\,x \text{ subject to } A\,x \leq b,\ x \geq 0 \qquad (1)$$

where Z is the total gross margin, w is the $n \times 1$ vector of the parameters of the linear part of the activities' gross margin, Q is the $n \times n$ matrix of the parameters of the quadratic part of the activities' gross margin, x is a $n \times 1$ vector of the simulated levels of the agricultural activities, A is a $m \times n$ matrix of technical coefficients, and b is a $m \times 1$ vector of available resources and upper bounds to the policy constraints.

A different model formulation has already been implemented and can be used if detailed agro-management information is available or if it is important to account for the risk averse attitude of the farmer explicitly. In this model formulation the farmer's utility is maximized. Utility is defined as gross margin minus risk. For this specification a linear gross margin function is assumed.

$$\text{maximise } U = w' x - \varphi \cdot \sigma \text{ subject to } A x \leq b, x \geq 0 \qquad (2)$$

where φ is the risk aversion parameter that assumes constant absolute risk aversion (Hazell and Norton, 1986), and σ is the standard deviation of the total gross margin.

FSSIM consists of four major components, i.e. arable production, livestock production, policies and regulations and the calibration and forecasting component which are described below.

Arable production

In FSSIM, arable agricultural activities are defined as crop rotations grown under specific soil and climate conditions and under well-defined management describing major field operations in detail. It is assumed that in each year, all crops of a rotation are grown on equal shares of the land. A model solution can include several crop rotations. The concept of crop rotations allows to account for temporal interactions between crops. The agricultural management of arable activities describes operations associated with fertilization, soil preparation, sowing, harvesting, irrigation and pest management of crops and results in different inputs and outputs.

FSSIM uses information available in FADN. This data source lacks detail in agro-management information which is needed to assess the environmental aspects of production. Therefore, a simple survey was performed within SEAMLESS to identify and quantify current production activities (Borkowski *et al.*, 2007; Zander *et al.*, 2010). For operational purposes and due to resource limitations the survey was conducted for a sample of 16 NUTS2 regions from the EU27 (NUTS: Nomenclature of Territorial Units for Statistics). Experts from the sampled regions were asked to specify the most important rotations and related management which are currently used by arable farms in their region. In total 87 rotations of 21 different crops were identified in the sampled regions.

The agricultural management component of FSSIM (FSSIM-AM) and the Agricultural Production Externalities Simulator (APES) (Janssen *et al.*, 2009b) can be used to quantify externalities of current activities (e.g. N-leaching) and complete sets of discrete input and outputs coefficients (e.g. costs, labour requirements, input of agrochemicals, yields, externalities) for alternative activities which have improved performance in one or more criteria. Alternative arable activities may include new crops and rotations, changes in crop management or their combination resulting in activities with different technical coefficients. Alternative activities are used to account for technological innovations in agriculture (e.g. new varieties, modern agricultural practices) and effects of future changes to bio-physical and climatic conditions (e.g. effects of climate change or soil degradation to production).

Arable farmers face a number of resource scarcities that affect their decision making. These resource scarcities have been taken into account in FSSIM by means of constraints. The available arable land constraint is specified per soil type and ensures that the sum of the area of the activities on a certain soil does not exceed the available farm land for this soil type. The available land is derived from FADN and hence imposed exogenously. Selling or buying of land is not considered in FSSIM. However, pre-determined scenarios with more (in case of buying) or less (in case of selling) available land can be tested. The available irrigated land constraint ensures that the area with irrigated activities does not exceed the available irrigable land. The available amount of irrigation water constraint ensures that the total volume of water required for the irrigated activities does not exceed the available water volume. Finally, the labour constraint is used to calculate the number of hours of hired labour, given the labour requirements of different activities and the availability of family labour. Hired labour is considered as an additional cost, the price of which is equal to the average region-specific wage rate. Allocation of family labour to off-farm activities is not considered in FSSIM. Scenarios can be used to assess consequences of allocating family labour to off farm activities by changing the availability of family labour for agricultural activities.

Livestock production

Three different animal activities are modelled in FSSIM, i.e. dairy, beef, and small ruminants (sheep and goats). The core element of a dairy activity is a productive cow, a bull and their off-springs. A replacement rate is based on the actual milk production per cow and sets the share of young animals in a dairy activity i.e. calves and heifers. For example, a typical dairy activity in Flevoland may consist of 60.5% cows, 17.5% heifers, 20.8% calves and 1.2% bulls. Increasing the activity level by 1 unit will cause an increase in the number of all animals so that the share of animals in the activity remains constant. Feed requirements of different animal types and decisions on the length of the grazing period are also taken into account in a dairy activity. The feed requirements of the herd in terms of fibre, energy and protein are covered by roughage produced on farm (fresh, hay or silage), purchased roughage (hay or silage), concentrates produced on-farm or purchased concentrates. Feed crops like grass and fodder maize are grown either in a rotation with other crops or as mono-crop activities. The quantities of on-farm produced and purchased feed depend mainly on prices of crop product (including feed) and input prices. Beef activities are modelled in a similar way. Two distinct methods of raising animals for beef production are available i.e. a suckler system comprising a cow and its off-springs, and a fattening system, which merely fattens purchased young animals till the moment of selling. The small ruminant activities for meat and milk production are modelled in a way similar to dairy and beef activities. The milk and meat production is used to determine an appropriate replacement rate and the feed requirements of different animals (Thorne *et al.*, 2009).

FADN data are used to identify the predominant livestock activities across the regions of EU, and to derive related animal shares, production levels and replacement rates. The SEAMLESS survey (Borkowski *et. al*., 2007) and a feed evaluation and animal nutrition system proposed by Jarrige (1989) were used to quantify the technical coefficients of animal activities like yields, total production costs, costs of feed, feed nutrient values and feed requirements (Thorne *et al.*, 2009).

A number of constraints were used to model the on-farm availability of resources, the feed production and the animal’s diet. Constraints relating feed availability to feed requirements are used to secure that the total requirements of energy, protein and fibre are met by the produced (on-farm) and purchased quantities of feed and concentrates. Another constraint (maximum amount of concentrates) is used to set an upper bound to the share of concentrates in the animal’s diet to prevent animal diseases related to high amounts of concentrate. The available amount of roughage constraint restricts the grazing period to a region specific maximum. Finally, the milk quota constraint restricts the produced quantity of milk to the available milk quota. Any milk production exceeding the milk quota is penalized. This constraint is the main limiting factor for a dairy farm and for that reason it is mentioned here as a resource constraint.

Policies and regulations

FSSIM is able to simulate a broad range of agricultural and environmental policy instruments, some of which have been already implemented in practice while others might be of interest to policy makers in the future. These policies are modelled as additional constraints and variables in a generic way to account easily for various products or region-specific policy implementation. The policy instruments which are currently modelled in FSSIM can be classified in a number of groups.

The first group of policies modelled in FSSIM includes the EU compensation payments which are taken into account as part of revenues in the objective function of the model. Existing compensation payments related to rain-fed and irrigated land, historical yield but also the degree of the payments that is linked to production (coupling) are taken into account in order to calculate the total amount of received payments according to the existing regime. Two farm support policies are already programmed in FSSIM, the farm support policy under the Agenda 2000 (CEC, 1999a,b) and the reform of the common agricultural policy of 2003 (CAP reform of 2003) (CEC, 2003; OECD, 2004). The first CAP reform of 1992 (CEC, 1991) and the market liberalization led to a reduction of product prices. Therefore, a regime of direct payments was developed to compensate farm income within the general context of the Agenda 2000. These direct payments were given to the highly affected arable and livestock sectors of the EU and they were linked either to production or to the area of different crops. The direct payments are financed by the EU and administered by the ministry or department of agriculture of each member state. Modelling the regional specific implementation of the Agenda 2000 requires two pieces of information: the way the payment was given (i.e. per activity level, per unit of main output) and the amount of the payment (basic premium) per hectare, slaughtered animal or tonne of product. The CAP reform of 2003 replaces the Agenda 2000 regime and involves mainly the partial (or total for some crops) decoupling of subsidies from production. To calculate subsidies under the CAP reform of 2003 in FSSIM, the subsidies received under the Agenda 2000 were (partially or totally) detached from production. To achieve this, the new coupling degree of each product was used. The decoupled part of the payment is based on the historical reference land and the total amount of subsidies received over the years 2000-2002. The coupled and the decoupled payment of each activity were used to calculate the total received subsidies per hectare of activity under the CAP reform of 2003.

The second group of policy instruments that has been modelled in FSSIM relates to quota based policies which are currently used in many EU countries to regulate the price and supply of certain products like milk and sugar beet. This kind of regulation was also used under Agenda 2000. In FSSIM quota based policies are taken into account with additional constraints. The part of production that exceeds the pre-determined quotas gets a lower price according to the specificities of the regulation. The same structure of the constraint set is used for all products that are currently under a quota regulation (or might be in the future).

Another policy that has been included in FSSIM is the obligatory set-aside policy which was introduced by the EU in 1988 (i) to reduce the large and costly cereal surpluses produced under the guaranteed price system of the CAP reform of 1992 and (ii) to provide environmental benefits following considerable damage to agro-ecosystems and nature as a result of the intensification of agriculture. Although the implementation of the set-aside policy differs across the EU, in general, the measure entails the obligation to leave a proportion of the farm land uncultivated or assigned to non-food purposes for a certain period in exchange for subsidy payments. The obligatory set-aside policy is taken into account in FSSIM by setting a lower bound to the area which is left as set-aside and by adding an extra source of revenues in the objective function for each hectare of set-aside. If the area of set-aside is less than 10% of the area of Cereals, Oil seed and Protein (COP) crops a subsidy cut is assumed.

The last group of policies modelled in FSSIM is related to the environmental conditions and cross-compliance regulations which aim at sustaining various agro-environmental conditions that must be respected to avoid reduced farm support payment under the CAP reform of 2003. Cross compliance regulations must be in line with a number of well-defined standards determined at EU level and cover environmental, food safety, crop protection, animal health and animal welfare issues. Cross-compliance regulations are taken into account mainly by additional constraints while in some cases binary variables are needed transforming the model into a Mixed Integer Non-Linear Programming (MINLP) model.

In addition to the above described policy instruments, a number of environmental indicators (e.g. total nitrogen use, water use, pesticide use), indicators related to biodiversity and multi-functionality (e.g. number of crop species on the farm), and socio-economic indicators (e.g. labour use per hour) are assessed. Those indicators can be easily used to evaluate future environmental policies.

2.3.2. Calibration and forecasting

A Positive Mathematical Programming (PMP) based approach is used to calibrate the model and guarantee exact reproduction of the observed (base year) situation without using additional calibration constraints which are difficult to justify in a way consistent with existing economic theory (Heckelei, 2003). PMP is a generic and fully automated procedure which means that it can be easily adapted and used for different regions and farm types without additional site specific information.

In PMP calibrated models, the observed activity levels of farm types are used to calculate unobserved non-linear costs which are omitted from the linear cost function of LP models because of data limitations and simplification purposes. Non-linear costs are related to issues like managerial capacity, fixed costs (e.g. machinery, buildings) and risk. PMP uses a two step approach. In the first step, a number of calibration constraints are added to the model, to ensure that the observed activity levels of the base year are reproduced. In the second step, the calibration constraints are taken out and their shadow prices are used to specify and include the non-linear costs in the objective function. Since the first introduction of PMP to bio-economic modelling by Howitt (1995) a number of PMP variants have been developed based on different assumptions resulting in different model forecasts (Heckelei and Wolff, 2003; Röhm and Dabbert, 2003; Kanellopoulos *et al.*, 2010). The appropriateness of PMP variants is case specific and depends on the available data and policy question. In FSSIM a number of PMP variants are programmed providing users with various options.

A different calibration procedure is used for the model presented in (2) where the risk aversion coefficient is the only unknown parameter. The risk aversion parameter is estimated in an iterative process that involves multiple model runs. In each model run a different value of the risk aversion coefficient is used; the value of the risk aversion coefficient that gives the best fit in terms of crop allocation is selected for simulations. In this case, exact calibration is not guaranteed.

After the model has been calibrated it can be used for forecasting. Inflation of input and output prices is considered, while exogenous to the model information on yield and price trends are used to account for possible technological innovations and price-supply fluctuations.

To facilitate the analysis of policy scenarios, FSSIM is setup in such a way that policy makers and model users can easily access and adapt the constraint set and the parameters of the model. New policy scenarios can be incorporated into the model by: (i) varying the available farm resources, (ii) changing the input and output coefficients for activities, (iii) abolishing base year policies and (iv) including new policies, constraints and parameters. A set of general policies has been pre-programmed and is ready to use after having provided the required data.

Figure 2 shows a simple presentation of the model set-up for a simulation of an arable farm in year 2003, where gross margin is maximized (risk aversion is not taken into account) subject to a number of resource and policy constraints. This presentation reveals the general structure of the model and summarizes the required information that is stored in an integrated database developed within SEAMLESS (Janssen *et al.*, 2009a). Switching on and off different components of the model allows different simulation of the same or a different farm type (e.g. the livestock component is switched on in the case where a livestock farm type is simulated).

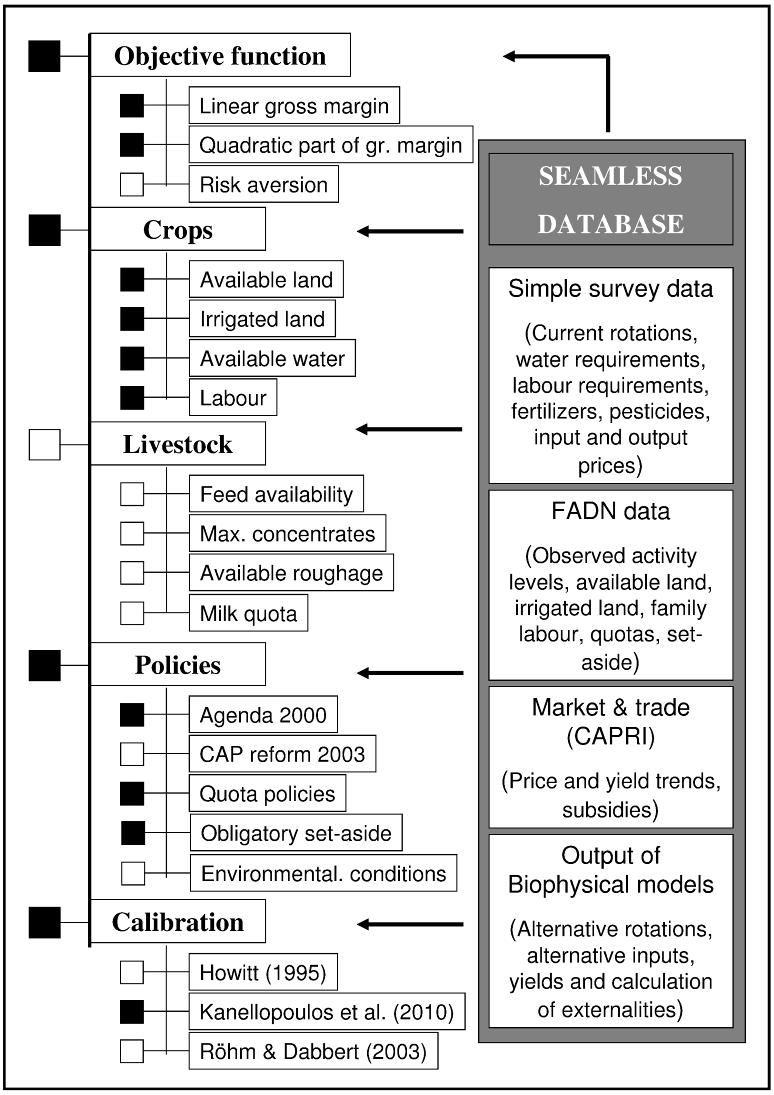

Figure 2: Set-up of FSSIM for simulating an arable farm type under Agenda 2000. Switching on (■) and off (□) components and constraints results in different simulations of a different farm type. Summary of the required information stored in the SEAMLESS database.

2.4. Set up of the calculations

Here, we present the application of FSSIM to arable and dairy farm types in Flevoland (The Netherlands) and Midi-Pyrenees (France). The bio-physical (climate and soil types) and socio-economic (different size, intensity and specialization of arable and livestock farms) conditions of these two regions differ substantially. We have chosen these regions to illustrate the applicability of the model under different bio-physical and socio-economic conditions and demonstrate the generic features of the model enabling the simulation of other farm types within the EU. For this exercise, we did not use a bio-physical model to estimate externalities because of data limitations and simplification purposes. Instead, we used total nitrogen input and total irrigation water input as environmental indicators. The model specification with a non-linear total gross margin function, described in (1), was selected for this exercise because exact calibration is guaranteed. This model specification is used for higher level analysis where data is limited and calibration only on the risk aversion parameter is not adequate to reproduce what is observed in reality. FSSIM was calibrated for the base year (2003) with the PMP variant proposed by Kanellopoulos *et al.* (2010), using activity specific supply elasticities from the literature (Jansson, 2007). For this exercise we used exogenous base year prices and consequently we did not use the full procedure described in Figure 1.

We use the four digits codes of the SEAMLESS farm typology to distinguish between the different farm types. The first digit of the farm type code refers to the farm size: (3) Large farms, i.e. size > 40 European Size Units (ESU), (2) Medium farms (16 ESU ≤ size ≤ 40 ESU), (1) small farms (size < 16 ESU). The second digit refers to farm intensity: (3) High intensity (output > 3000 €/ha), (2) Medium intensity (500 €/ha ≤ output ≤ 3000 €/ha), (1) Low intensity (output < 500 €/ha). The two last digits refer to farm specialization: (08) dairy cattle/others, (07) dairy cattle/land independent, (06) dairy cattle/temporary grass, (05) dairy cattle/permanent grassland, (04) arable/other, (03) arable specialized crops, (02) arable/fallow, and (01) arable/cereal. The set of constraints, used for the base year (2003) to simulate arable farm types consists of the resource constraints (available land, available irrigated land and labour) and policies (sugar beet quota regime and the obligatory set-aside). For dairy farms the constraints relate to the feed availability, the maximum amount of concentrates in animals' diet, and the grazing period were added. The data requirements for the base year simulations include the available farm resources (i.e. available farm land characterized by soil and climatic conditions, available irrigated land and available family labour) the inputs and outputs of current activities, the observed cropping patterns, the herd composition (Table 1 and Table 2), the economic data (i.e. variable costs of inputs, output prices and wages) and the policy data (i.e. compensation payments under Agenda 2000, quotas for sugar beet and milk production).

Table 1: Farm specific data of farm types in Flevoland in 2003, and, observed crop areas and animal numbers that are included in the current activities. Source: FADN.

		Arable farms					Dairy farms				
		FT 2303	FT 3203	FT 3303	FT 3304	Aver. farm	FT 3205	FT 3305	FT 3307	FT 3308	Aver. farm
Farm specific data											
Total available land	(ha)	17.9	66.3	68.7	33.9	56.4	49.7	44.6	33.1	48.9	44.6
Irrigated land	(ha)										
Family labour	(hrs)	3156	2997	5403	7641	4754	3325	4293	4440	3933	4196
Milk quota	(tons)						437	555	488	571	543
Costs of hired labour	(€/hr)	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5
Farms represented	(%)	13	29	44	15		8	78	6	8	
Crop areas											
Grass (perm.)	(ha)						45	35.9	18.8	19	34.2
Maize (silage)	(ha)	1.5	2	0.6	0.1	1	3.7	7.7	9.2	16.4	8.2
Onions	(ha)	2.2	3.2	9.7	23.4	8.9			0.2	0.6	0.1
Potatoes	(ha)	4.8	17.9	24.8	3.5	17.1		0.2	0.4	1	0.3
Set-aside	(ha)	1.7	1.8	1.3	0.9	1.4	0.1	0.2	3	5.4	0.8
Sugar beet	(ha)	3.1	11.2	9.1	1.3	7.8	0.3	0.3	0.3	1.6	0.4
Wheat (soft)	(ha)	2.7	10.4	11.5	2	8.7		0.1		1.7	0.2
Other crops (not simulated)	(ha)	2	19.8	11.7	2.8	11.5	0.6	0.3	1.2	3.1	0.6
Animals											
Bulls	(heads)						3	2	1	2	2
Calves	(heads)						24	25	23	31	25
Cows	(heads)						58	74	65	76	72
Heifers	(heads)						17	21	19	22	21
Total	(heads)						102	122	107	131	120

The calibrated model is used to predict changes in total gross margin, agricultural supply and environmental indicators as a consequence of the implementation of the 2003 CAP reform in a context of market liberalisation. The time horizon of the simulation is the year 2013 and takes into account (i) new exogenous prices generated by the CAPRI agricultural sector model (Britz *et al.*, 2007) under the market liberalization scenario, (ii) abolishment of the existing quota for sugar beet and milk, (iii) abolishment of the obligatory set-aside policy and (iv) new subsidies calculated under the CAP reform of 2003.

The data used for policy scenarios are yields and price trends for year 2013 as calculated by CAPRI (Britz *et al.*, 2007) in a market liberalization scenario. The market liberalization scenario in CAPRI assumes abolishing the export tariffs. It should be noted that this market liberalization scenario within CAPRI does not include abolishment of quota or of the obligatory set-aside policy. Input and output prices are inflated with a constant inflation rate (1.9%) in both regions for a period of 10 years. Historical yields, subsidy levels (as those determined in the CAP reform of 2003) and region-specific decoupling of subsidies from production, were also considered.

Table 2: Farm specific data of farm types in Midi-Pyrenees in 2003, and, observed crop areas and animal numbers that are included in the current activities. Source: FADN.

		Arable farms				Dairy farm
		FT 3201	FT 3202	FT 3304	Aver. farm	FT 2206
Farm specific data						
Total available land	(ha)	141.2	123.8	173.1	148.7	41.6
Irrigated land	(ha)	41.8	30.4	16.5	30.9	0.8
Family labour	(hrs)	2902	3260	3179	3067.2	2152
Milk quota	(tons)					171
Cost of hired labour	(€/hr)	7.5	7.5	7.5	7.5	7.5
Farms represented	(%)	46	20	34		100
Crop areas						
Barley	(ha)	4.1	1.6	2.4	3	2
Grass (permanent)	(ha)					28.4
Maize (grain)	(ha)	35.1	25.1	3.6	22.3	0.1
Maize (silage)	(ha)	0.3	0.5	0.7	0.5	6.3
Peas	(ha)	3.7	3.6	6.4	4.6	0.3
Rape seed	(ha)	1.7	1	1.6	1.5	
Set-aside	(ha)	9.3	18.9	9.4	11.2	0.5
Soya	(ha)	3	3.6	7.8	4.8	
Sunflower	(ha)	14.3	12.6	33.9	20.7	
Wheat (durum)	(ha)	17.3	11.4	31.6	21.1	
Wheat (soft)	(ha)	13.1	12.3	13.2	13	0.9
Other crops (not simulated)	(ha)	39.3	33.2	62.4	46	3.3
Total	(ha)	141.2	123.8	173.1	148.7	41.7
Animals						
Bulls	(heads)					1
Calves	(heads)					10
Cows	(heads)					29
Heifers	(heads)					7
Total	(heads)					47

Table 3: Crop product and animal product prices, yields, subsidies, costs and gross margins in 2003 and 2013 in Flevoland. Source: FADN, SEAMLESS survey and CAPRI model

	Price		Yield		Subsidy		Costs		Gross margin		
	(€/tonne)		(tons/ha or tons/head)		(€/ha or €/head)		(€/ha or €/head)		(€/ha or €/head)		
	2003	2013	2003	2013	2003	2013	2003	2013	2003	2013	Change (%)
Crop products											
Maize fodder	30	34	40.8	42.9		448	1098	1329	126	567	350
Onions	90	109	58.4	61.4		7	2158	2611	3098	4100	32
Potatoes	100	74	40.9	40.5		91	2252	2725	1838	340	-81
Set-aside					298		100	121	198	-121	-161
Soft wheat (spring)	120	142	7.8	8.7	298	234	527	638	707	836	18
Soft wheat (winter)	130	154	8.6	9.6	298	234	524	634	892	1082	21
Sugar beet	75	48	65.5	70.6			1150	1392	3763	2018	-46
Grass products (dry matter)											
Grass (grazed)			6.0	6.6			267	323	-267	-323	-21
Grass (silage)			4.0	4.4							
Animal products											
Bull (meat)	700	695	0.0	0.0							
Calves (meat)	108	143	0.0	0.0							
Cows (meat)	650	645	0.2	0.2							
Cows (milk)	320	275	7.5	8.9							
Herd unit					31	59	749[a]	906	720	633	-12

[a] Average costs before calibration, feed costs are not included.

Weighted average economic and policy data (prices, yields subsidies, costs and gross margins) for the base year and the 2003 CAP reform for Flevoland and Midi-Pyrenees under the market liberalization scenario are presented in Table 3 and Table 4, respectively. The weights are determined from the share of activities observed in each farm type and the share of farm types in the regions. Based on this information, in Flevoland the expected gross margins of silage maize, onions and soft wheat are projected to increase in 2013 while the expected gross margins of potatoes and sugar beet are projected to decrease substantially. With the CAP reform of 2003 silage maize receives a larger subsidy than other crops. The main reason for this is that most of the silage maize area is at dairy farms which receive a larger subsidy per ha because of the decoupled animal production. Grass products are assumed to be non-tradable products and thus have no price in the model. The expected gross margin decrease of grass is due to increasing costs because of inflation. The large decrease in the price of milk is associated with the market liberalization scenario and it is the reason for the lower expected gross margin per herd unit. In Midi-Pyrenees, the expected gross margins of most crops increase due to higher prices and subsidies. An exception is durum wheat for which the subsidy decreases by almost 67% resulting in a substantial decrease of expected gross margin. Inflation of the costs is the main reason for the lower expected gross margin of grass while the lower subsidy for set-aside is the main reason for lower expected gross margin of the fallow activity. Similar to Flevoland the average expected gross margin of a herd unit is reduced due to projected lower milk price.

Table 4: Crop product and animal product prices, yields, subsidies, costs and gross margins in 2003 and 2013 in Midi-Pyrenees. Source: FADN, SEAMLESS survey and CAPRI model

	Price (€/tonne)		Yield (tons/ha or tons/head)		Subsidy (€/ha or €/head)		Costs (€/ha or €/head)		Gross margin (€/ha or €/head)		
	2003	2013	2003	2013	2003	2013	2003	2013	2003	2013	Change (%)
Crop products											
Barley	94	101	5.0	5.2	304	452	340	411	434	567	23
Maize (grain)	120	152	11.0	10.5	304	431	859	1039	765	993	23
Maize (silage)	120	132	15.4	17.3		423	860	1041	988	1657	40
Peas	133	150	3.5	3.6	304	448	385	466	385	526	27
Rape seed	204	318	2.2	2.2	304	443	582	704	171	451	62
Set-aside					304	156			304	156	-95
Soya	196	318	2.3	3.1	304	450	331	401	424	1027	59
Sunflower	213	323	2.4	2.4	304	451	294	356	521	871	40
Wheat (durum)	135	148	5.0	5.8	592	198	421	509	846	546	-55
Wheat (soft)	116	137	6.5	7.0	304	444	430	520	628	879	29
Grass products (dry matter)											
Grass (grazed)			2.3	2.5							
Grass (hay)			3.1	3.4			72	87	-72	-87	-21
Grass (silage)			4.6	5.0							
Animal products											
Bull (meat)	1200	1191	0.0	0.0							
Calves (meat)	110	146	0.0	0.0							
Cows (meat)	600	595	0.2	0.2							
Cows (milk)	320	258	6.0	7.3							
Herd unit					30	31	405[a]	490	1023	800	-22

[a] Average costs before calibration, feed costs are not included.

Three model runs were designed to analyse the effects of the different changes during the period 2003 – 2013 (see Table 5). In the first model run (*price-yield change*) we included only price and yield changes and inflated input prices for year 2013, assuming market liberalisation. In the second model run (*set-aside & quota abolishment*) we added the abolishment of the obligatory set-aside policy and the quota regimes for both sugar beet and milk. In the third model run (*CAP 2003*) we added the CAP reform of 2003. In this model run we recalculated subsidies according to the CAP reform of 2003 where decoupling of subsidies from production was decided. Notice that only model run 3 can be considered as a complete policy scenario (all interrelated changes are taken into account simultaneously), the other model runs serve to analyse the effects of the individual changes during period 2003-2013.

Table 5: Definition of the base year and the model runs (price-yield change, *set-aside & quota abolishment* and *CAP 2003*)

	Exogenous assumptions	Price & Yield	Set-aside and quota policies	EU compensation payment
Base year [2003]		2003 price and yield	With obligatory set-aside and quota	Agenda 2000 (direct payment)
Price-yield change [2013]	Inflation rate of 1.9% per year	Projection in prices and yields from 2003 to 2013 accounting for market liberalization	With obligatory set-aside and quota	Agenda 2000 (direct payment)
***set-aside & quota abolishment* [2013]**	Inflation rate of 1.9% per year	Projection in prices and yields from 2003 to 2013 accounting for market liberalization	Abolishing set-aside obligation and quota	Agenda 2000 (direct payment)
***CAP 2003* [2013]**	Inflation rate of 1.9% per year	Projection in prices and yields from 2003 to 2013 accounting for market liberalization	Abolishing set-aside obligation and quota	2003 CAP reform (decoupled payment)

2.5. Results

In this section weighted average results of different farm types in the two regions are presented; the weights are determined from the relative share (based on number of farms represented) of the farm types in the region, i.e. first, the average farm of each farm type is simulated and then the results were used to calculate weighted average values of arable and dairy farms in each region. The regional average simulated crop levels, the regional average economic results and the calculation of the regional average nitrogen use of arable farms in Flevoland are presented in Figure 3. Because of the PMP calibration, the simulated crop levels for the base year are exactly the same as the actual levels observed in FADN (Table 1). In the *price-yield change* model run, the gross margin increase of maize silage, onions and wheat causes a substantial increase in the areas of these crops in arable farming. The gross margin decrease of potatoes and sugar beet causes a decrease in the area of these crops. The decrease of the area of sugar beet is also because of the yield trend (8% increase). Less area of sugar beet are needed to produce the same quota. The average total gross margin of arable farms decreases with more than 28%. The shift of crop production from spring soft wheat to winter soft wheat is the main reason for the increase of the total nitrogen use per ha in all farm types of Flevoland.

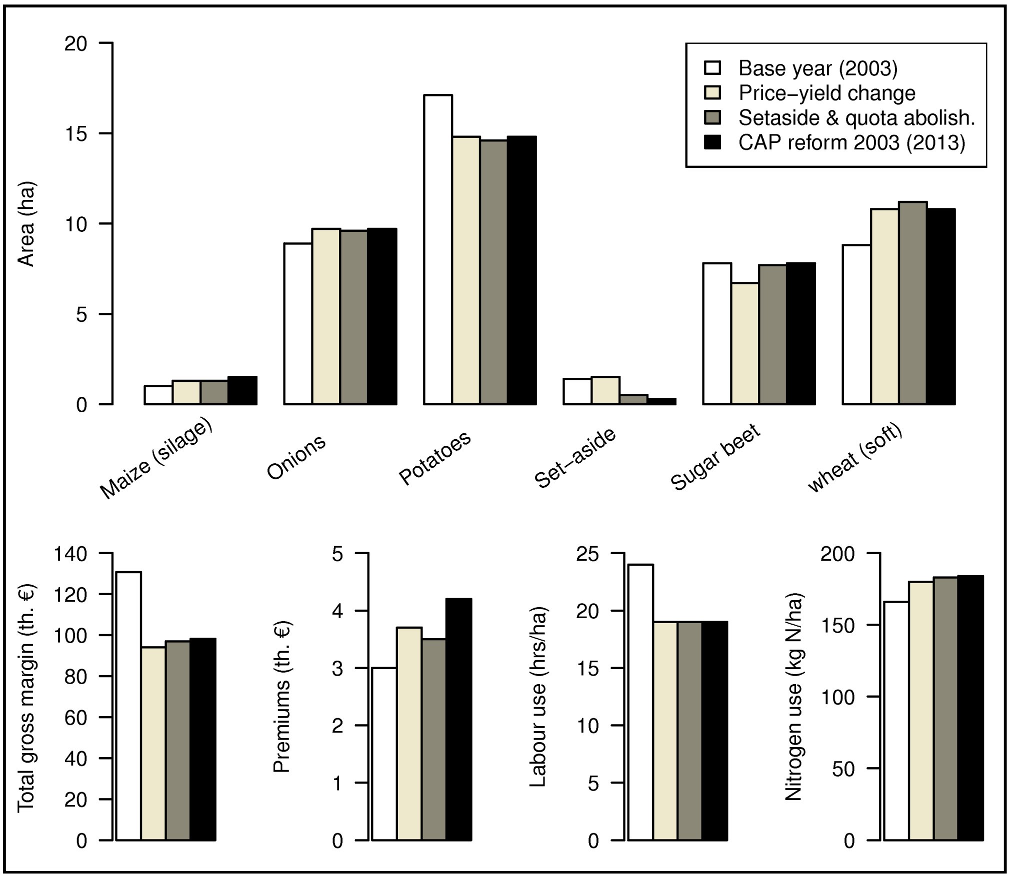

Figure 3: Simulated results for the base year (2003) simulation and 3 model runs (2013) for an average arable farm of Flevoland (*price-yield change*: yield and price trend, inflation of input prices, *set-aside & quota abolishment*: *price-yield change* +abolishment of obligatory set-aside policy and quotas, *CAP 2003*: *price-yield change* + *set-aside & quota abolishment* + CAP reform of 2003).

In the *set-aside & quota abolishment* model run for arable farms, the abolishment of the sugar beet quota system and the obligatory set-aside policy, are the reasons for the increase of the area of sugar beet and the decrease of the area of set-aside (compared to the simulated levels of these activities in the *price-yield change model run*). Putting the set-aside area in production causes an increase of the average total gross margin of arable farms. The total nitrogen use increased in all farm types because of the decrease of the area of set-aside and the increase of the area of the more nitrogen demanding winter soft wheat.

In the *CAP 2003* model run, overall effects on crop allocation are very modest compared to the *set-aside & quota abolishment*, model run and the associated effect on the total gross margin is negligible.

The regional average weighted results of the simulated dairy farms in Flevoland are presented in Figure 4. The produced feed reported in Figure 4 corresponds to on-farm feed production that is used on-farm excluding sold quantities of on-farm produced feed. Similar to the base year simulation of the arable farms and because of the PMP calibration, the observed activity levels of crops and animals of the dairy farms are reproduced exactly. In the *price-yield change* model run the total number of animals of the herd decreases because of the increased milk production per cow and the given quota. The area of permanent grassland decreases but the on-farm feed production of grass increases because of the assumed yield increase. The amount of silage maize sold increases because of the price increase. The share of grass in the diet increases and as a result the amount of concentrates also increases to fulfil the animals energy requirements while respecting their intake capacity. The gross production decreases mainly because of the decrease in the price of milk. The total costs increase because of the higher input and feed prices and the increased feed requirements. As a result, the total gross margin decreases by almost 35%. The total nitrogen use remains almost the same in all dairy farm types of Flevoland.

In the *set-aside & quota abolishment* model run where the milk quota is abolished, the total number of animals increases by 1.7% compared with the base year simulation and by almost 13% from the *price-yield change* model run. The increased feed requirements are covered by increasing purchases of concentrates and silage maize. The total gross margin increases by almost 16% from the total gross margin of the *price-yield change* model run while the total nitrogen use remained almost the same.

In the *CAP 2003* model run, the CAP reform of 2003 and mainly the large increase of the subsidy for maize silage causes a shift of production from grass to maize silage. The received premiums under the CAP reform of 2003 for dairy farms increase substantially, causing a modest increase of the farm’s total gross margin.

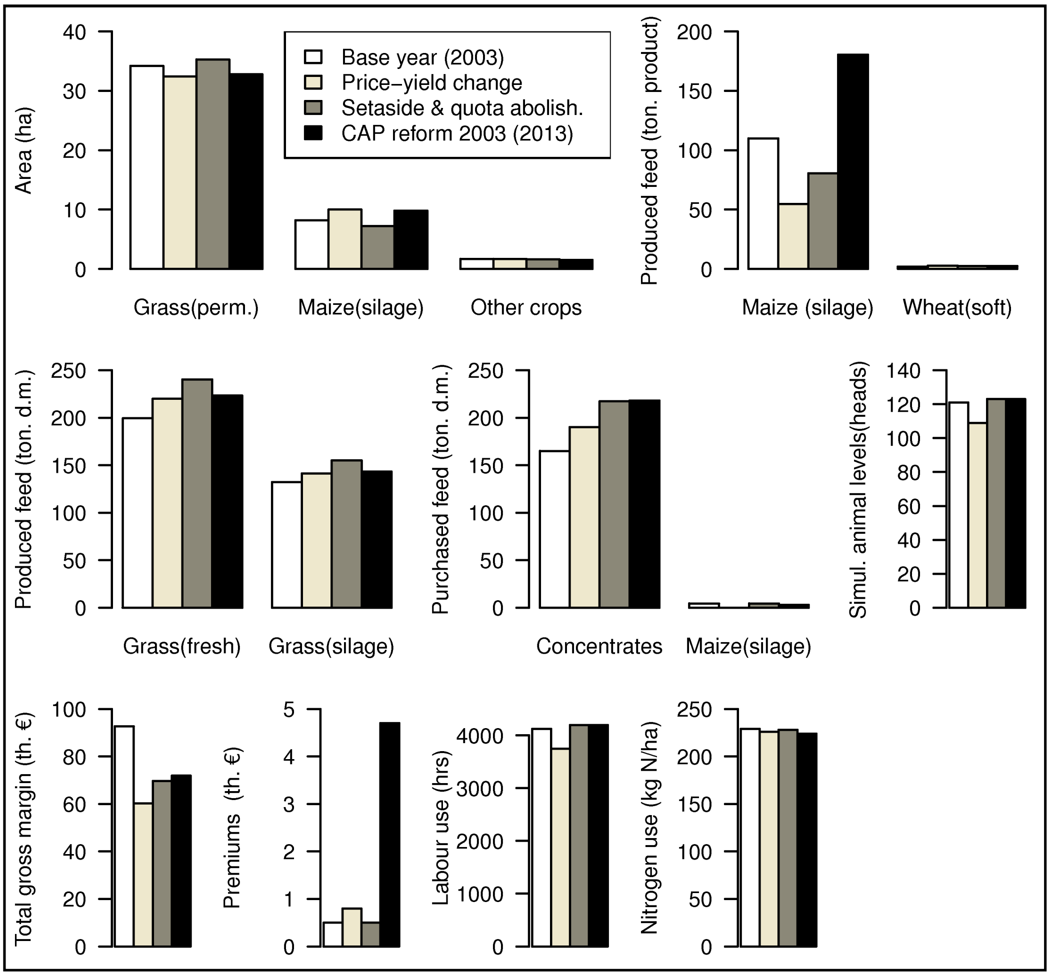

Figure 4: Simulated results for the base year (2003) simulation and 3 model runs (2013) for an average dairy farm of Flevoland (*price-yield change*: yield and price trend, inflation of input prices, *set-aside & quota abolishment*: *price-yield change* +abolishment of obligatory set-aside policy and quotas, *CAP 2003*: *price-yield change* + *set-aside & quota abolishment* + CAP reform of 2003).

The regional average weighted results from the application of FSSIM to the arable farms of Midi-Pyrenees are presented in Figure 5. In the *price-yield change* model run, the predicted changes for 2013 of gross margins resulted in an increase of the areas of soya, rape seed and silage maize and a decrease in the areas of barley and peas. The average total gross margin of arable farms increases by 24%. The main reason for this is the large price increase of oil seed crops.

In the *set-aside & quota abolishment* model run the set-aside obligation of arable farms is abolished putting almost 70% of the set-aside area of the *price-yield* change model run into production. The set-aside land is allocated to all other crops. The intensification of production caused a large increase of the average total gross margin of arable farms but also a substantial increase of the total nitrogen use compared to the *price-yield change* model run.

The recalculation of subsidies according to the CAP reform of 2003 caused a large increase of the received subsidies for most crops. Exceptions are the subsidies for durum wheat and set-aside land, which decrease by 67 and 49%, respectively, causing a decrease of the average area of these activities. The total gross margin decreased by 1.5% compared with the *set-aside & quota abolishment*, model run and increased by 24 and 54% compared with the *price-yield change model run* and base year, respectively.

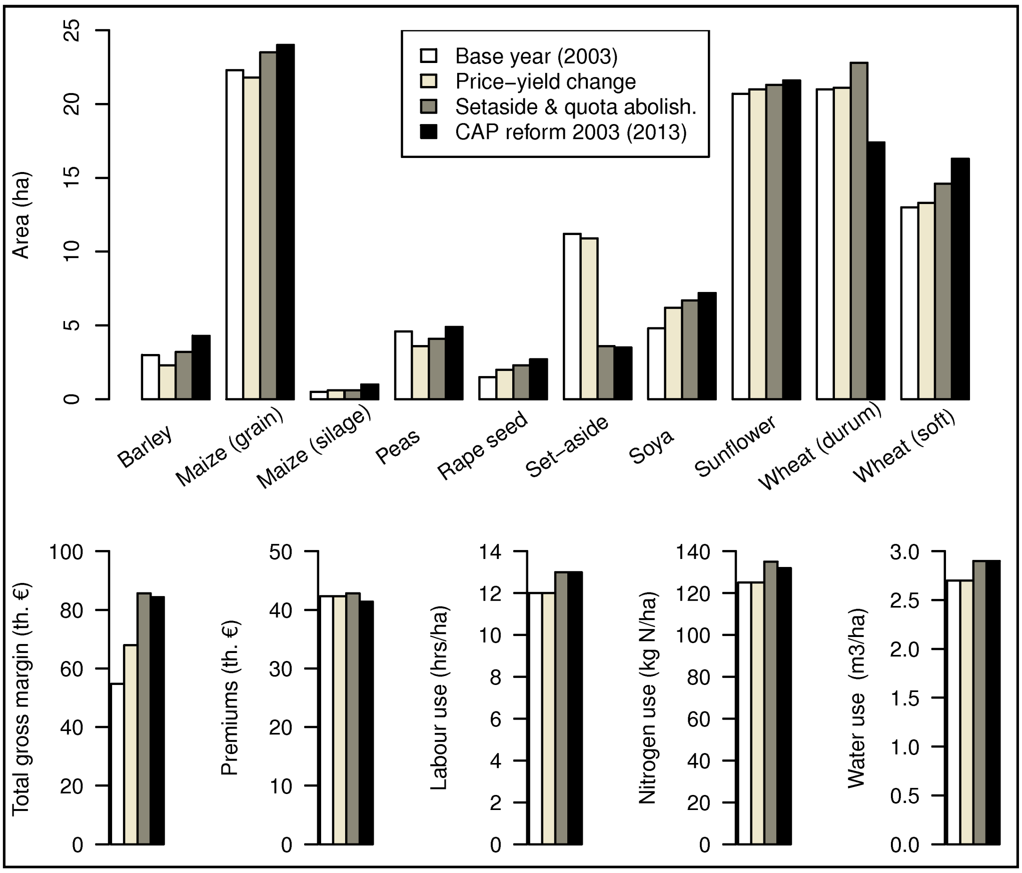

Figure 5: *Simulated* results for the base year (2003) simulation and 3 model runs (2013) for an average arable farm of Midi-Pyrenees (*price-yield change*: yield and price trend, inflation of input prices, *set-aside & quota abolishment*: *price-yield change* +abolishment of obligatory set-aside policy and quotas, *CAP 2003*: *price-yield change + set-aside & quota abolishment* + CAP reform of 2003).

Regional average weighted results from the application of FSSIM to the dairy farms of Midi-Pyrenees are presented in Figure 6. Similar to the dairy farm of Flevoland, the produced feed reported in Figure 6 corresponds to on-farm feed production that is used on-farm, excluding sold quantities of on-farm produced feed. In the *price-yield change* model run, the substantial increase of feeding costs and input prices; and the price decrease of milk caused a small decrease in the average herd size in Midi-Pyrenees. The area of permanent and temporary grasslands decreases and it is substituted mainly by silage maize and barley. On-farm produced grass and the more expensive purchased concentrates in this model run are substituted by cereals and silage maize to cover the animal's feed requirements. The gross margin decreases by 6%.

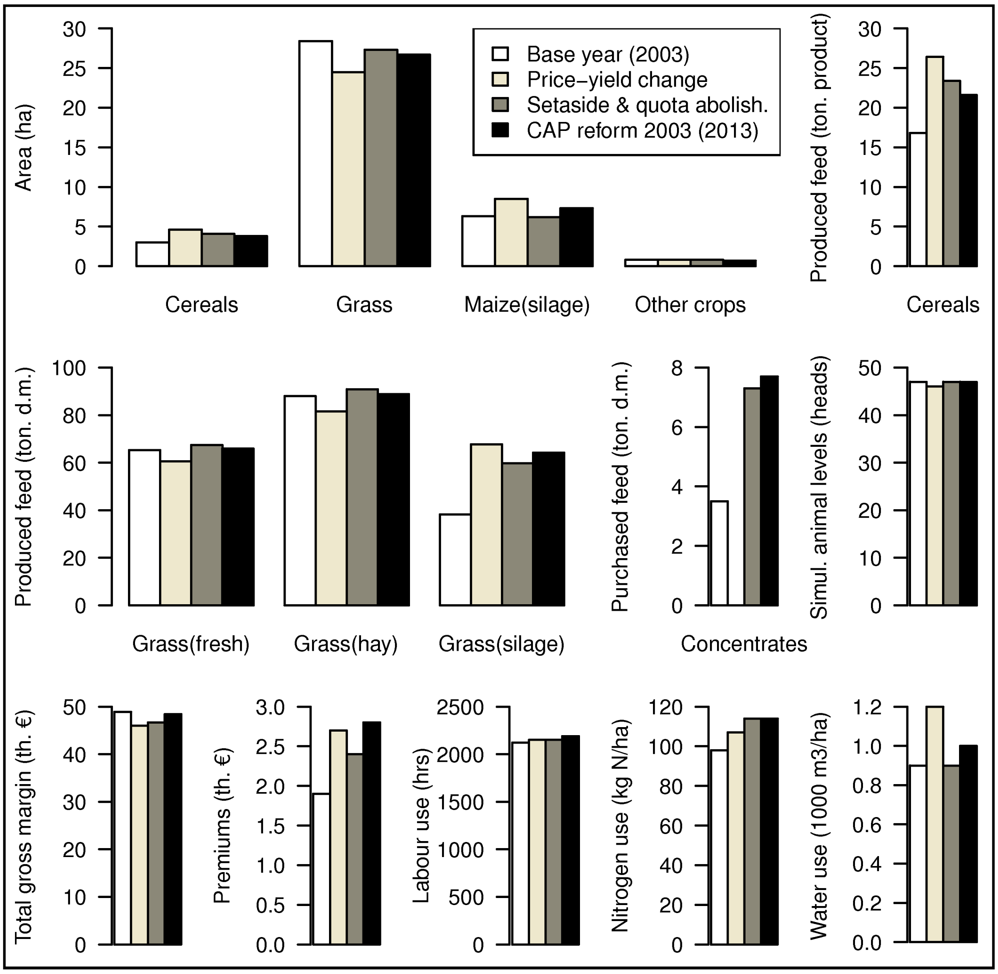

Figure 6: Simulated results for the base year (2003) simulation and 3 model runs (2013) for an average dairy farm of Midi-Pyrenees (*price-yield change*: yield and price trend, inflation of input prices, *set-aside & quota abolishment*: *price-yield change* +abolishment of obligatory set-aside policy and quotas, *CAP 2003*: *price-yield change* + *set-aside & quota abolishment* + CAP reform of 2003).

Abolishment of the milk quota policy in the *set-aside & quota abolishment* model run increases the average herd size back to the level of the base year. Labour availability becomes a binding constraint and therefore the number of animals does not exceed the number of animals observed in the base year. The grassland area increases compared to the *price-yield change* model run substituting the area with silage maize which is not fed to the animals. In both Flevoland and Midi-Pyrenees it is expected that the yield of milk in

2013 will have increased with 19 and 22%, respectively, causing an increase in the animals' feed requirements. To cover the additional requirements for feed, more grass silage and purchased concentrates are needed. Abolishment of the milk quota caused an increase to the farm's total gross margin.

The effects of the CAP reform 2003 (tested in the *CAP 2003* model run) relative to the results of the *set-aside & quota abolishment* model run are marginal. The large increase of subsidies on maize silage in the *CAP 2003* model run caused an increase of the area of silage maize and a decrease of the area of cereals (mainly barley) and grassland compared to the *set-aside & quota abolishment* model run. The decreased amount of barley fed to the animals is compensated by purchased concentrates. A small amount of hired labour is needed to cover the additional labour requirements of maize. Water and nitrogen use remain almost the same in all model runs.

2.6. Discussion and Conclusion

In this article, a bio-economic farm model has been presented that is modular and can be used to simulate the responses of farms to agricultural and environmental policies in a broad range of contexts that may occur in the EU27. This was achieved by: (i) separating model and data and creating a consistent European database for farm types, their locations and production activities, (ii) designing the model in a modular way, that allows switching on and off modules, constraints or calibration methods, (iii) providing adequate documentation, and (iv) ensuring public availability. The arable and dairy farms of two regions that differ substantially from a bio-physical and socio-economic point of view were simulated successfully, using information mainly available in a large EU-wide database (i.e. FADN) and a relatively simple survey conducted within SEAMLESS for a sample of regions representative for the EU27. The PMP based calibration of FSSIM does not require additional region-specific knowledge and detailed information on specific constraints to guarantee exact calibration. Nevertheless, availability of this kind of information could be easily exploited and used to improve the forecasting performance of the model.

The market of land, possibilities for off-farm labour and structural changes are usually issues exogenous to the system definition of bio-economic farm models (Janssen and Van Ittersum, 2007). This is how these issues were also treated in the model presented in this article, but we have indicated in Section 2 and 3 how they can be partially dealt with, using FSSIM and a scenario approach. To simulate farm structural change and land markets more comprehensively FSSIM needs to be combined with other models that account for market and sector level changes, as has been attempted in the SEAMLESS modelling framework (Perez Dominguez *et al.*, 2009; Zimmermann *et al.*, 2009).

In the present article we illustrate the standalone value of FSSIM using applications in two regions and different farming systems. The applications raise a number of discussion points because of a number of decisions concerning the set-up of the model. First, the presented applications were based on data available in the Farm Accountancy Data Network and a simple survey on agricultural management. This led to a restricted set of environmental indicators, i.e. the total amount of water used for irrigation, and the total amount of nitrogen used. This hinders a comprehensive overview of the environmental implications of the market liberalization under the CAP reform of 2003. The use of a bio-physical model to calculate technical coefficients that can easily be exploited in FSSIM would increase the number of environmental indicators and thus improve the overall assessment of the environmental impact of the tested scenario. However, this requires detailed agro-management data (timing and precise quantities of inputs per crop) that are not available in pan-European data-sets.

Second, we used an average farm type in our simulations to ensure that all important crop products that are produced by farms of a specific farm type will be part of the simulated production plan. This is very important for the type of analysis that requires full representation of agricultural production to determine equilibrium between supply and demand, such as in SEAMLESS (Van Ittersum *et al.*, 2008). However, simulating the average farm has also important drawbacks. An average farm and an average farmer do not actually exist and consequently, an average activity pattern also does not exist. The activity pattern of the average farm is much more diversified than that of individual farms. Reproducing such a cropping pattern using an LP model would require a large number of binding constraints. It is possible that such constraints do not even exist in reality and consequently they are difficult to define (e.g. rotational constraints of an “average” production plan). In such cases, calibration of the LP model is necessary for reproducing the observed activity levels and often calibration will dominate the simulations. It is possible that the impact of calibration on the results of the model would be reduced substantially if a number of individual farms were simulated instead of a single average farm. However, this would also have increased the computational requirements and individual farm data would have to be available which is usually not the case (individual farm data are usually confidential and not available for research).

Finally, we assume a yield trend (based on forecasts of the sector model CAPRI) to represent technological innovation. However, the rapid changes in the socio-economic and the bio-physical environment might lead to a broader variety of alternative activities that will become available to farmers in the future with even completely different inputs and outputs. Such alternative activities can not be ignored and should be taken into account in ex-ante evaluation of agricultural and environmental policies. Offering alternative activities in FSSIM is possible from a technical point of view. The difficulty is to identify a consistent and feasible set of alternative activities for all regions across the EU.

Apart from using our intuition to assess the model's forecasting performance, it is very difficult to evaluate the results in a quantitative and more objective way because they refer to future events and they use simulated data to account for price and yield trends. The quality of the results of FSSIM has been previously evaluated and assessed in ex-post experiments that demonstrate the capacity of the model to simulate the future behaviour of the farmer (Kanellopoulos *et al.*, 2010). Even though, the results of such ex-post exercises cannot be generalized they do increase the confidence in the model's predictions.

A well calibrated and tested bio-economic farm model can be used for ex-ante assessment of the impacts of new policies. Different farming systems across EU can be affected in different ways and consequently farmers respond differently when they are confronted with market and policy changes. This was confirmed by the results presented in this article. For example, price and yield changes are the main factor explaining the gross margin decrease of farms in Flevoland. In Midi-Pyrenees, simulated price and yield changes have the opposite effect on the total gross margin of arable farms, and for this region the abolishment of obligatory set-aside has an additional positive effect on the total gross margin of arable farms. In Flevoland farms showed an increase in premiums under the CAP 2003 reform scenario, whereas in Midi-Pyrenees the CAP 2003 scenario did not further increase the already high level of premiums. The variation in farm's behaviour should be taken into account for efficient and effective policy assessment. Bio-economic modelling can be a useful tool for exploring this variation.

FSSIM has been set-up such that it can readily simulate farm types in very different contexts (climate, soils and socio-economic conditions) and for different purposes. The presented examples in this paper show a fairly detailed analysis for the farm types of two regions. The reusability of the model was confirmed by the significant number of applications that have been published (Louhichi *et al.*, 2008; Kanellopoulos *et al.*, 2009; Louhichi *et al.*, 2009; Majewski *et al.*, 2009; Mouratiadou *et al.*, 2010; Traoré *et al.*2009). Pérez Domínguez *et al.* (2009) show how results of the model can be used for linking micro and macro level analysis of market changes. The model is available under an Open Source license (www.seamlessassociation.org) and through its broader use it can be further tested and new modules can be added.

2.7. Acknowledgements

The work presented in this publication is funded by the SEAMLESS Integrated Project, EU 6th Framework Programme for Research Technological Development and Demonstration, Priority 1.1.6.3 Global Change and Ecosystems (European Commission, DG Research, contract no.010036-2).

2.8. References

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D. (2007). Farm management indicators and farm typologies as a basis for assessment in a changing policy environment. Journal of Environmental Management, 82: 352-362.

Berentsen, P.B.M. (2003). Effects of animal productivity on the costs of complying with environmental legislation in Dutch dairy farming. Livestock Production Science, 84: 183-194.

Borkowski, N., Hecker, J.-M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G., Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 1: 33-34.

Britz, W., Pérez, I., Zimmermann, A., Heckelei, T. (2007). Definition of the CAPRI Core Modelling System and Interfaces with other Components of SEAMLESS-IF. SEAMLESS Report No. 26, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, pp. 114 (www.seamlessassociation.org).

CEC (1991). The Development and Future of the Common Agricultural Policy. COM(91) 258 Final.

CEC (1999[a]). Agricultural council: political agreement on CAP reform. DG Agriculture Newsletter, Special Edition March 11, 1999, Brussels.

CEC (1999[b]). Berlin European Council: Agenda 2000, Conclusions of the Presidency. DG Agriculture Newsletter, April 1999, Brussels.

CEC (2003). Reform of the Common Agricultural Policy: Medium-term Prospects for Agricultural Markets and Income in the European Union, December, Brussels (http://europa.eu.int/comm/agriculture /publi/caprep/prospects2003b/fullrep.pdf).

Donaldson, A.B., Flichman, G., Webster, J.P.G. (1995). Integrating agronomic and economic models for policy analysis at the farm level: the impact of CAP reform in two European regions. Agricultural Systems, 48: 163-178.

EC (2005). Impact Assessment Guidelines. European Commission, Brussels, pp. 99.

EEC, 1985. Commission Decision of the 7th of June 1985, Establishing a Community Typology for Agricultural Holdings (OJ L 220, 17.8.1985, p. 1).

Elbersen, B., Kempen, M., Van Diepen, K., Hazeu, G., Verhoog, D., Andersen, E. (2006). Protocols for Spatial Allocation of Farm Types. SEAMLESS Report No. 19, SEAMLESS integrated project, EU 6th Framework Programme, contract No. 010036, org, pp. 99 (www.seamlessassociation.org).

Flichman, G. (1996). L'evaluation micro-economique de la reforme de la PAC par un modele bio-economique: une illustration sur la region de Toulouse. In: Reyniers, F.N., Benoit-Cattin, M. (Eds.), Linking-agricultural-models:-seminar-proceedings. Montpellier, France: CIRAD-CA., IAM, Montpellier, France., pp 71-75.

Flichman, G., Jacquet, F. (2003). Couplage des modèles agronomiques et économiques: intérêt pour l'analyse des politiques. Cahiers d'économie et Sociologie Rurales, 67: 1-20.

Hazell P., and Norton R. (1986). Mathematical Programming for Economic Analysis in Agriculture. Macmillan, New York, 1986.

Hazeu, G., Elbersen, B., Andersen, E., Baruth, B., Van Diepen, K., Metzger, M. (2010). A bio-physical typology for a spatially-explicit agri-environmental modeling framework. In: Brouwer, F., Van Ittersum, M.K. (Eds.), Environmental and Agricultural Modelling: Integrated Approaches for Policy Impact Assessment. Springer Academic Publishing, Dordrecht.

Heckelei, T. (2003). Positive Mathematical Programming: Review of the Standard Approach, CAPRI Working Papers 97-03 (http://www.agp.uni-bonn.de/agpo/rsrch/ capri/ wrkpap_e.htm).

Heckelei, T., Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalized maximum entropy. European Review of Agricultural Economics, 30: 27-50.

Howitt, E.R. (1995). A calibration method for agricultural economic production models. Journal of Agricultural Economics, 46: 147-159.

Janssen, S., van Ittersum, M.K. (2007). Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems, 94: 622-636.

Janssen, S., Andersen, E., Athanasiadis, I., Van Ittersum, M. (2009a). A database for integrated assessment of European agricultural systems. Environmental Science & Policy, 12: 573-587.

Janssen, S., Van Ittersum, M.K., Meuter, E., Louhichi, K., Flichman, G., Blanco, M., Zander, P., Borkowski, N., Hecker, M., Stokstad, G., Romstad, E., Kanellopoulos, A., Berentsen, P., Hengsdijk, H., Van Keulen, H., Thorne, P., Thornton, P., Li, H., Rizzoli, A.E., Heckelei, T. (2009b). Integration of all FSSIM components within SEAMLESS-IF and a Standalone Graphical User Interface for FSSIM, SEAMLESS

Report No. 38, SEAMLESS Integrated Project, EU 6th Framework Programme, contract No. 010036-2 (www.seamlessassociation.org).

Jansson, T. (2007). Econometric Specification of Constrained Optimization Models. PhD dissertation, Bonn University, electronic publication (http://www.ilr1.uni-bonn.de/agpo/rsrch/capri/ wrkpape .htm).

Jarrige, R. (1989). Ruminant Nutrition. Recommended Allowances and Feeding Tables. INRA. Paris, France.

Judez, L., Chaya, C., Martinez, S., Gonzalez, A.A. (2001). Effects of the measures envisaged in "Agenda 2000" on arable crop producers and beef and veal producers: an application of Positive Mathematical Programming to representative farms of a Spanish region. Agricultural Systems, 67: 121-138.

Kanellopoulos, A., Hengsdijk, H., Janssen S., Van Ittersum, M.K., (2009), Ex-ante assessment of the abolishment of the EU set aside policy: results from a bio-economic farm analysis. In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP 2009). Egmond aan Zee, The Netherlands, 10-12 March 2009. Wageningen University and Research Centre, Wageningen, pp. 478-479.

Kanellopoulos, A., Berentsen, P.B.M., Heckelei, T., Van Ittersum, M.K., Oude Lansink, A.G.J.M. (2010). Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. Journal of Agricultural Economics, 61: 274-294.

Louhichi, K., Hengsdijk, H., Janssen, S., Bigot, G., Perret, E. (2009). EU dairy farming in the face of change: an exploration using a bio-economic farm model. In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Egmond aan Zee, The Netherlands, 10-12 March 2009 Wageningen University and Research Center, Wageningen, pp. 262-263.

Louhichi K., Belhouchette H., Flichman G., Therond O. and Wery J. (2008). Impact assessment of the 2003 CAP reform and the nitrate directive on the arable farming system in the Midi-Pyrénées region: bioeconomic Modelling at field, farm and regional levels. In 109th Seminar, November 20-21, 2008, Viterbo, Italy.

Majewski, E., Was, A., Belhouchette, H., Louhichi, K., Mouratiadou, I. (2009). Impact Assessment of policy changes on the arable sector using the FSSIM model: Case study of the Zachodniopomorskie NUTS region. In: Proceedings of the conference on

Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Egmond aan Zee, The Netherlands, 10-12 March 2009 Wageningen University and Research Center, pp. 484-485.

Mouratiadou, I., Russesll, G., Topp, C., Louhichi, K., Moran, D. (2009). Modelling common agricultural policy-water framework directive interactions and cost-effectiveness of measures to reduce nitrogen pollution. Water Science & Technology, 61: 2689-2697.

OECD (2004). Analysis of the 2003 CAP Reform. OECD Publication. 53p.

Onate, J.J., Atance, I., Bardaji, I., Llusia, D. (2006). Modelling the effects of alternative CAP policies for the Spanish high-nature value cereal-steppe farming systems. Agricultural Systems, 94: 247-260.

Pérez Domínguez, I., Bezlepkina, I., Heckelei, T., Romstad, E., Oude Lansink, A., Kanellopoulos, A. (2009). Capturing market impacts of farm level policies: a statistical extrapolation approach using biophysical characteristics and farm resources. Environmental Science & Policy, 12: 588-600.

Riesgo, L., Gomez-Limon, J.A. (2006). Multi-criteria policy scenario analysis for public regulation of irrigated agriculture. Agricultural Systems, 91: 1-28.

Röhm, O., Dabbert, S. (2003). Integrating agri-environmental programs into regional production models: an extension of positive mathematical programming. American Journal of Agricultural Economics, 85: 254-265.

Semaan, J., Flichman, G., Scardigno, A., Steduto, P. (2007). Analysis of nitrate pollution control policies in the irrigated agriculture of Apulia Region (Southern Italy): A bio-economic modelling approach. Agricultural Systems, 94: 357-367.

Thorne, P.J., Hengsdijk, H., Janssen, S., Louhichi, K., Keulen, H. van, Thornton, P.K. (2009). Modelling Livestock Component in FSSIM. SEAMLESS Report No. 35, SEAMLESS Integrated Project, EU 6th Framework Programme, Contract No. 010036-2, pp. 73 (www.seamlessassociation.org).

Topp, C.F.E,, Mitchell, M. (2003). Forecasting the environmental and socio-economic consequences of changes in the common agricultural policy. Agricultural Systems, 76: 227-252

Traoré, B., Louhichi, K., Rapidel, B. (2009). A bio-economic model to analyse the performance of the cotton sector in Mali. In: Proceedings of the conference on Integrated Assessment of Agriculture and Sustainable Development Setting the

Agenda of Science and Policy (AgSAP). Egmond aan Zee, The Netherlands, 10-12 March 2009 Wageningen University and Research Center, pp 276-277.

Van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E., Wolf, J. (2008). Integrated assessment of agricultural systems - a component-based framework for the European Union (SEAMLESS). Agricultural Systems, 96: 150-165.

Veysset, P., Bebin, D., Lherm, M. (2005). Adaptation to Agenda 2000 (CAP reform) and optimisation of the farming system of French suckler cattle farms in the Charolais area: a model-based study. Agricultural Systems, 83: 179-202.

Zander, P., Borkowski, N., Hecker, J.-M., Uthes, S., Stokstad, G., Rørstad, P.K. Bellocchi G. (2010). Procedure to identify and assess current activities. SEAMLESS Report No. 50, SEAMLESS integrated project, EU 6th Framework Programme, Contract No. 010036-2 (www.seamlessassociation.org).

Zimmermann, A., Heckelei, T., Pèrez Dominguez, I., 2009. Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants. Environmental Science & Policy, 12: 601–618.

Chapter 3

3. Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants

Argyris Kanellopoulos[1,2], Paul Berentsen[1], Thomas Heckelei[3], Martin van Ittersum[2] and Alfons Oude Lansink[1]

[1] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*

[2] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*

[3] *University of Bonn, Economic and Agricultural policy, Nussallee 21, 53115 Bonn, Germany.*

Journal of Agricultural Economics 61 (2010): 274-294

Abstract

Using Linear Programming in bio-economic farm modelling often results in overspecialised model solutions. The positive mathematical programming (PMP) approach guarantees exact calibration to base year data but the forecasting capacity of the model is affected by necessary but arbitrary assumptions imposed during calibration. In this paper, a new PMP variant is presented which is based on less arbitrary assumptions that, from a theoretical point of view, are closer to the actual decision making of the farmer. The PMP variant is evaluated according to the predictions of the bio-economic farm model, developed within the framework for integrated assessment of agricultural systems in Europe (SEAMLESS). The forecasting capacity of the model calibrated with the standard PMP approach and the alternative PMP variant, respectively, are tested in ex-post experiments for the arable farm types of Flevoland (the Netherlands) and Midi-Pyrenees (France). The results of the ex-post experiments, in which we try to simulate farm responses in 2003 using a model calibrated to 1999 data, show that the alternative PMP variant improves the forecasting capacity of the model in all tested cases.

Keywords: agricultural policy; bio-economic models; environmental policy; farming systems; mathematical programming.

3.1. Introduction

Bio-economic farm models are often used to integrate model formulations of bio-physical processes with economic evaluations to simulate management decisions about resource allocation (Barbier and Bergeron, 1999). In many bio-economic studies, Linear Programming (LP) models have been used (Berentsen *et al.*, 1997; Acs *et al*, 2007; Janssen and van Ittersum, 2007). Solutions of LP models are, by definition, corner points of the feasible decision space. This implies that the number of selected activities cannot exceed the number of binding, policy, rotational and resource constraints which are included in the model. In practice, the number of binding policy, rotational and resource constraints are kept relatively small to avoid complexity and reduce data requirements. As a result, overspecialised model solutions occur (Heckelei, 2003). Estimating non-linear models using traditional estimation methods could reduce the problem of unrealistic simulation behaviour. However, traditional statistical estimators require multiple observations of farm inputs, outputs and prices which are not always available. In these cases a calibration procedure for LP models could be used to exploit existing information more efficiently and to reduce the gap between observed data and simulated results of bio-economic models.

Howitt (1995a) presented positive mathematical programming (PMP) as an elegant calibration procedure which guarantees exact reproduction of the base year activity levels, without additional, poorly justified calibration constraints. A decreasing marginal gross margin function is used to ensure that the base year activity levels are reproduced. The decreasing marginal gross margin function is justified by increasing variable costs per unit of production because of inadequate machinery and management capacity and decreasing yields due to land heterogeneity (Howitt, 1995a). In typical LP models not calibrated with PMP, increasing marginal costs are either omitted from the analysis or taken into account in an oversimplified way, resulting in unrealistic model solutions. An attractive feature of PMP calibration is that the model’s solution is closer to observed reality.

A second very attractive feature of PMP is that it is a generic procedure that can be fully automated. This means that it can be easily adapted and used for different regions and farm types without additional site specific information. This feature is important for sector, national and higher level analysis, where the data are limited, the knowledge on relevant policies/constraints is fragmented and the resources invested in developing a fully specified bio-economic model are restricted (Heckelei, 2003).

The standard PMP (STPMP) approach has been substantially criticised in the past for a number of limitations, extensively reviewed by Heckelei (2002). An important limitation of the STPMP approach is related to the arbitrary restrictions imposed on some of the model's parameters, especially the assumption that the gross margin of the least preferable activity is constant whereas gross margins of all other activities are assumed to decrease with increasing activity level. These restrictions are necessary in order to estimate the remaining parameters based only on one year of observations. This limitation of STPMP is described more explicitly later in this article and an approach to improve the justification of those restrictions is proposed.

A second important limitation of STPMP is related to the values of the shadow prices of the limiting resource constraints that are enforced in the STPMP approach. In many cases the shadow prices, and consequently the values of the limiting resources are underestimated leading to misspecification of the model's parameters and to unrealistic forecasts. This limitation of STPMP is also discussed later in this study, where additional information is used to retrieve more realistic values of the limiting resources. Another limitation of PMP is that the recovered parameters essentially embody marginal model misspecification of technology, data errors, aggregation bias, and representation of risk behaviour. Explicit description of the modelling assumptions is necessary to ensure a good interpretation of the model's parameters and results. Reliance on one year's observations of activity levels to recover the unknown parameters has been also criticised, since it does not allow estimation of the real value of parameters underlying the observed response behaviour of producers. Nevertheless, in most cases, calibration is used instead of estimation because of the lack of multiple year observations.

In recent years, a number of PMP variants have been developed and used for bio-economic analysis (Howitt, 1995b; Gohin and Chantreuil, 1999; Heckelei and Wolf, 2003; Röhm and Dabbert, 2003). The main objective of these PMP variants is to overcome the limitations of the STPMP approach and improve the forecasting capacity of models by utilising additional available information. In many cases these PMP variants are not sufficiently generic and have additional data requirements which are not always available. Although all variants guarantee exact calibration, simulation models of future behaviour, calibrated with different PMP variants, still produce different results (Heckelei and Britz, 2005). In PMP-calibrated models, the values of the unknown parameters are estimated in a way that exact calibration is ensured. As different assumptions are used, the values of the parameters are different between different variants, which clearly affect the responsiveness of the model to policy or strategy experiments.

An evaluation procedure is necessary to assess the forecasting capacity of calibrated bio-economic models and to increase users' confidence in the results of the analysis. A model evaluation reveals to model users the consequences of certain simplifications and assumptions and gives them a good overview of when and how the model should be used. Despite the importance of ensuring the quality of bio-economic analyses, model evaluation and validation are often not addressed adequately in existing bio-economic literature (Janssen and van Ittersum, 2007). Very few, if any, studies are available that test the forecasting performance of calibrated bio-economic models.

The research presented in this study was part of SEAMLESS which was a sixth framework EU project (Van Ittersum *et al.*, 2008). The main objective of SEAMLESS was to develop a model framework to be used for ex-ante assessment of agricultural and environmental policies at EU25 level. The framework was designed to be generic and modular, to enable analysis at multiple scales, to make it possible to address a variety of policy questions and to demonstrate the socio-economic and environmental consequences of multi-functional agricultural systems.

The Farm System SIMulator (FSSIM) (Janssen *et al.*, 2009; Louhichi *et al.*, 2009) is a bio-economic farm model developed within SEAMLESS for farm level analysis. FSSIM is used to simulate farmers' behaviour and future decisions and hence, price-supply relationships at farm level. A distinctive feature of FSSIM is that crop rotations are included in the model as activities instead of using rotational constraints to account for the important agronomic interactions between crops. A model solution can include several crop rotations. All crops of a rotation are grown every year on the same share of land. A crop grown in different rotations can have different technical coefficients accounting for interactions between crops. PMP is used to calibrate the model to the base year data and to improve its forecasting performance. A number of arbitrary assumptions are required to estimate the parameters of a non-linear cost function. These assumptions affect the predictive capacity of the model, which is an essential feature for ex-ante policy assessment.

The objectives of this study are, first, to highlight some limitations of the STPMP approach for this type of analysis; second, to present a PMP variant which overcomes those limitations, and improves the predictive capacity of the model; third, to compare the forecasting capacity of FSSIM calibrated with the two different PMP variants in “back-casting” (ex-post) experiments, providing evidence on the quality of the model's results.

In Section 2, the FSSIM framework is described. In Section 3, the theoretical basis of this study is formulated by presenting two PMP methods. An ex-post experiment for the arable farming systems of Flevoland (the Netherlands) and Midi-Pyrenees (France) is designed to compare the forecasting performance of the model calibrated with the two PMP methods. In Section 4 the results of the ex-post experiment are presented and Section 5 contains the discussion and conclusions.

3.2. FSSIM for arable farming

The main purpose of FSSIM within the SEAMLESS framework is to simulate responses of farming systems within the EU25 to policy changes and technological developments and to calculate price-supply relationships at farm level (Janssen *et al*., 2010; Louhichi *et al*., 2009). The price-supply relationships of FSSIM are aggregated to higher levels and used to evaluate market impact of environmental policies and agricultural innovations at EU and global scale. For that reason FSSIM is designed to be generic and flexible, accounting in an easy way for region specific policies or alternative production activities to be used for a sample of representative NUTS2 regions (i.e. nomenclature of territorial units for statistics) across the EU25. FSSIM has been fully integrated in the whole SEAMLESS framework to facilitate the process of exchanging inputs and outputs with other models and databases (Van Ittersum *et al*., 2008).

Modelling all individual farms within the EU25 is not feasible because of the large number of farms and the existing diversification among different farming systems. For that reason, a farm typology was developed within SEAMLESS based on economic, environmental and social characteristics of EU farms, linking farm level data to environmental data (Andersen *et al*., 2007). The SEAMLESS farm typology is based on the existing EU farm typology (Decision 85/377/EEC, 1985) which classifies farms according to their income and specialization. This farm typology has been enriched with environmental criteria related to the land use and intensity of farming. A spatial allocation procedure adds a spatial dimension to the farm types and makes it possible to aggregate farms of the same farm type to both natural (territorial) and administrative regions (Elbersen *et al*., 2006). The total available land of each farm type is spread across a number of agro-environmental zones which are defined as combinations of climatic zones and soil types. The number of agro-environmental zones of a farm type depends on the diversity of climate and soil types of the region but also on the degree of dependence of a farm type to specific climatic and soil conditions. The ‘average farm’ is used to represent all farms that belong to the same farm type. This average farm is a virtual construction, derived by averaging historical data from farms that are grouped in the same farm type.

Farm **S**ystem **SIM**ulator for arable farming is a static, one year LP model which maximizes the total gross margin of an average farm of a certain farm type subject to a set of resource and policy constraints.

$$\max_{x \in [o..+\infty)} \{z = r'x - c'x\}, \quad s.t. \quad Ax \leq b\ [\pi], \quad x \geq 0, \tag{1}$$

where z is the objective value (e.g. total gross margin) of a certain farm type; x is an $n\times1$ vector of production activities; r is the $n\times1$ vector of activity revenues; c is the $n\times1$ vector of variable costs; A is the $m\times n$ matrix of the technical coefficients; b is the $m\times1$ vector of upper bounds of the resources, and policy constraints; and π is the $m\times1$ vector of shadow prices of the resource and policy constraints.

The total gross margin is defined as total revenues from crop production and subsidies minus variable costs including costs of agrochemicals, fertilizers, irrigation and hired labour. Costs related to machinery and buildings are not taken into account as they are assumed to be fixed within the time horizon of the model. The total gross margin is maximized subject to a number of basic resource and policy constraints relevant to all EU arable farms:

- The available land constraint restricts the simulated area to the available farm area (per soil type);
- The labour availability constraint determines the required hired labour on top of family labour;
- The irrigated land constraint restricts the area under irrigated activities to the available irrigated land;
- The obligatory set-aside constraint sets a lower bound to the area of fallow land.

In FSSIM for arable farming, production activities are specified as crop rotations and not as single crops. Consequently, although rotational constraints are not included in the model explicitly, the various agronomic rules and restrictions are taken into account during the construction of the production activities (Dogliotti *et al.*, 2005; Janssen *et al.*, 2009) and thus outside the optimization model. It is assumed that the areas of all crops that are part of the same rotation are equal. For example, it is assumed that in a four years rotation all different crops are grown on 25% of the area of the rotation. To make the concept of a rotation compatible to a static one period model like FSSIM we also assume that in each period all crops of the rotation are grown in the field. The technical coefficients of a particular crop can differ between rotations accounting for possible interactions between crops. A model solution can include several crop rotations simultaneously within one farm. Compared with other PMP applications in the literature, this reduces the burden on the calibration methodology for correctly representing the substitution between the single crops.

3.3. Methodology

The STPMP approach is briefly described here, some limitations are highlighted and an alternative PMP variant is presented addressing some limitations of the STPMP approach. Finally, an ex-post experiment is designed to compare the forecasting performance of two PMP variants.

3.3.1. The standard PMP approach

Positive mathematical programming approaches assume decreasing marginal gross margins of the beneficial activities, such that in the base year the model exactly reproduces the observed activity levels. To assume decreasing marginal gross margins, a non-linear cost or production function is estimated based on the activity levels of the base year.

The STPMP approach, described in Heckelei (2003), is a two-step approach. Step 1 is the extension of the model described in equation (1) by adding a set of calibration constraints which fix the simulated crop levels to the observed base year data. A small perturbation ε is allowed in order to guarantee that all binding resource constraints of the model described in (1) (the uncalibrated LP model) remain binding in the model described in equation (2):

$$\max_{x \in [o..+\infty)} \{z = r'x - c'x\}, \quad s.t. \quad Ax \leq b\ [\pi], \quad x \leq x^0 + \varepsilon\ [\lambda], \quad x \geq 0, \tag{2}$$

where x^0 is the $n \times 1$ vector of observed activity levels, ε is an $n \times 1$ vector of small positive numbers, and λ is the $n \times 1$ vector of the dual values of the calibration constraints.

In the solution of the model in equation (2), the preferable (high average gross margin) activities are bounded by the calibration constraints, while the non-preferable activities (with low average gross margin) are bounded by the resource and policy constraints (e.g., obligatory set-aside). The calibration constraints of the non-preferable activities are not binding and consequently their shadow prices are equal to 0.

In step 2 of STPMP the calibration constraints of model described in equation (2) are taken out although their shadow prices are used to estimate the parameters of a quadratic cost function [equation (3)] such that the model exactly calibrates to the base year data. Different functional forms with the required properties (i.e., positive semi-definite functions) can be used. For simplification purposes and because there are no strong arguments for using a different functional form, a quadratic cost function has been used in most PMP related studies (Heckelei, 2003). This functional form is also selected here.

$$C = d'x + 0.5x'Qx, \tag{3}$$

where, d is the $n \times 1$ vector of parameters associated with the linear term and Q is the symmetric ($n \times n$) positive semi-definite matrix of parameters associates with the quadratic terms. The general structure of the calibrated model is:

$$\max_{x \in [o..+\infty)} \{z = r'x - d'x - 0.5x'Qx\}, \quad s.t. \quad Ax \leq b\ [\pi], \quad x \geq 0. \tag{4}$$

To guarantee exact calibration, the parameters of the cost function must be estimated to satisfy the first order conditions of the quadratic optimization model:

$$c + \lambda = d + Qx^0. \tag{5}$$

Assuming that $d=c$, and that Q is a diagonal matrix with diagonal elements specified such that[2]: $q=\lambda./ x^0$ (q is the $n\times 1$ vector of diagonal elements of the Q matrix) we can estimate a set of parameters of the quadratic cost function that will ensure exact calibration at the base year.

Despite the fact that the standard PMP approach guarantees exact calibration to the base year data, it has some limitations which affect the predictive capacity of the calibrated model. The first limitation is related to the arbitrariness that dominates the estimation of the parameters of the non-linear cost function. The arbitrary assumptions of the STPMP approach that $d=c$ and $q=\lambda./x^0$, imply that the non-linear term (q) of the non-preferable activity will be equal to 0, since the shadow price λ of the calibration constraint of this activity is equal to 0. This means that the marginal gross margin of the non preferable activity is constant. On the contrary, the marginal gross margins of all other activities decrease and depend on the activity levels. As decreasing marginal gross margin applies also to the non-preferable activity, it is theoretically more appropriate to assume decreasing marginal gross margin for this activity too. A simple example illustrating the problems following from this implicit assumption of STPMP is the case where one additional unit of a scarce resource becomes available (e.g., one ha of land). The model calibrated with STPMP will allocate this resource in a way that the level of the preferable activities remains constant. The additional land will be allocated to the non-preferable activity only.

A second limitation of STPMP, which has implications for the forecasting performance of the model, is related to the implicit under estimation of the value of limiting resources. For example, in the specific case where the available land is the only limiting resource, the STPMP approach equalises the value of land at the observed activity levels to the gross margin of the non preferable activity (e.g., set-aside). This could be derived from the first order conditions of the calibrated model. However, the shadow price of land in the model setup considered should capture the average return of the production plan to fixed factors and management. Farmers decide for the optimal rotation based on a number of factors such as available resources, relative returns and restrictions on land use, rotational constraints, and there are also non-linearities involved in the decision making process. If the optimal rotation is presumed to be reflected in the observed activity levels, then a

---

[2] With './' it is meant element wise division. Each element of vector λ is divided by the respective element of vector x^0.

marginal change in available land is not probably to affect the shares of different activities in the observed rotation. The area of all crops will change accordingly so that the optimal farm plan is maintained. As a result, it is more realistic to assume that the value of land at the observed activity levels is equal to the observed average gross margin.

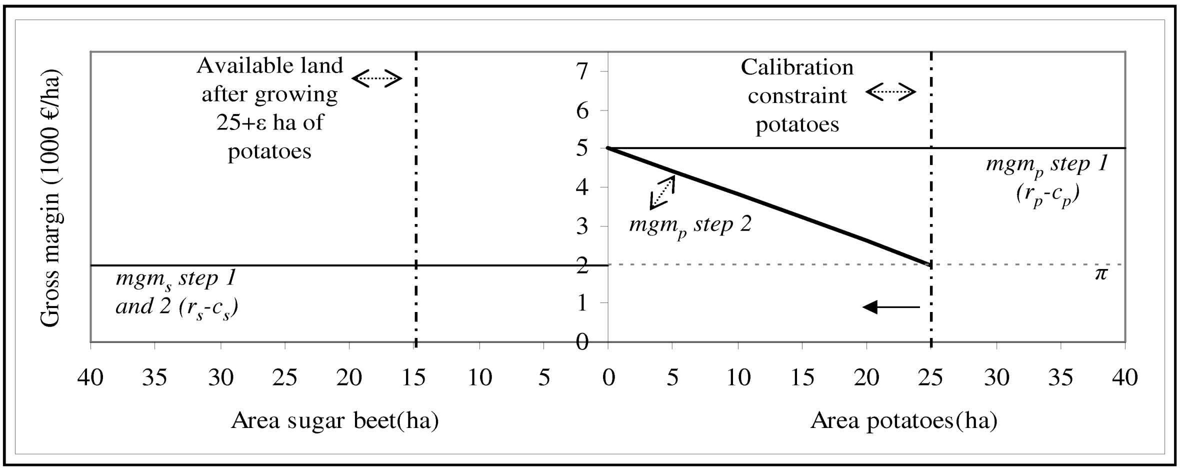

Figure 1: Graphical representation of standard PMP calibration of the two activities example. Marginal gross margins of potatoes (mgm_p) and sugar beet (mgm_s) in step 1 and step 2. See text for further explanation.

Like in Howitt (1995a), a two activities example could be used to present graphically (Figure 1) the two steps of the STPMP approach and reveal the limitations described above. We assume that at a certain moment in time, for which we calibrate the model, a farm grows 25 ha of potatoes and 15 ha of sugar beet. For this simple example, rotation constraints are not taken into account. Growing potatoes is the most profitable activity with average gross margin of 5,000 €/ha whereas sugar beet is less profitable with 2,000 €/ha. It is assumed that the available land is the only limiting resource and constraint of the model. The LP solution would be 40 ha of potatoes which is far from what is observed in reality. Calibrating the model with STPMP will involve two steps. In the first step, two calibration constraints (one for each activity) are added to the LP model to enforce exact calibration. The area of potatoes (most profitable activity) is restricted by the calibration constraint, whereas sugar beet is selected for the remaining land. The calibration constraint of sugar beet is not binding and consequently the shadow price of this constraint is equal to 0. The value of land (π) is equal to the average gross margin of sugar beet (2,000 €/ha). In the second step of STPMP, the shadow prices of land are used to retrieve the quadratic part of the objective function that ensures exact calibration. Given the assumptions of STPMP in step 2, the marginal gross margin of sugar beet is equal to 0 because the shadow price of the calibration constraint is equal to 0 and consequently the quadratic term of sugar beet is 0. Potatoes have a decreasing marginal gross margin with slope ($q_p = \lambda_p / x_p^0 = 3{,}000 / 25 = 120$).

3.3.2. An extended variant of PMP

An extended variant of the standard PMP approach (EXPMP) is presented here. The goal of this method is to overcome the shortcomings of the STPMP approach outlined above and thus to improve the predictive performance of the model.

Like the STPMP approach, the EXPMP variant is a two step approach. In the first step of EXPMP, the value of land is raised to the weighted average gross margin (calculated at the base year situation) for reasons explained above. To achieve this, we include a land renting activity, in which additional land is available at the farm's average gross margin for each hectare of used land. Consequently, the added activity is not really a land rent because it includes remuneration for capital, management and labour assets. This is incorporated in the model by adding the costs of rented land to the objective function of the model described in equation (2) and by replacing the resource constraint of the available land with a flexibility constraint where land is a decision variable [equation (6)]. The shadow price of the flexibility constraint and consequently the perceived value of the land are equal to the average gross margin at the base year. The set of activities is separated in two groups: (i) those activities that result in gross margins higher than the average gross margin at the observed activity level and (ii) those activities that result in gross margins lower than the average gross margin at the observed activity level. The first group of activities is restricted to their observed activity level by the set of calibration constraints of the STPMP approach. This set fixes an upper bound, equal to the observed activity levels, to each of the activities. The levels of the activities that belong to the second group are not restricted by those constraints because they have a gross margin lower than the average gross margin. To ensure exact calibration, a second set of calibration constraints is added to the model. Those constraints set a lower bound to each activity and restrict the area of activities with gross margin lower than the average. This lower bound is equal to the observed activity level plus a small positive number, so that we finally obtain:

$$\max_{x,y\in[o..+\infty)} \{z = r'x - c'x - g\cdot y\}, \quad s.t. \quad A^{-}x \leq b^{-} \quad \lfloor \pi^{-} \rfloor, \; I'x - y \leq 0,$$
$$x \leq x^0 + \varepsilon \;\; [\lambda], \quad x \geq x^0 - \varepsilon \;\; [\lambda'], \quad x \geq 0, \tag{6}$$

where *g* is the average gross margin at the observed level, *y* is the rented land (a variable equal to the total used land when the average gross margin is positive), A^{-} is the $(m\text{-}1)\times n$ matrix of the technical coefficients of resource and policy constraints except from the available land constraint, b^{-} is the (m-1)×1 vector of upper bounds to the model's constraints, π^{-} is the $(m\text{-}1)\times 1$ vector of shadow prices of the resource and policy constraints except from the available land constraint, *I* is a $n\times 1$ vector of ones and λ' is the $n\times 1$ vector of shadow prices of the second set of calibration constraints. For each activity, only one of the two calibration constraints is binding. Consequently, either λ or λ' will be non-zero. To guarantee exact calibration in step 2, the parameters of equation (3) need to be specified such that:

$$c + \lambda + \lambda' = d + Qx^0. \tag{7}$$

As in the STPMP approach, the *Q* matrix is assumed to be diagonal. All diagonal elements of the *Q* matrix should be positive to ensure positive semi-definiteness and consequently satisfaction of the second order conditions of the calibrated model. Contrary to the STPMP approach, where the intercepts of the quadratic cost functions are equal to the respective constant average costs of the LP model, in EXPMP the intercepts of the quadratic cost functions differ from average costs in the LP model. Calculating parameters *Q* and *d* of the quadratic cost function, as in equations (8) and (9) will satisfy equation (7) for any value of α.

$$q = \alpha|\lambda + \lambda'|./x^0 \tag{8}$$

$$d = c + (\lambda + \lambda') - \alpha|\lambda + \lambda'| \tag{9}$$

where, α is an $n\times 1$ vector of parameters that determines the weights of the linear and the non-linear costs of the activities in the objective function. Later, in this article, it is shown that the value of α is related to the own supply elasticity of different activities. The larger the value of α, the less sensitive the model becomes to price changes. A large value of α

can result in a negative intercept of the marginal cost function. In the case of activities with low marginal gross margin (below the average gross margin) and when $c \leq \lambda' - \alpha|\lambda'|$, the marginal costs are negative at the observed activity levels. From the first order conditions of the calibrated model it can be shown that the supply elasticity of the activity levels is reciprocally related to the respective α parameter:

$$\eta_i = \frac{r_i}{\alpha_i \left| \lambda_i + \lambda_i^{'} \right|} \tag{10}$$

where, η_i is the own-price elasticity of supply for activity i. From equation (10) it can be concluded that as the value of α parameter increases, the supply becomes more inelastic. The value of α can differ between regions, farm types and crops. One way to determine the value of α is to use elasticities that have been estimated in existing econometric studies in equation (10) and solve the equation for α. This procedure will result in a different value of α for each different activity. It is important to notice, that elasticities of supply estimated in econometric studies at sector level are not always comparable with farm level elasticities. This is mainly because the former include structural changes and the effects of the industry whereas the latter do not. For this reason, the own-price elasticities which are usually estimated at regional or industry level are not used to fix the farm-price elasticities. They are used only as prior information which is used together with the farm and activity specific shadow prices of the calibration constraints to recover the value of the unknown parameters *(*Helming *et al.*, 2001; Gocht, 2005). Alternatively, in cases where supply elasticities are not available, α can be estimated from ex-post experiments. With this approach the same value of α is assumed for all activities. The model is calibrated and used iteratively with different values of α in each iteration. The value of α that gives the best forecast can then be used for the actual simulations and scenario testing. Using the α parameter in this way to estimate the parameters of the cost function reduces the arbitrariness of the STPMP approach by attaching a better empirical justification to the necessary assumptions.

The graphical example with two activities of the section 3.3.1. could be also used to summarise the differences between STPMP approach and the extended variant of PMP (Figure 2). In the first step of the EXPMP a land renting activity is offered to the model in order to raise the value of land to the average gross margin ($\pi = g$). In the simple example, the average gross margin of the farm at the observed activity levels is 3,875 €/ha. The farmer is confronted with an additional cost of 3,875 €/ha of used land. As a result, the gross margin of potatoes is equal to 5,000-3,875 = 1,125 €/ha whereas the gross margin of sugar beet is equal to 2,000-3,875 = -1,875 €/ha. Two sets of calibration constraints are used (instead of one in STPMP) to enforce exact calibration. The first set of calibration constraints sets an upper bound to the observed areas of the two activities. The area of potatoes is restricted by this constraint, whereas the area of sugar beet is not (the gross margin of sugar beet is below the average). The second set of calibration constraints imposes a lower bound to the level of sugar beet (this constraint is not binding for potatoes). In Figure 2, only the two binding calibration constraints are presented. In step 2 of EXPMP, the relationships of (8) and (9) are used to calculate quadratic terms for both activities. As a result the marginal gross margin of both activities decreases with increasing the area of the activities.

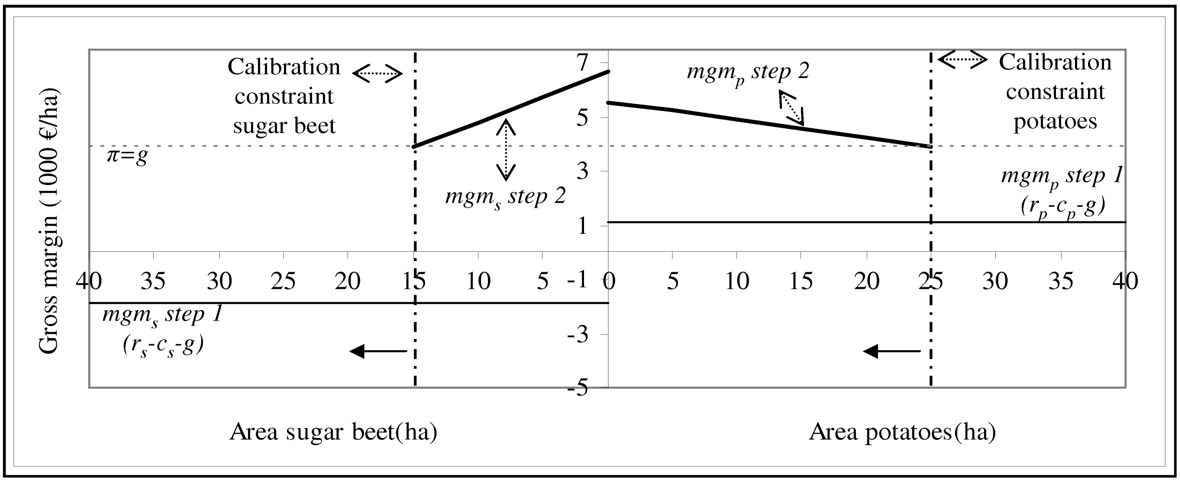

Figure 2: Graphical representation of the extended PMP calibration of the two activities example. Marginal gross margins of potatoes (mgm_p) and sugar beet (mgm_s) in step 1 and step 2. See the text for further explanation.

3.4. Ex-post application to arable farm types of Flevoland and Midi-Pyrenees

The forecasting performance of FSSIM, calibrated with the EXPMP and STPMP variants, was tested with a number of ex-post experiments. For operational purposes and because of data availability and data quality, the farm types of Flevoland (the Netherlands) and Midi-Pyrenees (France) were used. For each farm type identified in the SEAMLESS farm typology (Andersen *et al.*, 2007), an FSSIM model was developed. The FSSIM models were calibrated, with the two PMP variants for the year 1999 and used to predict the changes in the activity levels of year 2003. It was assumed that only prices changed and technology remained constant. The same constraint structure was used for the simulations of both year 1999 and 2003.

Table 1: Observed crop levels (ha), areas per agro-environmental zone* (ha), and farm resources for two arable farm types in Flevoland in 1999 and 2003

	FT3203		FT3303	
Crop levels (ha)	1999	2003	1999	2003
Maize (silage)	1.6	2.0	1.7	0.6
Onion	3.4	3.2	9.1	9.7
Potatos (ware & seed)	18.0	17.9	24.8	24.8
Set-aside	1.5	1.8	1.3	1.3
Soft wheat	8.6	10.4	6.1	11.5
Sugar beet	10.3	11.2	12.2	9.1
Other crops (not simulated)	14.9	19.8	12.2	11.7
Total	58.3	66.3	67.4	68.7
Available land per agro-environmental zone (ha)				
Flevoland agro-env. zone 1	7.8	8.9	6.4	6.5
Flevoland agro-env. zone 2	0.4	0.5	0.3	0.3
Flevoland agro-env. zone 3	47.1	53.5	59.0	60.1
Flevoland agro-env. zone 4	2.4	2.7	0.8	0.8
Flevoland agro-env. zone 5	0.6	0.7	1.0	1.0
Other Farm resources				
Available family labour (hrs)	2,997	2,997	5,403	5,403

Note: * An agro-environmental zone is characterized by a climatic zone, the soil organic carbon (SOC) content and the region

In the SEAMLESS farm typology for Flevoland, we used observations for two arable farm types and the years 1999 and 2003. The first farm type (FT3203)[3] is a large size, medium intensity arable farm, whereas the second one (FT3303) is a large size, high intensity farm. The observed crop levels, the available area per soil type and the available family labour of each farm type are shown in Table 1. Available farm resources and data on observed activity levels were taken from the farm accounting data network (FADN). According to expert knowledge, some of the crops that were observed in FADN data were not important and were not considered typical for the region (Borkowski *et al.*, 2007). Those crops were not taken into account in the simulations and the corresponding land was treated as fixed

---

[3] The first digit of the farm type code refers to the farm size: (3) Large farms (size > 40 ESU), (2) Medium farms (16 ESU ≤ size ≤ 40 ESU), (1) small farms (size < 16 ESU). The second digit refers to farm intensity: (3) High intensity (output > 3,000 €/ha), (2) Medium intensity (500 €/ha ≤ output ≤ 3,000 €/ha), (1) Low intensity (output < 500 €/ha). The two last digits refer to farm specialization: (04) arable/other, (03) arable specialised crops, (02) arable/fallow, (01) arable/cereal.

land. In Table 1 these crops are referred to as ‘other crops not simulated’. In total, five agro-environmental zones were identified in Flevoland which are combinations of two different climate zones and three soil types.

Table 2: Observed crop levels (ha), areas per agro-environmental zone (ha), and farm resources for two arable farm types in Midi-Pyrenees in 1999 and 2003

	FT3201		FT3202	
Crop levels (ha)	1999	2003	1999	2003
Barley	2.7	4.1	1.9	1.6
Maize (grain)	32.1	35.1	20.1	25.1
Maize (silage)	0.1	0.3	0.5	0.5
Rape seed	3.8	1.7	4.3	1.0
Set-aside	8.0	9.3	16.3	18.9
Soya	4.8	3.0	5.2	3.6
Sunflower	14.3	14.3	12.3	12.6
Wheat (durum)	11.5	17.3	6.9	11.4
Wheat (soft)	20.5	13.1	12.1	12.3
Other crops (not simulated)	38.1	43.0	37.3	36.8
Total	135.9	141.2	116.9	123.8
Available land per agro-environmental zone *(ha)*				
Midi-Pyrenees agro-env. zone 1	2.1	2.2	1.6	1.7
Midi-Pyrenees agro-env. zone 2	1.2	1.3	1.5	1.6
Midi-Pyrenees agro-env. zone 3	4.1	4.3	4.8	5.1
Midi-Pyrenees agro-env. zone 4	3.0	3.1	1.0	1.1
Midi-Pyrenees agro-env. zone 5	0.1	0.1	0.0	0.0
Midi-Pyrenees agro-env. zone 6	38.9	40.5	24.3	25.8
Midi-Pyrenees agro-env. zone 7	57.2	59.5	43.9	46.7
Midi-Pyrenees agro-env. zone 8	4.8	5.0	7.0	7.4
Midi-Pyrenees agro-env. zone 9	1.9	2.0	0.8	0.8
Midi-Pyrenees agro-env. zone 10	0.4	0.4	0.1	0.1
Midi-Pyrenees agro-env. zone 11	0.3	0.3	0.2	0.2
Midi-Pyrenees agro-env. zone 12	3.1	3.2	6.7	7.1
Midi-Pyrenees agro-env. zone 13	15.1	15.7	21.5	22.9
Midi-Pyrenees agro-env. zone 14	2.2	2.3	1.9	2.0
Midi-Pyrenees agro-env. zone 15	1.1	1.1	1.0	1.1
Midi-Pyrenees agro-env. zone 16	0.1	0.1	0.0	0.0
Midi-Pyrenees agro-env. zone 17	0.1	0.1	0.2	0.2
Other Farm resources				
Available family labour (hrs)	2,901	2,901	3,260	3,260

In Midi-Pyrenees, we also used observations for two arable farm types of 1999 and 2003, where both types are large farms of medium intensity and differ only in specialisation. The first is a cereal farm (FT3201), and the second is an arable-fallow farm (FT3202). The observed crop levels, the available area per agro-environmental zone and the available family labour of each farm type are presented in Table 2. In total, 17 agro-environmental zones were identified in Midi-Pyrenees which are combinations of three different climate zones and six soil types.

The most common rotations and respective managements that are currently used in Flevoland and Midi-Pyrenees have been identified in a survey conducted within SEAMLESS (Borkowski *et al*., 2007). Expert knowledge was used to quantify the input-output coefficients (e.g., yields, costs, externalities) of these activities for year 2003. The same coefficients were used for the base year (1999). The short-term horizon justifies the assumption that the input and output coefficients of activities do not change. EUROSTAT and national databases were used to determine crop product prices for 1999 and 2003. Average crop product prices of years 1996, 1997 and 1998 were used for the base year simulations, whereas average prices from years 2000, 2001 and 2002 were used for the year 2003. EUROSTAT data were also used for estimating the received subsidies in years 1999 and 2003. The average prices and the received subsidies per crop of Flevoland and Midi-Pyrenees are presented in Table 3 and Table 4, respectively. Three different simulations were performed:

1. The model was calibrated with the STPMP approach and used to forecast 2003.
2. The model was calibrated with EXPMP with different values of α for each crop (EXPMP α = dif) and used to forecast 2003. The value of α of each crop was estimated based on supply elasticities from existing literature (Jansson, 2007).
3. The model was calibrated with EXPMP with the same value of α for all activities and used to forecast 2003. The value of α was determined in an iterative process. In each iteration, the FSSIM model was calibrated with the EXPMP approach with a different value of α and the simulation results were compared with the 2003 observed crop levels. The percentage absolute deviation (PAD)[4] was used as

---

[4] The percentage absolute deviation (PAD) is defined as the absolute deviation between simulated and observed activity levels per unit of actual activity level:

$$PAD(\%) = 100 \cdot \frac{\left(\sum_i \left| x_i - x_i^0 \right|\right)}{\left(\sum_i x_i^0\right)}$$

measurement of the model's performance. The value of α which obtains the minimum PAD value was selected.

The PAD value obtained with the STPMP approach for each farm type was compared with the PAD value obtained with the EXPMP variant using simulated supply elasticities, and with the PAD value obtained with the EXPMP variant using the iterative process described above under the third simulation.

Table 3: Crop product prices (€/ton), subsidies (€/ha), gross margins (€/ha) and supply elasticities (η) for 1999 and relative changes in 2003 in Flevoland

	Price		Subsidies		Gross margin		η
Crop	1999	Change(%)	1999	Change (%)	1999	Change (%)	1999
Maize (silage)	22	2	336	9	135	51	0.1
Onion	150	-40	-	-	6602	-53	0.5
Potatoes (ware)	129	-40	-	-	4975	-57	0.4
Potatoes (seed)	247	0	-	-	2237	0	0.4
Set-aside	0	0	408	7	307	10	0.1
Soft wheat	121	-6	334	46	817	12	0.9
Sugar beet	54	-6	-	-	1892	-9	1.0

Table 4: Crop product prices (€/ton), subsidies (€/ha), gross margins (€/ha) and supply elasticities (η) for 1999 and relative changes in 2003 in Midi-Pyrenees

	Price		Subsidies		Gross margin		η
Crop	1999	Change (%)	1999	Change (%)	1999	Change (%)	1999
Barley	117	-11	315	15	570	-29	1.9
Maize (grain)	131	-9	316	14	902	-8	1.5
Maize (silage)	131	-9	310	16	896	-8	3.8
Rape seed	222	-2	537	-35	925	-21	0.8
Set-aside	0	0	440	-23	340	-29	0.1
Soya	199	-8	538	-35	810	-29	0.4
Sunflower	224	11	537	-35	781	-16	0.1
Wheat (durum)	177	-14	540	8	806	-8	1.3
Wheat (soft)	128	-12	309	17	603	-6	0.9

3.5. Results

The PAD values of the simulations of FSSIM calibrated with EXPMP, where the value of α is estimated in an iterative process (third simulation), for the farm types of Flevoland and Midi-Pyrenees are presented in Figure 3. It appears that the model forecasts improve as α increases and consequently as the model becomes less responsive to price changes. At some point, the PAD value is minimal and then it starts increasing again slowly. The minimum PAD values of the simulations for farm type FT3203 and FT3303 of Flevoland were achieved for $\alpha = 10.8$ and $\alpha = 11.8$ respectively, whereas the minimum PAD values of the simulations of farm types FT3201 and FT3202 of Midi-Pyrenees were achieved for $\alpha = 5.5$ and $\alpha = 3.4$ respectively.

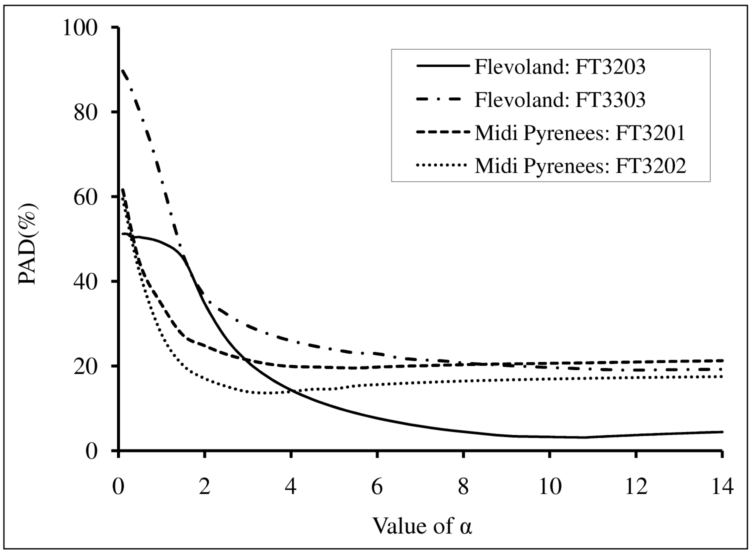

Figure 3: The percentage absolute deviation (%) for different values of α for the farm types

The results of the ex-post experiment of farm type FT3203 and FT3303 of Flevoland are presented in Table 5. The observed cropping patterns in 1999 (x^0_{1999}) and 2003 (x^0_{2003}) indicate a 7.2% (3.1 ha) increase of the total available farm land of farm type FT3203. This additional farm land was covered mainly by soft wheat and sugar beet. A 3.3% (1.8 ha) increase of available farm land is observed in the cropping pattern of farm type FT3303 which was covered mainly by soft wheat. The changes in areas of individual crops from year 1999 to 2003 are different between the two farm types. In farm type FT3203, the area of onions decreases slightly whereas the areas of maize, soft wheat and sugar beet increases. In farm type FT3303 the areas of maize for silage and sugar beet decrease, whereas the areas of onions and soft wheat increase substantially. The areas of potatoes remained the same in both farm types despite the large price decrease.

For both farm types of Flevoland, the STPMP simulation resulted in the highest PAD values in 2003, indicating a low forecasting capacity. The main reason for this is the large simulated set-aside area. The results of the model calibrated with STPMP show a reduction of the simulated areas of onion and potatoes because of the substantial price decrease of these crops. The reduction together with the additional available farm land was allocated mainly to set-aside, the gross margin of which increased with almost 10% compared with the base year. In STPMP, it is assumed that the marginal gross margin of set-aside, which is the non-preferable activity, is constant and independent from the simulated area of set-aside. On the contrary, the marginal gross margin of preferable activities decreases as the simulated level of the activity increases. This is the reason that the reduction of the simulated areas of onions and potatoes and the additional available farm land was all allocated to set-aside. The substantial increase of the simulated area of set-aside was not observed in any of the EXPMP simulations presented in Table 5, which gave forecasts much closer to the observed cropping pattern of year 2003. The main reason for the more realistic simulations of EXPMP is the assumption of decreasing marginal gross margin of set-aside (non-preferable activity) as opposed to the STPMP approach where the marginal gross margin of set-aside was assumed to be constant. The problem of STPMP described above is not observed in Midi-Pyrenees because in this region the gross margin of set-aside reduced substantially in 2003 because of a subsidy decrease. The gross margin decrease of set-aside is larger than the gross margin decrease of other crops. Some of the additional available farm land and the reduction of the simulated area of crops with low gross margin was first allocated to crops like wheat and maize, the gross margin of which decreased but not as much as that of set-aside. Once the marginal gross margin of these crops falls below the constant gross margin of set-aside, the area of set-aside started to increase and captured the remaining land. Nevertheless, the remaining land was not sufficient to create the same problem as for Flevoland.

Table 3 shows that the base year gross margin of silage maize is lower than that of set-aside which would make silage maize the non-preferable activity. However, maize for silage is part of a rotation with other profitable crops, that is sugar beet, potatoes and soft wheat. This rotation has a relatively high marginal gross margin because of higher yields and lower costs of wheat and sugar beet. As a result, the shadow price of the calibration constraint of silage maize becomes positive and a decreasing marginal gross margin is assumed. Set-aside becomes the non-preferable activity with the lowest and consequently constant marginal gross margin in STPMP. In 2003, because of decrease of marginal gross margin of potato and sugar beet, the rotation which includes silage maize becomes less profitable and the area of maize observed in 1999 is replaced by set-aside.

Using supply elasticities to determine a different value of α for each crop in the EXPMP calibrated model of farm type FT3203 resulted in higher PAD values than the minimum achieved PAD value of the third simulation with the same α for all activities (absolute difference is 19%). For farm type FT3303 the EXPMP calibrated model with different value of α for each crop resulted in PAD values close to the minimum achieved in the third simulation.

Table 5: Observed crop levels (x^0) in 1999 and 2003, and forecasted crop levels (x_i) for 2003 with the standard PMP approach and the EXPMP variant for farm types FT3203 and FT3303 of Flevoland.

			STPMP		EXPMP		EXPMP	
	x^0_{1999}	x^0_{2003}	x_i	$\lvert x_i-x^0_{2003}\rvert$	x_i	$\lvert x_i-x^0_{2003}\rvert$	x_i	$\lvert x_i- x^0_{2003}\rvert$
Results for FT3203								
Value of α					α=dif*		α=10.8	
Crop area								
Maize (silage)	1.6	2	0.0	2.0	1.5	0.5	1.7	0.3
Onion	3.4	3.2	1.5	1.7	2.7	0.5	3.3	0.1
Potatoes (ware+seed)	18	17.9	8.1	9.8	15	2.6	17.9	0.0
Set-aside	1.5	1.8	12.9	11.1	1.5	0.3	1.6	0.2
Soft wheat	8.6	10.4	14.1	3.7	16	5.2	10.2	0.2
Sugar beet	10.3	11.2	10.1	1.1	9.9	1.3	11.8	0.6
Total area (ha)	43.4	46.5	46.5	29.4	47	10.4	46.5	1.5
PAD (%)				63		22		3
Results for FT3303								
Value of α					α=dif		α=11.8	
Crop area								
Maize (silage)	1.7	0.6	0.0	0.6	1.8	1.2	1.8	1.2
Onion	9.1	9.7	4.0	5.7	8.9	0.8	8.8	0.9
Potatoes (ware+seed)	24.8	24.8	9.7	15.1	25	0.2	24.8	0.0
Set-aside	1.3	1.3	20.5	19.2	1.4	0.1	1.4	0.1
Soft wheat	6.1	11.5	11.3	0.2	6.8	4.7	7.0	4.5
Sugar beet	12.2	9.1	11.5	2.4	13	4.0	13.2	4.1
Total area (ha)	55.2	57.0	57.0	43.3	57.0	11.0	57.0	10.8
PAD (%)				76		19		19

* Parameter α was estimated based on supply elasticities from existing literature (simulation 2)

Table 6 presents the results of the ex-post application of FSSIM to farm types FT3201 and FT3202 of Midi-Pyrenees. The total available farm land of farm type FT3201 increased slightly from 1999 to 2003 (0.3 ha). The areas of rape, soya and soft wheat decreased, whereas the areas of barley, maize and durum wheat increased. The available farm land of farm type 3202 increased, from 1999 to 2003, by almost 9.2% (7.4 ha). The areas of rape seed and soya were replaced by maize and winter durum wheat.

Table 6: Observed crop levels (x^0) in 1999 and 2003, and forecasted crop levels (x_i) for 2003 with the standard PMP approach and the EXPMP variant for farm types FT3201 and FT3202 of Midi-Pyrenees.

			STPMP		EXPMP		EXPMP	
	x^0_{1999}	x^0_{2003}	x_i	$\lvert x_i - x^0_{2003}\rvert$	x_i	$\lvert x_i - x^0_{2003}\rvert$	x_i	$\lvert x_i - x^0_{2003}\rvert$
Results for FT3203								
Value of α					α=dif*		α=5.5	
Crop area								
Barley	2.7	4.1	3.6	0.5	3.1	1.0	3.0	1.1
Maize (silage)	0.1	0.3	0.1	0.2	0.1	0.2	0.1	0.2
Maize (grain)	32.1	35.1	32.1	3.0	31.9	3.2	34.1	1.0
Rape seed	3.8	1.7	2.8	1.1	3.5	1.8	2.3	0.6
Set-aside	8.0	9.3	7.5	1.9	8.0	1.4	8.0	1.3
Soya	4.8	3.0	3.2	0.2	4.5	1.5	3.0	0.0
Sunflower	14.3	14.3	12.8	1.5	14.2	0.1	12.8	1.5
Wheat (durum)	11.5	17.3	11.9	5.4	11.6	5.7	12.9	4.4
Wheat (soft)	20.5	13.1	24.4	11.3	21.2	8.1	22.1	9.0
Total area (ha)	97.9	98.2	98.2	25.1	98.2	22.9	98.2	19.2
PAD (%)				26		23		20
Results for FT3202								
Value of α					α=dif		α=3.4	
Crop area								
Barley	1.9	1.6	2.7	1.1	2.9	1.3	2.4	0.8
Maize (silage)	0.5	0.5	0.5	0.0	0.7	0.2	0.5	0.0
Maize (grain)	20.1	25.1	21.4	3.8	22.7	2.4	24.8	0.3
Rape seed	4.3	1.0	3.5	2.5	4.4	3.4	3.5	2.5
Set-aside	16.3	18.9	21.0	2.1	16.8	2.1	16.7	2.2
Soya	5.2	3.6	3.6	0.0	5.2	1.6	3.7	0.1
Sunflower	12.3	12.6	11.6	1.1	12.4	0.2	12.4	0.2
Wheat (durum)	6.9	11.4	7.4	4.0	8.0	3.4	8.1	3.3
Wheat (soft)	12.1	12.3	15.4	3.1	14.1	1.8	14.8	2.5
Total area (ha)	79.6	87.0	87.0	17.5	87.0	16.4	87.0	11.9
PAD (%)				20		19		14

* Parameter α was estimated based on supply elasticities from existing literature (simulation 2)

Contrary to the results of STPMP in Flevoland, in Midi-Pyrenees, the forecasts of the model calibrated with STPMP are relatively close to the forecasts of the model calibrated with EXPMP. In Midi-Pyrenees, in the STPMP simulation, the area of set-aside was increased marginally compared with the increase of the area of set-aside in the STPMP simulation of Flevoland. The main reason for this is the gross margin decrease of set-aside in 2003 (Table 4). This allowed the areas of crops such as wheat and barley to increase despite the decrease of their gross margins in 2003. This is because the decrease of gross margins of wheat and barley is lower than that of set-aside. The gross margins of other crops like rape seed, soya and sunflower decreased more than the gross margin decrease of set-aside, and hence their simulated land decreases in 2003. The activity substitution in the simulated cropping pattern is more diverse resulting in more realistic predictions with lower PAD values than that observed in Flevoland.

3.6. Discussion & Conclusions

The new PMP variant (EXPMP) presented in this study resulted in lower PAD values than the PAD values achieved by STPMP, in all ex-post exercises. Two major limitations of STPMP (i.e. the underestimation of the value of limiting resources and the assumption of constant marginal gross margin of the non-preferable activity) are overcome and a better justification is attached to the necessary assumptions. As a result, the forecasting capacity of the model improves. The two approaches used in the EXPMP variant to estimate the value of α resulted in similar quality of predictions in both Flevoland and Midi-Pyrenees. Using additional information on supply elasticities to estimate a different value of α for each activity increased the data requirements of the model but also resulted in slightly higher values of PAD compared to the minimum achieved PAD value. Nevertheless, the procedure of determining the value of α is better justified from an empirical point of view. The appropriateness of one of the two approaches depends on data availability. If good quality information on supply elasticities is available, that is, if estimation of supply elasticities is based on longer time series of a dataset relevant for this farm type, then it can be utilized to improve the predictions of the model and to strengthen the economic justification of the assumptions of PMP.

From the ex-post experiments of all farm types calibrated with EXPMP, it can be concluded that given the same values of the model parameters, the model predictions improve as α increases. As α is reciprocally related to the supply elasticities, it can be stated that for this exercise, more inelastic models result in better model predictions. Machinery and managerial capacity of farms do not change that quickly in the short run and for that reason less elastic models are needed. In cases of long term model applications and forecasts, a more elastic response might be more relevant. However, in such cases, factors exogenous to the model, such as changes in the structure of farming systems and the industry, might be more important for good predictions than the elasticity of farm's supply to price changes. The models presented here are calibrated with PMP and consequently exact calibration is guaranteed. We can only assess the performance of the model based on its forecasting capacity. Hazell and Norton (1986) suggest that, in practice, a model that reproduces the base (calibration) year activity levels with PAD values not > 15% can be used for forecasting purposes. It is to be expected that the PAD values of the forecasts of such models will be substantially greater than the PAD values for the calibration year. All farm types tested in this study with the model calibrated by EXPMP resulted in PAD values only marginally > 15% for the forecasting year. We conclude that the forecasting capacity of the resulting model is acceptable.

In this study, the quality of the model predictions is evaluated by comparing observed and simulated cropping patterns. However, assessing other important economic (e.g. average farm income) and environmental (e.g., nitrogen leaching) indicators could be of great interest to model users and policy-makers, because this not only evaluates the modelling methods but also the technical coefficients of the model and hence the quality of the data. FSSIM is used to simulate different farm types across the EU and calculate a number of different indicators relevant for the assessment of a large variety of policy questions. In some cases, the simplifications and the mismatch of data are such that large PMP terms are needed to achieve a satisfactory forecast.

The objective of the SEAMLESS models is to simulate farming systems across Europe. To achieve this, given the available resources, a farm typology was developed and the average farms were simulated with FSSIM. Despite the increased detail of the SEAMLESS typology compared with what is available at EU level, still a lot of the existing diversity between individual farms is not taken into account. In general, the observed cropping pattern of average farms includes more activities than the observed cropping patterns of individual farms. Issues related to farm specific constraints, accessibility of resources and the decision making of individuals are averaged and hence only partially considered. This affects the values of the calibrated parameters of all PMP variants and the results of the analysis. Researchers should be careful with the interpretation of the PMP calibrated parameters since they capture modelling misspecifications.

The suitability of a PMP variant for specific bio-economic analysis depends on various issues, such as justification of PMP assumptions, model characteristics, data availability, type of policy and strategy questions addressed by the model. Gocht (2005), for example, evaluates a number of existing PMP variants with ex-post experiments in Germany, whereas Blanco *et al*. (2008) design ex-post experiments to test models calibrated with different variants of PMP including activities not observed in the base year. Ex-post experiments and validation of the model predictions are clearly necessary to determine the

PMP variant that is more appropriate for each specific case and to increase user's confidence in the model results. From the results of the ex-post exercises presented here, it appears that the EXPMP variant outperformed the STPMP, indicating that EXPMP is an attractive calibration procedure for a bio-economic farm model such as FSSIM.

3.7. References

Acs, S., Berentsen, P., and Huirne, R. (2007). Conversion to organic arable farming in the Netherlands: A dynamic linear programming analysis. Agricultural Systems, 94: 405-415.

Andersen, E., Elbersen, B., Godeschalk, F., and Verhoog, D. (2007). Farm management indicators and farm typologies as a basis for assessment in a changing policy environment. Journal of Environmental Management, 82: 352-362.

Barbier, B. and Bergeron, G. (1999). Impact of policy intervention on land management in Honduras: result of a bioeconomic model. Agricultural Systems, 60: 1-16.

Berentsen, P., Giesen, G and Renkema, J. (1997). Economic and environmental consequences of technical and institutional change in Dutch dairy farming. Netherlands Journal of Agricultural science, 45: 361-379.

Blanco, M., Cortignani, R., and Severini, S. (2008). Evaluating changes in cropping patterns due to the 2003 CAP reform. An ex-post analysis of different PMP approaches considering new activities. 107th Seminar, January 30-February 1, 2008, Sevilla, Spain, European Association of Agricultural Economists.

Borkowski, N., Hecker, J.-M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G., Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy Vol. 1: 33-34.

Decision 85/377/EEC (1985). Commission decision of the 7 June 1985 establishing a Community typology for agricultural holdings (OJ L 220, 17.8.1985, p. 1).

Dogliotti, S., Van Ittersum M.K. and Rossing, W.A.H. (2005). A method for exploring sustainable development options at farm scale: A case study for vegetable farms in South Uruguay. Agricultural Systems, 86: 29-51.

Elbersen, B., Kempen, M., van Diepen, K., Hazeu, G., Verhoog, D. and Andersen E. (2006). Protocols for spatial allocation of farm types PD 4.7.1', SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

Gocht, A. (2005). Assessment of simulation behaviour of different mathematical programming approaches', in Proceedings of the 98th EAAE symposium: State of the art and new challenges, Parma, Italy, pp.166-187.

Gohin, A., and Chantreuil, F. (1999). La programmation mathématique positive dans les modèles d'exploitation agricole: principes et importance du calibrage. Cahiers d'Economie et Sociologie Rurales, 52: 59-78.

Hazell P., and Norton R. (1986). Mathematical Programming for economic analysis in agriculture. Macmillan, New York, 1986.

Heckelei, T. (2002). Calibration and Estimation of Programming Models for Agricultural Supply Analysis'. Habilitation Thesis, University of Bonn, Germany (http://www.ilr1.uni- bonn.de/agpo/ staff/heckelei/heckelei_hab.pdf).

Heckelei, T. (2003). Positive Mathematical Programming: Review of the standard approach'. CAPRI working paper (http://www.ilr1.uni-bonn.de/agpo/rsrch/capri/ wrkpap_e.htm).

Heckelei, T. and Britz, W. (2005). Models based on Positive Mathematical Programming: State of the art and further extensions', in Proceedings of the 98th EAAE symposium: State of the art and new challenges, Parma, Ital.

Heckelei, T. and Wolf, H. (2003). Estimation of constraint optimization models for agricultural supply analysis based on generalised maximum entropy. European Review of Agricultural Economics, 30: 27-50.

Helming, J.F.M., Peeters, L. and Veendendaal P.J.J. (2001) Assessing the consequences of environmental policy scenarios in Flemish agriculture. In: Heckelei, T., Witzke, H.P., and Henrichsmeyer, W. (Ebs): Agricultural sector Modelling and policy information systems. Proceedings of the 65th EAAE seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel, pp. 237-245.

Howitt, R. E. (1995a). Positive Mathematical Programming. American Journal of Agricultural Economics, 77: 329-342.

Howitt, R. E. (1995b). A calibration method for agricultural economics production models. Journal of Agricultural Economics, 46: 147-159.

Janssen, S. and van Ittersum, M. (2007). Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems, 94: 622-636.

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Meuter, E., Andersen, E., Belhouchette, H., Blance, M., Borkowski, N., Heckelei, T., Hecker, M., Li, H., Oude Lansink, A., Stokstad, G., Thorne, P., Van Keulen, H., Van Ittersum, M.K. (2009). A generic bio-economic farm model for environmental and economic assessment of agricultural systems. Submitted for publication in Environmental Management.

Jansson, T. (2007). Econometric specification of constrained optimization models', (PhD dissertation, Bonn University, 2007), electronic publication (http://www.ilr1.uni-bonn.de/agpo/rsrch/capri/ wrkpape .htm).

Louhichi, K., Flichman, G., and Blanco, M. (2009). A generic template for FSSIM for all farming systems. PD3.3.11, SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

Röhm, O. and Dabbert, S. (2003). Integrating agri-environmental programs into regional production models: an extension of Positive Mathematical Programming. American Journal of Agricultural Economics, 85: 254-265.

Van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Olsson, J. A., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A. E., van der Wal, T., Wien, J. E., and Wolf, J. (2008). Integrated Assessment of agricultural systems: a component-based framework for the European Union (SEAMLESS). Agricultural Systems, 96: 150-165.

Chapter 4

4. Estimating Risk Attitude and Production Structure in Ill-posed Bio-economic Farm Models using Maximum Entropy

Argyris Kanellopoulos[1,2], Paul Berentsen[1], Alfons Oude Lansink[1] and Martin van Ittersum[2]

[1] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*

[2] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*

European Review of Agricultural Economics. Submitted.

Abstract

Bio-economic farm models used for higher level policy analysis usually deal with ill-posed problems, calibrated using Positive Mathematical Programming. PMP-based calibration methods do not use available panel data to their full potential and they require strong arbitrary assumptions. In this paper, Maximum Entropy was used to estimate the risk attitude of farmers and the production parameters of a bio-economic farm model. The application focuses on panel data of arable farm types in Flevoland and Midi-Pyrenees. The ME method resulted in better forecasts than PMP. Complementarity and substitution between activities was quantified while the farmer's attitude towards risk was assessed.

Keywords: maximum entropy; bio-economic modelling; integrated assessment; arable farming.

4.1. Introduction

Bio-economic farm models that are proposed for ex-ante integrated assessment of policies across the European Union (EU) usually suffer from a lack of data on agro-management and activity specific practices. The number of observations per farm type on activity levels, production and output is usually not enough to allow for traditional econometric estimation, i.e. the problem is ill-posed (Oude Lansink *et al*., 2001).

In order to address the ill-posed problem, researchers have frequently employed Positive Mathematical Programming (PMP) (Howitt, 1995) in bio-economic studies (Helming *et al*., 2001; Röhm and Dabbert, 2003; Buysse *et al*., 2007). PMP is popular because it guarantees an exact calibration based on just a single observation of activity levels. Nevertheless, a number of limitations can be identified. A first important limitation of PMP is that a number of arbitrary assumptions are imposed on the production structure. A commonly made assumption is that the gross margin of each activity is independent from the simulated level of other activities. This means that complementarity or substitution between different activities is assumed to be absent. This assumption is realistic only in the unlikely case where there is no competition and/or synergy for resources and management between activities. Another assumption of PMP based calibration (Heckelei, 2002) is that at the observed activity levels, the value of the land is assumed to be constant and equal to the gross margin of the least profitable activity. Farmers decide for the optimal farm plan based on a number of factors like available resources, policies, rotational constraints and non-linearities involved in the decision making process. The optimal farm plan is reflected in the observed activity levels. A marginal change of the available farm land will not affect the shares of different activities in the observed rotation. The area of all crops will change proportionally so that the optimal farm plan is maintained. Therefore, it is more realistic to assume that the value of land at the observed activity levels is equal to the observed average gross margin (Kanellopoulos *et al*., 2010).

Another limitation of the PMP approaches is that multiple year observations of activity levels that are available in EU level data bases are not used in the estimation. The use of one year observations of activity levels to recover unknown parameters has also been criticized as it results in poor estimation of parameters reflecting the behaviour of producers (Heckelei, 2002). Observed variation of income because of periodical price and yield changes is not taken into account and consequently in many cases the risk attitude of farmers is ignored and is not taken into account explicitly (Helming *et al.*, 2001; Júdez *et al.*, 2001; Röhm and Dabbert, 2003; Buysse *et al.*, 2007). Those important limitations of PMP have consequences for the model's forecasting capacity and the interpretation of the model's parameters.

Maximum Entropy (ME) is an estimation procedure (Golan *et al.*, 1996) that can be used to estimate problems where the amount of available information is not enough to estimate all unknown parameters (i.e. ill-posed problems). Paris and Howitt (1998) demonstrated the applicability of ME in bio-economic modelling of ill-posed problems while Oude Lansink (1999[a]) used ME to estimate farm-specific output-supply and input-demand relationships to capture technological heterogeneity between farms. Heckelei and Wolff (2003) used ME to estimate bio-economic farm models based on the optimality conditions of a sector gross margin maximization problem. Tonini and Jongeneel, (2008) used ME to estimate supply responses including adjustment dynamics for dairy farming in eastern and central European countries where the quality and quantity of official statistics is not sufficient for traditional econometric estimation. The advantage of ME estimation compared to PMP calibration procedures is that available information in EU level data bases can be utilised more efficiently while a number of calibration restrictions of the PMP approach are relaxed. Moreover, the ME method allows for integrating expert knowledge on some of the model's parameters.

The objective of this paper is to use an ME estimation procedure to estimate the production structure and farmers' risk attitude in ill-posed, bio-economic farm models. Available information in EU level databases and expert knowledge-intuitions for some of the model's parameters are used to relax a number of strong and to some extent arbitrary assumptions that are often made in PMP models. The use of multiple years of observations presumably allows for a better reflection of the farmers response to price and subsidy changes. Complementarities and substitutions between activities are allowed in the model specification. Income variation is explicitly incorporated in the model and the risk attitude of the farmer is estimated along with the production parameters. The empirical application focuses on panel data of arable farm types in Flevoland (the Netherlands) and the Midi-Pyrennees (France). The farm types are simulated with the Farm System SIMulator (FSSIM) (Janssen *et al*, 2009; Louhichi *et al*., 2010), which is the bio-economic model developed within the modelling framework of the System for Environmental and Agricultural Modelling: Linking European Science and Society (SEAMLESS-IF) (Van Ittersum *et al*, 2008).

In section 2, the FSSIM model for arable farming is briefly presented. In section 3, the ME estimation procedure and the setup of the ex-post experiment are described. In section 4, the results of the ex-post exercise are presented. Section 5, discusses and concludes.

4.2. FSSIM for Arable Farm Types

FSSIM is an optimization model which maximizes a farm’s utility subject to a set of resource and policy constraints (model 1). The mean-standard deviation approach (Hazell and Norton, 1986) is used to account for the risk attitude of the farmer.

$$\max \{U = z - \varphi \cdot \sigma_z\}, \quad s.t. \quad Ax \le b\ [\pi], \quad x \ge 0 \qquad (1)$$

Where U is the farmer’s utility defined as total gross margin (z) minus risk, φ is the risk aversion coefficient, and σ_z is the standard deviation of income, x is the $n \times 1$ vector of activities, A is an $m \times n$ matrix of technical coefficients, b is an $m \times 1$ vector of available resources and upper bounds to the policy constraints and π is the $m \times 1$ vector of shadow prices of the resource and policy constraints. Total gross margin is defined as total revenues including sales from agricultural products and compensatory payments (subsidies) minus total variable costs from crop production. A quadratic objective function is used to allow for increasing marginal costs of production that may arise from compensating for inadequate management and machinery capacity (see model 2). A quadratic gross margin function is employed as a functional form because it is flexible and allows for assessing and imposing curvature conditions globally (i.e. positive-definite Hessian):

$$\max_{x \in [0,+\infty)} \{U = r'x - c'x - d'x - 0.5x'Qx - \varphi \cdot \sigma_z\}, \quad s.t. \quad Ax \le b\ [\pi], \quad x \ge 0 \qquad (2)$$

Where r is the $n \times 1$ vector of activity revenues, c is the $n \times 1$ vector of average costs, d is an $n \times 1$ vector of a correction factor to activity’s average gross margin which is estimated, Q is an $n \times n$ matrix of the quadratic part of the cost function which is also estimated. The major sources of variation of income are variation in input prices, output prices and yields. For simplification purposes and because of lack of data in FSSIM, the input prices are treated as constant in the short run. Consequently, the standard deviation of income is a function of revenue variation:

$$\sigma_z = \left(V(z)\right)^{\frac{1}{2}} = \left(V\left(r'x - c'x - d'x - 0.5x'Qx\right)\right)^{\frac{1}{2}} = \left(V\left(r'x\right)\right)^{\frac{1}{2}} = \left(x'\Sigma_r x\right)^{\frac{1}{2}} \qquad (3)$$

Where, Σ_r is the $n \times n$ variance-covariance matrix of the activity's revenues. After substitution in (2) the final form of the model is:

$$\max_{x \in [0,+\infty)} \left\{ z = r'x - c'x - d'x - 0.5x'Qx - \varphi \cdot \left(x'\Sigma_r x\right)^{\frac{1}{2}} \right\}, \quad s.t. \quad Ax \le b \; [\pi], \quad x \ge 0 \qquad (4)$$

The farmer's utility is maximized subject to a number of basic resource and policy constraints relevant to all EU arable farms:

- The available land constraint restricts the simulated area to the available farm area (per soil type).
- The irrigated land constraint restricts the area under irrigated activities to the available land that can be irrigated.
- The labour availability constraint determines the required hired labour on top of family labour.
- The obligatory seta-aside constraint sets a lower bound to the area of fallow land.

Arable agricultural activities are defined as crop rotations grown under specific soil and climate conditions and under a specific management (including soil preparation, sowing, irrigation, fertilization). It is assumed that every year, all crops of a specific rotation are grown on equal shares of the land allocated to this rotation. Consequently, although rotational constraints are not included in the model explicitly, the various agronomic rules and restrictions are taken into account during the construction of the agricultural activities outside the optimization model (Dogliotti *et al.*, 2005, Janssen *et al.*, 2009). A model solution can include several crop rotations simultaneously within one farm.

FSSIM uses information on farm resources and farm economic performance across the EU available in the Farm Accounting Data Network (FADN) and EUROSTAT (Janssen *et al.*, 2009). This data source lacks detail in agro-management information which is needed to assess the environmental aspects of production. Therefore, a simple survey was conducted within SEAMLESS to identify and quantify existing (current) production activities (Borkowski *et al*., 2007; Zander *et al*., 2009). In most cases, the available information is not sufficient to estimate the parameters of (4) using traditional econometric procedures. For that reason a new variant of PMP was used by Kanellopoulos *et al*. (2010) to calibrate the model. Nevertheless, the model still imposes several of the aforementioned restrictions of PMP based calibration approaches: (i) the risk aversion parameter has been set to zero, (ii) only one year of observations has been used and (iii) complementarity and substitution is not taken into account, which means that all off-diagonal elements of Q in (4) are set to zero.

A method based on ME is proposed to relax the restrictions of the PMP calibration procedure of FSSIM. The ME method uses multiple year observations to enable the estimation of the risk aversion coefficient and allows for more flexibility in retrieving possible interactions (i.e. complementarity or substitution) between different activities. Improving the specification of the model must result in improving the robustness and forecasting capacity of the model.

4.3. Methods

4.3.1. ME estimation

The ME estimation procedure described in this paper is based on the approach presented by Paris and Howitt (1998) and it is a two-step approach. In the first step, a linear version of the model is used using the average gross margins of the activities which is data available in the survey conducted within SEAMLESS (Borkowski *et al*., 2007; Zander *et al*., 2009), while setting all unknown parameters (d, Q and φ) equal to zero. Calibration constraints are used to restrict the value of the simulated activity levels to the observed ones. An additional activity similar to a land rent is introduced to raise the value of land to the farm's average total gross margin. The linear model with the calibration constraints and the renting land activity is optimized multiple times one for each year used for estimation. The first step of this method is the same as the first step of the PMP method presented in Kanellopoulos *et al*. (2010). The only difference is that here the model is used multiple times, one for each observed year. Similar to all PMP methods, the objective is to calculate the marginal costs of the activities using information from the shadow prices of the calibration constraints. This is done not only for the base year (like in PMP) but for each year that is used for estimation.

In the second step of the method, the calculated shadow prices of the calibration constraints of step 1 are used to estimate the unknown parameters of the non-linear model in (4) using ME. In ME estimation, the unknown parameters are specified as an additive function of a number of support points and their probabilities. The support points are defined *a priori* by the researcher and can integrate expert knowledge, while the range of values covered by the support points are selected to be wide enough to include the actual value of the parameter. The shadow prices of the calibration constraints calculated in step 1 can be seen as additional information that is used to decrease the uncertainty for the actual value of the parameters and contribute in recovering the actual value of the unknown parameters. Using multiple year observations to estimate the unknown parameters of the model do not allow for exact calibration like in PMP but improves the robustness and forecasting capacity of the model. The two steps of the ME estimation procedure are presented below in more detail.

Step 1: calculating marginal costs of activities at the observed levels

In the first step the marginal costs of each activity in each observed year are calculated. To achieve this, the quadratic objective function of FSSIM is replaced by a linear function using the average costs estimated by experts while risk is not included in the objective function. A land rent activity is used to raise the value of land to the weighted average gross margin. Raising the value of land to the weighted average gross margin was proven to be closer to the actual decision making of the farmer and improves the model's forecasts (Kanellopoulos *et al*., 2010). The farmer will have to pay an amount equal to the farm's average gross margin for each hectare of used land. Consequently, the added activity is not really a land rent because it includes remuneration for capital and labour assets. This is incorporated in the model by adding the costs of rented land to the objective function and by replacing the resource constraint of the available land with a flexibility constraint where the used (rented) land is a decision variable. Next, the set of activities is separated in two groups: (i) those activities with a gross margin higher than the average gross margin at the observed activity level and (ii) those activities with a gross margin lower than the average gross margin at the observed activity level. To calculate marginal costs of all activities at the observed activity levels we used two sets of calibration constraints. The first set of calibration constraints is used to impose an upper bound to the first group of activities which have gross margins higher than the average. The levels of the activities that belong to the second group are not restricted by those constraints because they have a gross margin below the average. To ensure exact calibration, a second set of calibration constraints is added to the model, setting a lower bound to each activity. This lower bound is equal to the observed activity level minus a small positive number, so that we finally obtain:

$$\max_{x,y\in[0,+\infty)} \{z = r'x - c'x - g \cdot y\}, \quad s.t. \quad A^{-}x \le b^{-} \quad [\pi^{-}], \; l'x \le rl, \\ x \le x^{0} + \varepsilon \quad [\lambda], \qquad x \ge x^{0} - \varepsilon \quad [\lambda'], \qquad x \ge 0 \tag{5}$$

where g is the average marginal gross margin at the observed level, y is the rented land (a variable equal to the total used land), A^{-} is the $(m\text{-}1)\times n$ matrix of the technical coefficients of resource and policy constraints except from the available land constraint, b^{-} is the $(m\text{-}1)\times 1$ vector of upper bounds to the model's constraints, π^{-} is the $(m\text{-}1)\times 1$ vector of shadow prices of the resource and policy constraints except from the available land constraint, l is a $n\times 1$ vector of ones, x^{0} is the $n\times 1$ vector of observed activity levels, and λ and λ' are the $n\times 1$ vectors of shadow prices of the first and second set of calibration constraints. For each activity, only one of the two calibration constraints is binding. Consequently, either λ or λ' will be non-zero. Calculation of marginal costs of each activity in all observed years implies multiple model runs (one for each year) with model (5). The marginal costs of each activity for each year are given by:

$$mc_{it} = c_{i} + \lambda_{it} + \lambda'_{it} \quad \forall \; i \in \{1...N\} \; and \; t \in \{1...T\} \tag{6}$$

Where mc_{it} is the marginal costs of (5) of activity i in year t.

Step 2: Estimating the value of unknown parameters with ME

In step 2, a number of support points are defined for each of the unknown parameters (d, Q and φ). The parameters are expressed as a function of probabilities of those support points:

$$d_{i} = \sum_{k} sd_{ik} \cdot pd_{ik} \,, \qquad q_{ij} = \sum_{k} sq_{ijk} \cdot pq_{ijk} \;, \quad \varphi = \sum_{k} s\varphi_{k} \cdot p\varphi_{k} \tag{7}$$

Where d_i is the ith element of the d vector, sd_{ik} is the support point k of the ith element of the d_i parameter, pd_{ik} is the probability of sd_{ik}, q_{ij} is the i,j element of the Q matrix, sq_{ijk} is the kth support point of the q_{ij} parameter, pq_{ijk} is the probability of sq_{ijk}, $s\varphi_k$ is the support point of the risk aversion coefficient and $p\varphi_k$ is the probability of $s\varphi_k$. From the first order conditions of (3) and (4) it can be easily proven that exact calibration would mean that the following relationship is satisfied:

$$mc_{it} = c_i + \lambda_{it} + \lambda'_{it} =$$

$$c_i + d_i + \sum_j q_{ij} \cdot x^0_{it} + \frac{\varphi \cdot \sum_j^{j \le i} \sigma_{ij} x_j}{\sqrt{\sum_{ij} \sigma_{ij} x_i x_j}} + \varepsilon_{it} \quad \forall\ i \in \{1 \ldots N\}\ and\ t \in \{1 \ldots T\} \tag{8}$$

Where σ_{ij} is the i,j^{th} element of the variance covariance matrix of activity's marginal revenues (Σ_r), and ε_{it} is an error term of the marginal costs of activity i at time t. The error term is also written as the sum of the product of probabilities and support points. The centre of the support interval of the error term is set equal to 0 and the width is set to plus and minus a number of standard deviations of the shadow prices of the calibration constraints of (4). Although Golan *et al*. (1996) proposed the three standard deviations rule, Preckel (2001) suggested larger support intervals for the error term so that the ME estimates approximate the Ordinary Least Squares (OLS) estimates. Sensitivity of results to the width of the support range of the error term is analysed to determine the appropriate support range that should be used for estimation.

For $$\gamma_i = \frac{\sum_j^{j \le i} \sigma_{ij} x_j}{\sqrt{\sum_{ij} \sigma_{ij} x_i x_j}} \tag{9}$$

$$\lambda_{it} + \lambda'_{it} = d_i + \sum_j q_{ij} \cdot x^0_{it} + \cdot \gamma_i + \varepsilon_{it} = \sum_k sd_{ik} \cdot pd_{ik} + \sum_{jk} sq_{ijk} \cdot pq_{ijk} \cdot x^0_{it} +$$
$$\sum_k s\varphi_k \cdot p\varphi_k \cdot \gamma_i + \sum_k se_{itk} \cdot pe_{itk} \qquad \forall\ i \in \{1 \ldots N\}\ and\ t \in \{1 \ldots T\} \tag{10}$$

The entropy criterion of such a problem is a function of all probabilities:

$$H = -\sum_{i,k} pd_{ik} \cdot \ln\left(pd_{ik}\right) - \sum_{i,j,k} pq_{ijk} \cdot \ln\left(pq_{ijk}\right) - \sum_{i,t,k} pe_{itk} \cdot \ln\left(pe_{itk}\right) - \sum_{k} p\varphi_k \cdot \ln\left(p\varphi_k\right) \quad (11)$$

Where H is the entropy which is maximized when all probabilities are equal to 1/K (Golan *et al.*, 1996 pp. 21). The maximum entropy estimator could be derived by maximizing H subject to (10). A number of additional constraints are used to ensure that the sum of probabilities used to estimate the unknown parameters is equal to 1.

$$\sum_k pd_{ik} = 1 \ \forall \ i \in \{1...N\} \quad (12)$$

$$\sum_k pq_{ijk} = 1 \ \forall \ i, j \in \{1...N\} \quad (13)$$

$$\sum_k pe_{itk} = 1 \ \forall \ i \in \{1...N\} \ and \ t \in \{1...T\} \quad (14)$$

$$\sum_k p\varphi_k = 1 \ \forall \ i \in \{1...N\} \quad (15)$$

To ensure that Q is a positive semi-definite matrix we used a Cholesky decomposition procedure (Q is rewritten as LDL') according to which each positive semi-definite matrix can be written as a function of an upper triangular matrix ($L_{n\times n}$) with elements L_{ij}, the diagonal elements of which are equal to 1 and a diagonal matrix $D_{n\times n}$ with non-negative diagonal elements D_{ij}.

$$L_{ij} = 1 \ \forall i, j \in \{1..N\} \ and \ i = j \quad (16)$$

$$L_{ij} = 0 \ \forall i, j \in \{1..N\} \ and \ i > j \quad (17)$$

$$D_{ii} > 0 \ \forall i \in \{1..N\} \quad (18)$$

$$D_{ij} = 0 \ \forall i, j \in \{1..N\} \ and \ i \neq j \quad (19)$$

A recursive constraints is added to make sure that the estimated Q matrix will be positive semi-definite.

$$Q_{ij} = \sum_m \left(\sum_l L_{il} \cdot D_{lm} \right) \cdot L'_{mj} \quad \forall i, j, l, m \in \{1..N\} \quad (20)$$

A number of assumptions are made to define the necessary support points for the unknown parameters. There are no general rules for the number of required support points. Nevertheless, there is a trade-off between accuracy and computational requirements. In general, 5 support points for each unknown parameter are sufficient for the type of analysis aimed in this paper (Golan *et al*., 1996 p. 139-140). One central support point is defined for each parameter according to our expectation. The other four support points are defined symmetrically below and above this central support point. It is important that the support interval is wide enough and includes the actual value of the parameters.

Hazell and Norton (1986), suggested that a reasonable range of the risk aversion parameter is between 0 and 1.65. Defining support points within this narrow range will result in estimates very close to 0.825 which makes the comparison between farm types difficult. Using a narrow support range for φ implies high penalties for deviations from the initial expectation of equal probabilities for the different support points. A values of φ close to 0 or 1.65 will never be the outcome. Given that a measurement of variance of such a parameter is not available to help defining a more realistic support range for φ, an iterative procedure that involves multiple model runs is proposed. The objective is to restrict the estimated value of the risk aversion parameter between 0 and 1.65 but at the same time to retrieve risk aversion coefficients that could be used to compare between farms. In the first model run, the support range is set to 0-1.65. The 5 support points are defined as 0, 0.4125, 0.825, 1.2375 and 1.65. This model run will result in a value of risk aversion, close to the central support point (i.e. 0.825). In the next model run, the estimated value of φ is used as the central support point for the risk aversion parameter. The support range is redefined to be symmetrical around the central support point but still within the upper and lower bound of 1.65 and 0 respectively. As a result in each new iteration the support range becomes narrower. For example, if the resulted estimate of the risk aversion parameter from the first iteration is 1 then in the second iteration 1 is set as the central support point while the support range is redefined to 0.35-1.65. The iterative process continues until the risk aversion parameter converges to a single value between 0 and 1.65.

The central support points of the non-diagonal parameters of the Q matrix are set equal to 0. This implies that we do not know a priori whether activities are complements or substitutes. The width of the support range of the non-diagonal element of Q is set equal to the average support range of the diagonal elements. Given that the expected value of q_{ij} for $i \neq j$ is 0 equation (10) can be re-written as:

$$\lambda_i + \lambda_i^{'} = d_i + q_{ii} \cdot x_{it}^{0} + \varphi \cdot \gamma_i \quad \forall\ i \in \{1 ... N\} \qquad (21)$$

This equation is satisfied for any positive value of α that satisfies the relationship below:

$$q_{ii} = \frac{\alpha_i \left| \lambda_i + \lambda_i' - \varphi \cdot \gamma_i \right|}{x_i^{0}} \quad and \quad d_i = \left(\lambda_i + \lambda_i' - \varphi \cdot \gamma_i \right) - \alpha_i \left| \lambda_i + \lambda_i' - \varphi \cdot \gamma_i \right| \qquad (22)$$

It can be proven that parameter α is related to the own price elasticity according to equation (23) (Heckelei and Britz, 2005; Kanellopoulos *et al.*, 2010).

$$\eta_i = \frac{r_i}{\alpha_i \cdot \left| \lambda_i + \lambda_i^{'} - \varphi \cdot \gamma_i \right|} \qquad (23)$$

Where η_i is the price elasticity of activity i. The values of q_{ii} and d_i related to the price elasticity of each activity are calculated by equations 24.

$$q_{ii} = \frac{r_i}{\eta_\iota x_i^{0}} \quad and \quad d_i = \left(\lambda_i + \lambda_i' - \varphi \cdot \gamma_i \right) - \frac{r_i}{\eta_\iota} \quad \forall\ i \in \{1 ... N\}\ and\ t \in \{1 ... T\} \qquad (24)$$

The expected value of the risk aversion parameter (i.e. 0.825) is used to calculate values of parameters q_{ii} and d_i in each observed year. The average values of q_{ii} and d_i over the observed years are used as the central support points of these parameters while the standard deviation of the parameters are used to define the remaining four support points. The first and the second support points were defined as 1.5 and 3 standard deviations above the central support points while the other two support points were defined as 1.5 and 3 standard deviations below the central support point. It is common that the three standard deviations rule is used for defining the support range of the unknown parameters (Golan *et al.*, 1996). Table 1 summarizes the selected support points of each estimated parameter.

4.3.2. Setup of the ex-post experiment

The application uses panel data of one arable farm type in Flevoland (the Netherlands) and one in the Midi-Pyrenees (France). For both farm types, the model was estimated based on observed data of the years 1999-2001 and was used to predict the cropping patterns of year 2002 and 2003. No major policy changes occurred in the period 1999-2003 in both regions, and consequently the same set of policy constraints (see section 2) were used in the estimation and forecasting phase.

Average prices and yields[5], and changes in subsidy levels for years 1999-2003 were found in the data base of EUROSTAT. A three years moving average was used to calculate the expected average prices and yields for each of the years 1999-2003. EUROSTAT prices, yields and received subsidies of period 1996-2001 were used to calculate the variance covariance matrix of revenues (APPENDIX 3). Observed crop levels, and available farm resources (e.g. land, family labour) were taken from FADN. Other technical coefficients related to agro-management (e.g. labour requirements, fertilization requirements, irrigation) were taken from the survey conducted within SEAMELSS (Borkowski *et al.*, 2007; Zander *et al.*, 2009). It was assumed that those coefficients, which were collected for the year 2003, are the same for all simulated years.

Elasticities from Jansson (2007) were used to define support points for the unknown parameters *d* and *Q*. The Percent Absolute Deviation (PAD)[6] was used to measure differences between observed and simulated crop levels and assess the forecasting performance of the model. Input data of the farm types of Flevoland and Midi-Pyrenees are presented in Table 2 and Table 3 respectively.

---

[5] Within the SEAMLESS survey, crop yields have been quantified per rotation, soil type and management for year 2003. In EUROSTAT, annual crop yields are aggregated to crop levels, which means that differences between rotations, soils and managements are lost. The disaggregated yield data of the SEAMLESS survey for year 2003 was used to disaggregate the average crop yields reported in EUROSTAT for a number of years. The model in (5), i.e. the linear version of FSSIM with calibration constraints, was optimized for 2003, using the SEAMLESS survey data, to find the optimal set of activities (combinations of rotation, soil type and management) that result in the observed crop levels. The average yield in 2003 was calculated for each crop. The disaggregated yield of each crop per rotation soil and management is seen as a percentage difference from the simulated average yield. For some of the activities the percentage difference is a positive number while for others, it is a negative number. Assuming that the aggregated average yields of EUROSTAT correspond to the average yield of the optimal farm plan in each year and using the activity specific percent difference from the average of the year 2003 we calculated the average yield per rotation, soil and management for each simulated year.

[6] The Percentage Absolute Deviation (PAD) is defined as the absolute deviation between simulated and observed activity levels per unit of actual activity level.: $PAD(\%) = 100 \cdot \left(\sum_i \left| x_i - x_i^0 \right| \right) / \left(\sum_i x_i^0 \right)$

The support range of the non-diagonal elements of Q were set approximately equal to the average support range of the diagonal elements of Q. The support range of the diagonal elements of maize fodder in Flevoland and set-aside in Midi-Pyrenees were excluded from the calculations. This is because of the much larger standard deviation (and consequently support ranges) of these activities than the standard deviation of the other activities. The resulting support range for the non-diagonal elements of Q was set to -300-300 for both farm types. To determine appropriate support points for the error term of (8) we performed a sensitivity analysis that involved multiple model runs. In each model run, we used different support range for the error term. In all model runs the support ranges were symmetric around 0 which is the expected value of the error term. The support range was defined as a number of standard deviations of the activity's marginal costs above and below 0. The minimum number of standard deviations that gave acceptable fit to observed data for years 1999-2001 was used for simulations.

Table 1: Definition of support point for each unknown parameter.

	Support points				
Parameter	1	2	3	4	5
ε_{it}	$-Nr \cdot \sigma_i^\lambda$	$-0.5 \cdot Nr \cdot \sigma_i^\lambda$	0	$0.5 \cdot Nr \cdot \sigma_i^\lambda$	$Nr \cdot \sigma_i^\lambda$
φ	0	0.4215	0.825	1.2375	1.65
q_{ij} for i≠j	-300	-150	0	150	300
q_{ij} for i=j	$\overline{q}_{ii} - 3 \cdot \sigma_{ii}^q$	$\overline{q}_{ii} - 1.5 \cdot \sigma_{ii}^q$	$\overline{q}_{ii} = \frac{\sum_t \frac{r_{it}}{\eta_{it} x_{it}^0}}{T}$	$\overline{q}_{ii} + 1.5 \cdot \sigma_{ii}^q$	$\overline{q}_{ii} + 3 \cdot \sigma_{ii}^q$
d_i	$\overline{d}_i - 3 \cdot \sigma_i^d$	$\overline{d}_i - 1.5 \cdot \sigma_i^d$	$\overline{d}_i = \frac{\sum_t \lambda_{it} + \lambda'_{it} - \phi \cdot \gamma_i - \frac{r_{it}}{\eta_{it}}}{T}$	$\overline{d}_i + 1.5 \cdot \sigma_i^d$	$\overline{d}_i + 3 \cdot \sigma_i^d$

Where $\overline{q}_{ii}$ is the average value of q_{ii}, $\overline{d}_i$ is the average value of d_i, σ_i^λ is the standard deviation of the shadow price of the calibration constraint of activity i, σ_i^d is the standard deviation of parameter d_i, σ_{ii}^q is the standard deviation of parameter q_{ii} of each activity, Nr is the number of standard deviations that determines the upper and lower bound of the support space of the error term, and *T* is the number of years used for estimation (in our case *T*=3). Only the initial support points of φ are reported. The value of φ is determined in an iterative procedure and the support points of φ change in each iteration.

Table 2: Three year moving average yields (tonnes/ha), three year moving average prices (€/tonnes), received subsidies (€/ha), average costs (€/ha) and gross margins (€/ha) for years 1999-2003 of the arable farm type in Flevoland.

	Avg. yield (tonnes/ha)					Prices (€/tonne)					Subsidies (€/ha)					Costs	Gross margin (€/ha)				
	99'	00'	01'	02'	03'	99'	00'	01'	02'	03'	99'	00'	01'	02'	03'	(€/ha)	99'	00'	01'	02'	03'
Maize (fodder)	44.8	47.8	46.7	46.9	45.2	19	22	24	24	23	336	434	467	424	364	1098	89	388	490	452	306
Onions	35.4	32.7	31.5	35.5	37.4	173	173	153	149	163						2158	3966	3499	2662	3132	3938
Potatoes	43.2	43.8	44.1	44.5	44.2	96	126	115	96	70	53	60	72	58	66	1787	2413	3792	3357	2543	1373
Set-aside											334	290	297	488	488	100	234	190	197	388	388
Sugar beet	53.9	55.9	56.8	58.8	57.5	54	50	48	48	49						1150	1761	1645	1576	1672	1668
Wheat (soft)	8.1	7.9	8.1	8.2	8	117	110	108	110	105	334	290	297	488	488	524	758	635	648	866	804

Table 3: Three year moving average yields (tonnes/ha), three year moving average prices (€/tonnes), received subsidies (€/ha), average costs (€/ha) and gross margins (€/ha) for years 1999-2003 of the arable farm type in Midi-Pyrenees.

	Avg. yield (tonnes/ha)					Prices (€/tonne)					Subsidies (€/ha)					Costs	Gross margin (€/ha)				
	99'	00'	01'	02'	03'	99'	00'	01'	02'	03'	99'	00'	01'	02'	03'	(€/ha)	99'	00'	01'	02'	03'
Barley	6.3	6.3	6.4	6.1	6.3	117	113	111	110	105	315	331	354	360	361	340	712	703	724	691	683
Maize (fodder)	38.5	41.4	41.1	42.0	42.2						310	335	365	359	361	860	-550	-525	-495	-501	-499
Maize (grain)	8.6	8.8	8.8	8.9	8.9	131	123	122	123	120	316	323	354	359	361	860	583	545	568	594	569
Rape seed	3.4	3.4	3.2	3.0	2.9	222	204	194	196	218	537	460	409	365	351	583	709	571	447	370	400
Set-aside											309	331	355	360	361		309	331	355	360	361
Soya	2.7	2.7	2.6	2.6	2.7	196	183	181	193	208	538	461	409	350	352	472	595	483	408	380	442
Sunflower	2.2	2.3	2.3	2.4	2.4	224	221	214	222	249	537	460	409	351	351	294	736	674	607	590	655
Wheat (durum)	4.4	4.4	5.0	4.7	4.7	177	167	148	147	153	309	331	355	360	361	421	667	645	674	630	659
Wheat (soft)	7.3	7.3	7.5	7.1	7.2	128	120	114	114	112	309	331	355	360	361	430	813	777	780	739	737

4.4. Results

The sensitivity of results to the support range of the error term is presented in Figure 1. The support range is determined as plus and minus a number of standard deviations from 0. It was found that in both farm types support ranges larger than ±6 standard deviations resulted in satisfactory simulations. For that reason, the support range of the error term was set to ± 6 standard deviations.

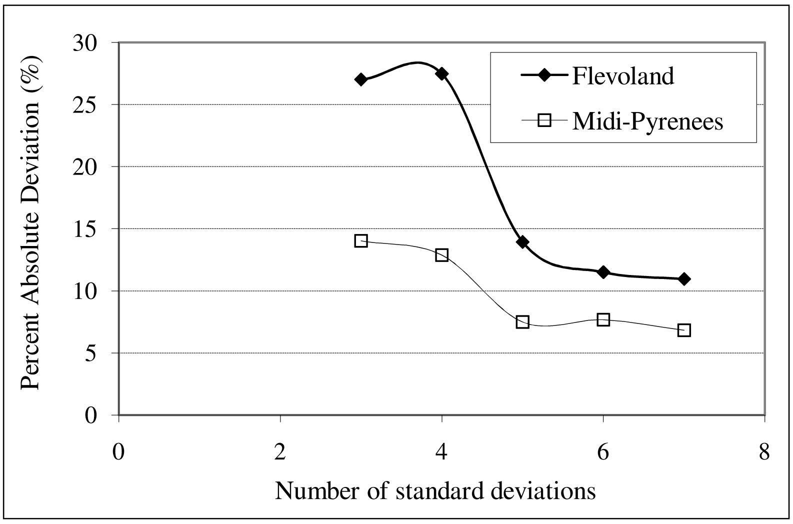

Figure 1: Average percent absolute deviation (PAD) from years 1999-2001 achieved with different widths of the support range of the error term of the marginal cost constraint (equation 8). On the Y-axis the Percent Absolute Deviation (PAD). On the X-axis the radiant of the support range in standard deviations from 0.

Results from the iterative process that was used to estimate the risk aversion coefficient in both farm types are presented in Figure 2. After 20000 iterations the estimated value of φ for the simulation of Flevoland converged to 1.21, which is 47% higher than the center of the initial support range. The risk premium calculated as $\varphi \cdot \sigma_z$ was found to be 8.6% of the total gross margin. In Midi-Pyrenees, after 20000 iterations the estimate of φ had not completely converged. Nevertheless, the change of value of φ after 20000 iterations is negligible (Figure 2). The estimated value of the risk aversion parameter of the arable farmer in Midi Pyrenees is 1.62, which is 96% above the initial centre of the support range of φ. The risk premium in Midi-Pyrenees is 7.5%. The larger risk aversion coefficient in Midi-Pyrenees suggests that the farmer of the arable farm type in Midi-Pyrenees is more risk averse than the farmer of the arable farm type in Flevoland. The larger risk premium in Flevoland suggests that the variation of farm income in Flevoland is larger than that of Midi-Pyrenees.

The estimated parameters d and Q of the quadratic gross margin function of FSSIM for the arable farm type of Flevoland are presented in Table 4.

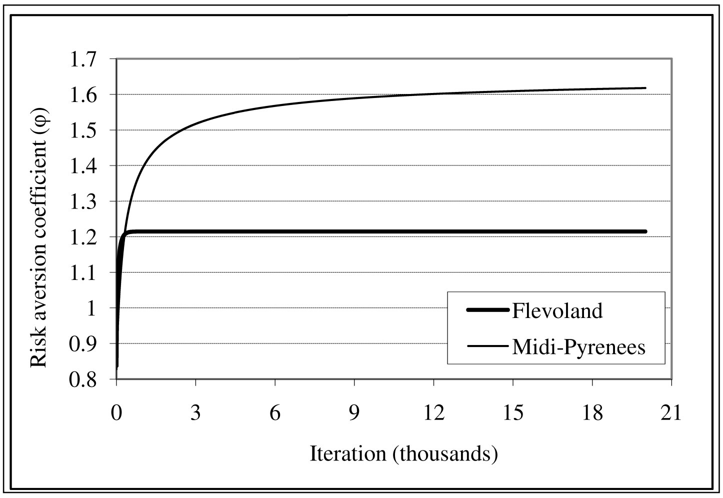

Figure 2: The estimated risk aversion coefficient for the tested farm types of Flevoland and Midi-Pyrenees for different number of iterations (thousands).

Table 4: Estimated parameters of the quadratic objective function of FSSIM for the arable farm type of Flevoland.

		d vector	Q matrix					
			MAIF	FVEO	POTA	FASE	SUGB	SWHE
Maize (fodder)	MAIF	-3480	204	14	24	1	-2	38
Onions	FVEO	-17380	14	2074	-1	0	5	-8
Potatoes	POTA	-13824	24	-1	538	-1	-9	-5
Set-aside	FASE	-3921	1	0	-1	1181	-1	4
Sugar beet	SUGB	-3974	-2	5	-9	-1	281	17
Wheat (soft)	SWHE	-2555	38	-8	-5	4	17	90

Complementarities and substitution between activities can be identified from the non-diagonal elements of the estimated *Q* matrix. A positive value of the non diagonal element (q_{ij}) of the *Q* matrix implies that activity *i* and *j* are substitutes since costs of activity *i* increase by increasing the level of activity *j*. Complementarity occurs when the non-diagonal elements of two activities are negative. This means that increasing the level of one activity reduces the costs of the other. Reasons for this could be beneficial effects of pest and disease management (which implies lower costs) or more efficient use of machinery and labour. Substitution can occur when activities compete for the same machinery and management or when two activities have common pests and diseases and consequently growing both crops increases the costs of pest and disease management. In many cases it is difficult to identify specific interactions between different activities because more than one interactions can occur. Only the total effect is reflected in the non-diagonal elements of the Q matrix. In the farm type of Flevoland, the major interaction between crops is the relationship of substitution between maize for fodder and potatoes, onions and soft wheat. The production costs of fodder maize increase with each extra hectare of potatoes, onions or soft wheat (and vice versa). The sowing and harvesting dates of these crops in Flevoland are close to each other and consequently competition for machinery, management and labour occurs. Substitution also occurs between sugar beet and soft wheat. The soil preparation period for winter soft wheat overlaps with the harvesting period of sugar beet resulting in competition for management and labour. Soft wheat complements the production of onion and potatoes. A possible explanation for such a relationship is the fact that costs for pest and disease management of onions and potatoes decrease with an increasing share of wheat in the rotation.

Detailed results on activity levels for years 1999-2003 for the arable farm type of Flevoland are presented in Table 5. As expected, the average achieved PAD value for years used for estimation i.e. 1999-2001 is lower than the PAD values of the forecasted years i.e. 2002 and 2003. In general, the average gross margin of each crop as those are reported in Table 2 and the available farm land explain the changes in the simulated cropping patterns. The simulations of years 2001 and 2003 resulted in higher PAD values. The higher PAD value of the year 2001 compared to the other years which were used for estimation is due to the poor performance of the model in simulating the area of soft wheat. The reason for the poor performance is explained from the increasing observed area (by almost 60%) and the decreasing gross margin of soft wheat from year 1999 to 2001. The gross margin of soft wheat in 2001 is higher than that of 2000 but the total farm area decreases substantially and that explains the larger simulated area of soft wheat in 2000. The PAD value for the year 2003 is 19% and it is that high mainly because of the poor simulation of the area of potato and sugar beet. In 2003, the observed area of potatoes was the same with the area of potato in years 1999, 2000 and 2002. Nevertheless, the gross margin of potatoes decreased by almost 55% compared to the average gross margins in the previous years. As a result in 2003, the simulated area of potatoes is smaller than the observed area of potatoes. The simulated area of sugar beet in 2003 increased because of the increase in the expected gross margin. The observed area of sugar beet in 2003 did not follow the increase of the gross margin because of quota restrictions that are not included in the model specification. The reason for this is lack of good quality information on quota levels and on penalties for exceeding the quota in FADN and

EUROSTAT. The estimated parameters d and Q of the objective function of FSSIM for the arable farm type of Midi-Pyrenees are presented in Table 6.

Complementarity and substitution between activities is more common in Midi-Pyrenees than in Flevoland. The estimated non-diagonal elements of the *Q* matrix are in general higher in Midi-Pyrenees than in Flevoland. This is mainly because the crops grown in Midi-Pyrenees are cereals, oil seeds or legumes. Activities of the same group have similar requirements for machinery and management and they are threatened by the same pests and diseases. As a result, their costs are dependent. Many different interactions can occur between crops of the same or similar crop groups which complicates the interpretation of the estimated non-diagonal elements of *Q*. An important interaction between activities that can be easily identified is the complementarity between barley and soya. If barley follows soya in the rotation, the nitrogen fixed by soya is released to barley. This means that for barley less nitrogen input from fertilizer is required. Soya and rape seed are substitutes. A possible reason for this relationship is the sowing date of rape seed which is in general a few weeks earlier than the harvesting date of soya. As a result competition for machinery and labour but also inefficiencies in management can occur while rapeseed cannot benefit from the nitrogen fixation of soya because it cannot follow soya in the rotation. Inefficiency in management and competition for labour and machinery can be also the explanation for the substitution between rape seed and durum wheat. The harvesting date of durum wheat coincides with the sowing date of rape seed. Maize and set-aside appear to be strong substitutes which is difficult to explain from the available information. Results from the model simulations for years 1999-2003 for the arable farm type of Midi-Pyrenees are presented in Table 7.

The achieved PAD values for 1999-2001 range between 3% and 12% while the PAD values of the years 2002 and 2003 are 11% and 15% respectively. The higher PAD value achieved in the year 2003 is mainly because of poor prediction of the areas of wheat (soft and durum), and maize. The observed area of maize for grain in 2003 increased by more than 7% from 2002 while the expected gross margin decreased by more than 4%. The observed area of soft wheat decreased substantially while the gross margin for the year 2003 did not change much. As a result the simulated area of soft wheat did not follow the changes in the observed levels of this crop. The deviations between predicted and observed areas of maize, and soft wheat, the changes to the available land, the interactions between crops (complementarity, substitution) and the changes in gross margin are responsible for the poor predictions of soya and durum wheat in 2003.

Table 5: Observed and simulated activity levels for years 1999-2003 for the arable farm type of Flevoland.

	Simulations of years used for estimation									Simulations of years used for forecasting					
	1999			2000			2001			2002			2003		
	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$
Maize (fodder)	1.7	0.9	0.8	1.9	2.1	0.2	0.6	2.3	1.7	0.7	2.3	1.6	0.6	2.7	2.1
Onions	9.1	9.2	0.1	9.1	9.0	0.1	8.5	8.5	0.0	7.9	8.8	0.9	9.7	9.3	0.4
Potatoes	24.8	22.6	2.2	24.7	25.0	0.3	21.9	23.8	1.9	24.6	22.6	2.0	24.8	20.8	4.0
Set-aside	1.3	1.4	0.1	1.3	1.3	0.0	1.3	1.2	0.1	1.7	1.5	0.2	1.3	1.7	0.4
Sugar beet	12.2	12.0	0.2	12.1	11.4	0.7	9.3	10.7	1.4	10.0	11.4	1.4	9.1	12.1	3.0
Wheat (soft)	6.1	9.1	3.0	6.0	6.3	0.3	9.7	4.8	4.9	10.7	8.9	1.8	11.5	10.6	0.9
Total (ha)	55.2	55.2	6.6	55.1	55.1	1.6	51.3	51.3	10.1	55.6	55.6	7.9	57.0	57.0	10.8
PAD (%)			12			3			20			14			19

Table 6: Estimated parameters of the quadratic objective function of FSSIM for the arable farm type of Midi-Pyrenees.

		d vector	Q matrix								
			BARL	MAIF	MAIZ	RAPE	FASE	SOYA	SUNF	DWHE	SWHE
Barley	BARL	-464	147	1	10	16	-8	-50	-7	-14	10
Maize (fodder)	MAIF	-1308	1	1128	1	0	0	4	-3	-3	1
Maize (grain)	MAIZ	-1088	10	1	29	-16	22	-2	2	-2	-14
Rape seed	RAPE	-1297	16	0	-16	390	17	31	-9	29	-12
Set-aside	FASE	-3556	-8	0	22	17	304	-23	2	-37	26
Soya	SOYA	-831	-50	4	-2	31	-23	52	23	19	0
Sunflower	SUNF	-505	-7	-3	2	-9	2	23	37	-7	-5
Wheat (durum)	DWHE	-627	-14	-3	-2	29	-37	19	-7	37	13
Wheat (soft)	SWHE	-1259	10	1	-14	-12	26	0	-5	13	69

Table 7: Observed and simulated activity levels for years 1999-2003 for the arable farm type of Midi-Pyrenees.

	Simulations of years used for estimation									Simulations of years used for forecasting					
	1999			2000			2001			2002			2003		
	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$	x^0	x_i	$\|x^0-x_i\|$
Barley	2.7	3.3	0.6	3.1	3.5	0.4	5.1	3.6	1.5	3.0	2.5	0.5	4.1	2.6	1.5
Maize (fodder)	0.1	0.1	0.0	0.2	0.2	0.0	0.2	0.2	0.0	0.3	0.2	0.1	0.3	0.2	0.1
Maize (grain)	32.1	31.1	1.0	32.9	32.1	0.8	29.7	30.9	1.2	31.2	33.9	2.7	35.1	31.4	3.7
Rape seed	3.8	3.4	0.4	3.3	3.2	0.1	2.1	2.6	0.5	0.9	3.0	2.1	1.7	2.8	1.1
Set-aside	8.0	8.4	0.4	9.0	9.0	0.0	9.7	9.5	0.2	9.1	8.9	0.2	9.3	9.1	0.2
Soya	4.8	5.8	1.0	2.7	3.7	1.0	7.7	3.4	4.3	3.2	1.0	2.2	3.0	1.5	1.5
Sunflower	14.3	13.6	0.7	16.2	16.1	0.1	11.1	13.6	2.5	17.8	15.3	2.5	14.3	15.5	1.2
Wheat (durum)	11.5	12.3	0.8	16.8	16.1	0.7	15.8	17.1	1.3	17.0	16.7	0.3	17.3	16.8	0.5
Wheat (soft)	20.5	19.7	0.8	19.6	19.8	0.2	18.1	18.4	0.3	18.5	19.5	1.0	13.1	18.2	5.1
Total (ha)	97.8	97.8	5.6	103.8	103.8	3.2	99.5	99.5	11.9	101.0	101.0	11.5	98.2	98.2	14.7
PAD (%)			6			3			12			11			15

4.5. Discussion and conclusions

In both tested farm types the forecasting performance of the models estimated with ME is better than the forecasting performance of models calibrated with PMP. In Kanellopoulos *et al*. (2010) the different PMP calibrated models achieved PAD values of 19-76% in Flevoland and 20-26% in the Midi-Pyrenees. The advantage of ME estimation is that multiple years of observations, which are available in EU level databases, are utilised in the estimation of the unknown model parameters. Consequently, the ME estimated parameters are expected to give a better representation of farmer's behaviour than the PMP parameters. Moreover, unlike the PMP approach, the ME approach allows for estimating the non-diagonal elements of the quadratic cost function and the farmer's risk attitude, which provides useful information on production structure (complementarity and substitution between crops) and behaviour of the farmer.

In ME estimated models, the Normalized Entropy (NE) is used to measure the importance of the data in reducing uncertainty about the values of the unknown parameters. The NE is defined as the achieved entropy divided by the maximum possible entropy and takes values between 0 and 1. In cases where the NE is different than one it is concluded that some information has been extracted from the dataset.

The value of NE for both tested farm types is close to 0.99, which means that some information from the multiple year observations was used in the ME estimation procedure. Large numbers of normalized entropy are common in the literature (Oude Lansink, 1999). A reason for the high NE values is that the chosen support points of the unknown parameters, which were used as prior information, were biased towards the value of the parameters that would best fit the average observed cropping patterns of years 1999-2001. These values did not have to change much during the estimation phase to respect the imposed optimality rules at the observed crop levels (equations 8) of the ME procedure. Moreover, it is important to notice that the magnitude of the measure of the NE depends on the width of the support range of the unknown parameters. The absolute value of NE can be decreased by decreasing the width of the support range of some parameters. In both tested farm types we defined the support range for the parameters in a uniform way. This is how they would have been defined in bio-economic studies that aim at levels of analysis (e.g. EU) where region specific information about the actual value of these parameters is scarce.

An important advantage of the ME estimation procedure is the capacity to estimate the farm-specific risk aversion parameter. In LP studies, this parameter is either completely ignored from the analysis or it is used as a calibration parameter. In the latter case, the value of risk aversion parameter is chosen such that the best model fit is achieved. The problem with this approach is that any kind of model misspecification is captured by the calibrated risk aversion parameter. Moreover, this approach cannot be easily combined with PMP calibration because of the exact calibration feature. In risk programming models, calibrated with PMP, exact calibration would be achieved independent from the value of risk aversion. Using ME to estimate the risk aversion parameter of the model together with the parameters reflecting the production structure is a superior alternative. For both the farm types of Flevoland and Midi-Pyrenees, we used information from the literature in determining the support interval of the risk aversion parameter. In this paper, it is argued that the value of the risk aversion parameters can be used to compare the risk averse attitude of different farm types. It was found that farmers in Midi-Pyrenees are substantially more risk averse than farmers in Flevoland, a result that is in line with the cropping choices of farmers in the two regions. Oude Lansink (1999^b) used econometric procedures to estimate the utility functions of arable farmers in Flevoland which resulted in lower risk premiums (i.e. 3-5%) than what is found here. However, in Oude Lansink (1999^b) yield variation was not taken into account in the calculations of income variation and it was concluded that the measurement of relative risk aversion was smaller than the findings of other studies (i.e. Saha *et al*., 1994).

Using multiple year observations on activity levels to estimate the objective function of the model, requires a good knowledge of drastic changes or structural breaks in the period covered by the data. Such changes could be the result of policy changes at national or regional scale that may explain the observed behavior of farmers. Therefore, it is important that these changes are included in the specification of the model. Failure to do so may result in poor simulations of current and, consequently, future behavior. The larger the number of observations (longer periods) the higher the risk of omitting essential policies or region-specific policies from the specification of the model. In cases where exact calibration is important for a single year, a PMP based calibration procedure might be easier to implement.

ME estimation does not require strong arbitrary assumptions to estimate the unknown parameters, like PMP, but it still requires prior information that can dominate the whole estimation process. As argued before, the centre of the support interval is a critical choice, that may affect the value of the estimated parameters. Also, the value of the parameters depends on the validity of the behavioural constraints imposed. For the ME estimations presented here, we imposed the first order conditions from the utility maximisation problem as behavioural constraints. These constraints assume that the farmer operates at the optimum and consequently that the observed activity levels correspond to the optimal farm plan. Aggregated information on farm inputs and total costs that are available in EU level databases could be also used in a ME estimation framework to improve the estimates of the model and ameliorate the forecasts. In this study we did not use such information to avoid the mismatch with the detailed agro-management data coming from the survey conducted within SEAMLESS (Borkowski *et al.*, 2007; Zander *et al.*, 2009).

Maximum Entropy appears to be a promising estimation procedure for estimating bio-economic farm models and this paper shows that it is capable of estimating risk attitudes. Also, the approach utilises the available information more efficiently and allows for relaxing a number of strict assumptions that are made in currently known PMP approaches.

4.6. Acknowledgments

The work presented in this publication was partly funded by the SEAMLESS integrated project, EU 6th Framework Programme for Research Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2).

4.7. References

Borkowski, N., Hecker, J.M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G., Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 1: 33-34.

Buysse, J., Fernagut, B., Harmignie, O., Henry de Frahan, B., Lauwers, L., Polomé, P., Van Huylenbroeck, G., and Van Meensel J. (2007). Farm-based modelling of the EU sugar reform: impact on Belgian sugar beet suppliers. European Review of Agricultural Economics, 34: 21 - 52.

Dogliotti, S., M.K. van Ittersum & W.A.H. Rossing (2005). A method for exploring sustainable development options at farm scale: A case study for vegetable farms in South Uruguay. Agricultural Systems, 86: 29-51.

Golan A., Judge, G.G. Miller, D. (1996). Maximum Entropy econometrics: Robust estimation with limited data. John Wiley & Sons, pp. 324.

Hazell, B. R. P., Norton, D. R. (1986). Mathematical Programming for economic analysis in agriculture. New York: Macmillan publishing company.

Heckelei, T. (2002). Calibration and Estimation of Programming Models for Agricultural Supply Analysis. Habilitation Thesis, University of Bonn, Germany (http://www.ilr1.uni- bonn.de/agpo/ staff/heckelei/heckelei_hab.pdf).

Heckelei, T. (2003). Positive Mathematical Programming: Review of the standard approach. CAPRI working paper (http://www.ilr1.unibonn.de/agpo/rsrch/capri/ wrkpape.htm).

Heckelei, T. and Britz, W. (2005). Models based on Positive Mathematical Programming: State of the art and further extensions. In Proceedings of the 98th EAAE symposium: State of the art and new challenges, Parma, Italy.

Heckelei, T. and Wolf, H. (2003). Estimation of constraint optimization models for agricultural supply analysis based on generalised maximum entropy. European Review of Agricultural Economics, 30: 27-50.

Helming, J.F.M., Peeters, L., Veendendaal P.J.J. (2001). Assessing the consequences of environmental policy scenarios in Flemish agriculture. In: Heckelei, T., Witzke, H.P., and Henrichsmeyer, W. (Ebs): Agricultural sector Modelling and policy information systems. Proceedings of the 65th EAAE seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel, pp. 237-245

Howitt, R. E. (1995). Positive Mathematical Programming. American Journal of Agricultural Economics, 77: 329-342.

Janssen S., Louhichi K., Kanellopoulos A., Zander P., Flichman G., Hengsdijk H., Meuter E., Andersen E., Belhouchette H., Blanco M., Borkowski N., Heckelei T., Hecker M., Li H., Oude, Lansink A., Stokstad G., Thorne P., van Keulen H. and van Ittersum M. (2009). A generic bio economic farm model for environmental and economic assessment of agricultural systems. Submitted paper in special issue of Environmental management.

Janssen, S., van Ittersum, M., Louhichi, K., Zander, P., Borkowski, N., Kanellopoulos, A., and Hengsdijk, H. (2009). Integration of all FSSIM components within SEAMLESS-IF and a standalone Graphical User Interface for FSSIM, D 3.3.12. SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

Jansson, T. (2007). Econometric specification of constrained optimization models. PhD dissertation, Bonn University, electronic publication (http://www.ilr1.uni-bonn.de/agpo/rsrch/capri/ wrkpape .htm).

Júdez, L., Chaya, C., Martínez, S., González, A. A. (2001). Effects of the measures envisaged in “Agenda 2000”on arable crop producers and beef and veal producers: an application of Positive Mathematical Programming to representative farms of a Spanish region. Agricultural Systems, 67: 121-138.

Kanellopoulos, A., Berentsen, P.B.M., Heckelei, T., Van Ittersum, M.K., Oude Lansink, A.G.J.M. (2010). Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. Journal of Agricultural Economics, 61: 274-294.

Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Oude Lansink, A., and Van Ittersum, M. (2010). FSSIM, a Bio-Economic Farm Model for Simulating the Response of EU Farming Systems to Agricultural and Environmental Policies. Accepted in Agricultural Systems.

Oude Lansink, A. (1999[a]). Generalized maximum entropy estimation and heterogeneous technologies. European Review of Agricultural Economics, 26:101–115.

Oude Lansink, A. (1999[b]). Area Allocation Under Price Uncertainty on Dutch Arable Farms. Journal of Agricultural Economics, 50: 93:105.

Oude Lansink, A., Silva, E. and Stefanou, S. (2001). Inter firm and Intra firm efficiency measures. Journal of Productivity Analysis, 15: 185-199.

Paris, Q. and Howitt, R. E. (1998). An Analysis of Ill-posed Production Problems Using Maximum Entropy. American Journal of Agricultural Economics, 80: 124-138.

Röhm, O. and Dabbert, S. (2003). Integrating agri-environmental programs into regional production models: an extension of Positive Mathematical Programming. American Journal of Agricultural Economics, 85: 254-265.

Saha, A., Shumway, C. R., and Talpaz, H. (1994). Joint estimation of risk preference structure and technology using expo-power utility. American Jounal of Agricultural Economics, 76: 173-184.

Tonini, A., and Jongeneel R. (2008). Modelling dairy supply for Hungary and Poland by generalized maximum entropy using prior information. European Review of Agricultural Economics, 35: 219-246.

Van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Olsson, J. A., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A. E., van der Wal, T., Wien, J. E., and Wolf, J. (2008). Integrated Assessment of agricultural systems: a component-based framework for the European Union (SEAMLESS). Agricultural Systems, 96: 150-165.

Zander, P., Borkowski, N., Hecker, J. M., Uthes, S., Stokstad, G., Rørstad P. Kr., and Bellocchi, G. (2009). Conceptual Approach to Identify and Assess Current Activities. P.D 3.3.9. SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

APPENDIX 1: The support points of parameters d, Q and error terms for the arable farm type of Flevoland.

	Support points				
	1	2	3	4	5
Support points of parameters d					
FASE	-4625	-4273	-3921	-3569	-3218
FVEO	-22146	-19797	-17448	-15099	-12750
MAIF	-4382	-3949	-3515	-3081	-2648
POTA	-18068	-15969	-13869	-11770	-9671
SUGB	-4872	-4431	-3991	-3551	-3111
SWHE	-3487	-3007	-2528	-2049	-1569
Support points of parameters Q					
FASE.SUGB	-300	-150		150	300
FASE.POTA	-300	-150		150	300
FASE.SWHE	-300	-150		150	300
FASE.MAIF	-300	-150		150	300
FASE.FVEO	-300	-150		150	300
FASE.FASE	908	1044	1181	1317	1454
FVEO.SUGB	-300	-150		150	300
FVEO.POTA	-300	-150		150	300
FVEO.SWHE	-300	-150		150	300
FVEO.MAIF	-300	-150		150	300
FVEO.FVEO	1538	1803	2068	2334	2599
FVEO.FASE	-300	-150		150	300
MAIF.SUGB	-300	-150		150	300
MAIF.POTA	-300	-150		150	300
MAIF.SWHE	-300	-150		150	300
MAIF.MAIF	-1753	-251	1250	2752	4254
MAIF.FVEO	-300	-150		150	300
MAIF.FASE	-300	-150		150	300
POTA.SUGB	-300	-150		150	300
POTA.POTA	279	423	566	710	854
POTA.SWHE	-300	-150		150	300
POTA.MAIF	-300	-150		150	300
POTA.FVEO	-300	-150		150	300
POTA.FASE	-300	-150		150	300
SUGB.SUGB	169	226	282	339	396
SUGB.POTA	-300	-150		150	300
SUGB.SWHE	-300	-150		150	300
SUGB.MAIF	-300	-150		150	300
SUGB.FVEO	-300	-150		150	300
SUGB.FASE	-300	-150		150	300
SWHE.SUGB	-300	-150		150	300
SWHE.POTA	-300	-150		150	300
SWHE.SWHE	17	54	92	129	167
SWHE.MAIF	-300	-150		150	300

SWHE.FVEO	-300	-150		150	300
SWHE.FASE	-300	-150		150	300
Support points of error terms					
FASE.1999	-2085	-1043	0	1043	2085
FASE.2000	-2085	-1043	0	1043	2085
FASE.2001	-2085	-1043	0	1043	2085
FVEO.1999	-2022	-1011	0	1011	2022
FVEO.2000	-2022	-1011	0	1011	2022
FVEO.2001	-2022	-1011	0	1011	2022
MAIF.1999	-3208	-1604	0	1604	3208
MAIF.2000	-3208	-1604	0	1604	3208
MAIF.2001	-3208	-1604	0	1604	3208
POTA.1999	-4352	-2176	0	2176	4352
POTA.2000	-4352	-2176	0	2176	4352
POTA.2001	-4352	-2176	0	2176	4352
SUGB.1999	-1887	-944	0	944	1887
SUGB.2000	-1887	-944	0	944	1887
SUGB.2001	-1887	-944	0	944	1887
SWHE.1999	-1682	-841	0	841	1682
SWHE.2000	-1682	-841	0	841	1682
SWHE.2001	-1682	-841	0	841	1682

APPENDIX 2: The support points of parameters d, Q and error terms for the arable farm type of Midi-Pyrenees.

	Support points				
	1	2	3	4	5
Support points of parameters d					
BARL	-513	-489	-464	-440	-415
DWHE	-943	-787	-632	-476	-321
FASE	-57231	-36484	-15737	5010	25757
MAIZ	-1147	-1117	-1088	-1058	-1029
MAIF	-1447	-1378	-1308	-1239	-1169
RAPE	-1729	-1515	-1300	-1086	-872
SOYA	-1273	-1036	-800	-563	-327
SUNF	-795	-654	-512	-371	-229
SWHE	-1342	-1301	-1260	-1218	-1177
Support points of parameters Q					
BARL.BARL	34	91	149	206	264
BARL.DWHE	-300	-150		150	300
BARL.FASE	-300	-150		150	300
BARL.MAIZ	-300	-150		150	300
BARL.OFPL	-300	-150		150	300
BARL.RAPE	-300	-150		150	300
BARL.SOYA	-300	-150		150	300
BARL.SUNF	-300	-150		150	300
BARL.SWHE	-300	-150		150	300
DWHE.BARL	-300	-150		150	300
DWHE.DWHE	-4	18	40	61	83
DWHE.FASE	-300	-150		150	300
DWHE.MAIZ	-300	-150		150	300
DWHE.OFPL	-300	-150		150	300
DWHE.RAPE	-300	-150		150	300
DWHE.SOYA	-300	-150		150	300
DWHE.SUNF	-300	-150		150	300
DWHE.SWHE	-300	-150		150	300
FASE.BARL	-300	-150		150	300
FASE.DWHE	-300	-150		150	300
FASE.FASE	-3509	-830	1849	4527	7206
FASE.MAIZ	-300	-150		150	300
FASE.OFPL	-300	-150		150	300
FASE.RAPE	-300	-150		150	300
FASE.SOYA	-300	-150		150	300
FASE.SUNF	-300	-150		150	300
FASE.SWHE	-300	-150		150	300
MAIZ.BARL	-300	-150		150	300
MAIZ.DWHE	-300	-150		150	300
MAIZ.FASE	-300	-150		150	300
MAIZ.MAIZ	23	26	29	32	35

MAIZ.OFPL	-300	-150		150	300
MAIZ.RAPE	-300	-150		150	300
MAIZ.SOYA	-300	-150		150	300
MAIZ.SUNF	-300	-150		150	300
MAIZ.SWHE	-300	-150		150	300
OFPL.BARL	-300	-150		150	300
OFPL.DWHE	-300	-150		150	300
OFPL.FASE	-300	-150		150	300
OFPL.MAIZ	-300	-150		150	300
OFPL.OFPL	-5	562	1129	1696	2263
OFPL.RAPE	-300	-150		150	300
OFPL.SOYA	-300	-150		150	300
OFPL.SUNF	-300	-150		150	300
OFPL.SWHE	-300	-150		150	300
RAPE.BARL	-300	-150		150	300
RAPE.DWHE	-300	-150		150	300
RAPE.FASE	-300	-150		150	300
RAPE.MAIZ	-300	-150		150	300
RAPE.OFPL	-300	-150		150	300
RAPE.RAPE	267	330	393	456	520
RAPE.SOYA	-300	-150		150	300
RAPE.SUNF	-300	-150		150	300
RAPE.SWHE	-300	-150		150	300
SOYA.BARL	-300	-150		150	300
SOYA.DWHE	-300	-150		150	300
SOYA.FASE	-300	-150		150	300
SOYA.MAIZ	-300	-150		150	300
SOYA.OFPL	-300	-150		150	300
SOYA.RAPE	-300	-150		150	300
SOYA.SOYA	-63	35	132	229	327
SOYA.SUNF	-300	-150		150	300
SOYA.SWHE	-300	-150		150	300
SUNF.BARL	-300	-150		150	300
SUNF.DWHE	-300	-150		150	300
SUNF.FASE	-300	-150		150	300
SUNF.MAIZ	-300	-150		150	300
SUNF.OFPL	-300	-150		150	300
SUNF.RAPE	-300	-150		150	300
SUNF.SOYA	-300	-150		150	300
SUNF.SUNF	16	26	36	47	57
SUNF.SWHE	-300	-150		150	300
SWHE.BARL	-300	-150		150	300
SWHE.DWHE	-300	-150		150	300
SWHE.FASE	-300	-150		150	300
SWHE.MAIZ	-300	-150		150	300
SWHE.OFPL	-300	-150		150	300

SWHE.RAPE	-300	-150		150	300
SWHE.SOYA	-300	-150		150	300
SWHE.SUNF	-300	-150		150	300
SWHE.SWHE	58	63	69	74	80
Support points of error terms					
BARL.1999	-234	-117	0	117	234
BARL.2000	-234	-117	0	117	234
BARL.2001	-234	-117	0	117	234
DWHE.1999	-218	-109	0	109	218
DWHE.2000	-218	-109	0	109	218
DWHE.2001	-218	-109	0	109	218
FASE.1999	-342	-171	0	171	342
FASE.2000	-342	-171	0	171	342
FASE.2001	-342	-171	0	171	342
MAIZ.1999	-170	-85	0	85	170
MAIZ.2000	-170	-85	0	85	170
MAIZ.2001	-170	-85	0	85	170
OFPL.1999	-368	-184	0	184	368
OFPL.2000	-368	-184	0	184	368
OFPL.2001	-368	-184	0	184	368
RAPE.1999	-592	-296	0	296	592
RAPE.2000	-592	-296	0	296	592
RAPE.2001	-592	-296	0	296	592
SOYA.1999	-323	-162	0	162	323
SOYA.2000	-323	-162	0	162	323
SOYA.2001	-323	-162	0	162	323
SUNF.1999	-208	-104	0	104	208
SUNF.2000	-208	-104	0	104	208
SUNF.2001	-208	-104	0	104	208
SWHE.1999	-96	-48	0	48	96
SWHE.2000	-96	-48	0	48	96
SWHE.2001	-96	-48	0	48	96

APPENDIX 3: Variance covariance matrix of marginal revenues for the farm type of Flevoland and Midi-Pyrenees

Flevoland

	FASE	FVEO	MAIF	POTA	SUGB	SWHE
FASE	10163	18724	1384	9679	-7706	7906
FVEO	18724	1619271	141039	-1076893	133145	-23633
MAIF	1384	141039	16350	-131797	7641	-7296
POTA	9679	-1076893	-131797	3350643	-179445	-28414
SUGB	-7706	133145	7641	-179445	36372	6281
SWHE	7906	-23633	-7296	-28414	6281	22622

Midi-Pyrenees

	BARL	DWHE	FASE	MAIZ	MAIF	RAPE	SOYA	SUNF	SWHE
BARL	4232	6395	-978	6048	-1085	2563	5550	2254	6974
DWHE	6395	28360	-1097	12757	-977	4476	9027	7774	14831
FASE	-978	-1097	585	-2241	598	-2569	-3060	-1599	-2149
MAIZ	6048	12757	-2241	15442	-2305	10887	13136	7774	12588
MAIF	-1085	-977	598	-2305	623	-2555	-3092	-1510	-2214
RAPE	2563	4476	-2569	10887	-2555	14251	14385	8561	7911
SOYA	5550	9027	-3060	13136	-3092	14385	18732	10100	12356
SUNF	2254	7774	-1599	7774	-1510	8561	10100	6772	6760
SWHE	6974	14831	-2149	12588	-2214	7911	12356	6760	13858

Chapter 5

5. A Method to Select Alternative Agricultural Activities for Future-Oriented Land Use Studies

Argyris Kanellopoulos[1,2], Paul Berentsen[1], Martin van Ittersum[2] and Alfons Oude Lansink[1]

[1] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*
[2] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*

European Journal of Agronomy. Submitted

Abstract

Agricultural systems, all over the world, are challenged with substantial changes in the climatic, bio-physical and socio-economic environment. Agricultural and environmental policies are necessary to restrict the negative consequences of these changes and to facilitate the diffusion of technological innovations and alternative agricultural activities in order to achieve sustainable production of food and fibres. Ex-ante integrated assessment can ensure the effectiveness of such policies. Research can support ex-ante assessment of policies through bio-economic models which can be used to explore alternative activities and technological innovations. An approach that has been used in existing bio-economic studies for identifying alternative activities in a consistent and reproducible way is based on combinatorics and agronomic filtering rules. One important limitation of this approach is that the number of generated, feasible activities can increase exponentially with the number of crops, management options and bio-physical conditions of the region. Many of these activities are inferior with respect to their input-output relationships or irrelevant given a specific policy question. However, the multi-dimensional nature of the input-output relationships of such activities do not allow for a straight-forward selection. The objective of this research is to propose a methodology based on Data Envelopment Analysis (DEA) for identifying a manageable set of representative alternative activities out of the large set of possible alternatives which are interesting from both an economic and a policy point of view. The method is applied in a fertilization problem of arable farming in Flevoland (the Netherlands). In total 831 activities were selected with the proposed DEA method out of the 16,514 generated activities. The smaller set of activities is further analyzed using the optimization part of a bio-economic farm model. Subsequent use of the 16,514 activities and the 831 activities in the same farm model resulted in exact same results showing that the selecting method is valid. Especially when repeated calculations need to be done the selection procedure contributes in reducing the total time required for computation and facilitates the analysis of the results. The proposed method can be a complementary component for existing and future combinatorial tools that aim to identify and quantify alternative activities for policy assessment

Keywords: data envelopment analysis; agricultural activity; bio-economic models; land use; future studies.

5.1. Introduction

The last decades, interrelated changes in the social, economic and bio-physical environment affect the livelihood and the welfare of millions of people all over the world. Agricultural systems are challenged to deal with those changes by reducing their own environmental impact and by maintaining sustainable production of food and fibre. To achieve this, technological innovations and alternative agricultural activities that improve the efficiency of existing agricultural systems must be adopted. The diffusion of such technological innovations and alternative agricultural activities can be supported by agricultural and environmental policies (Olesen and Bindi, 2002). Ex-ante evaluation of such agricultural and environmental policies is a necessary step in the development of efficient and effective policy measures with desirable consequences at social, economic and environmental level. The European Commission has formalized this through a mandatory ex-ante impact assessment of new agricultural and environmental policies (EC, 2005). Research can support such requirements through future-oriented land use studies that employ an integrated and multi-disciplinary approach that involves analysis at multiple levels (Van Ittersum *et al*., 2008). Bio-economic models are defined as integrated economic evaluations of model formulations of biophysical processes that aim to simulate management decisions on resource allocation (Barbier and Bergeron, 1999; Janssen and Van Ittersum, 2007). Bio-economic models have been widely used for ex-ante assessment of policies.

Ex-ante assessment of agricultural and environmental policies using bio-economic models is not complete without exploring alternative activities and technological innovations at farm level. The production opportunities available to a farmer today are not the same as those available in the future because of changes in the social, economic, institutional and bio-physical environment. For meaningful ex-ante assessment of future policies a set of representative activities, which is adequate to satisfy all possible targets of different objectives, is needed. Selecting a representative set of alternative activities and opportunities given a specific policy framework is a challenging procedure because it can involve multiple and conflicting objectives of the different stakeholders. Also, the assessed policy regime and the available farm resources can restrict the feasible "window of opportunities" from which farmers can choose to make decisions for the future.

Procedures for the identification and quantification of alternative activities have been proposed in this journal (Hengsdijk and Van Ittersum, 2003). Existing bio-economic studies have used combinatorial approaches and filtering agronomic rules to identify alternative activities in a uniform and reproducible way (Dogliotti *et al*., 2003; Janssen, 2009). Crops, livestock, rotation requirements and management options are combined to agricultural activities that have specific input requirements. Outputs and externalities are quantified using bio-physical models and/or expert rules. The filtering rules used in this kind of tools are mainly related to crop frequency, crop sequence and management and they are used to filter out those combinations which are not feasible from an agronomic point of view. The quantified set of activities is then offered to a farm level optimization model to simulate the farmer’s behaviour. This approach assures that no feasible option from an agronomic point of view, is excluded a priory and that the set of generated activities includes a wide variety of options that will or may become available to farmers in the future. One important limitation of this approach is that the number of feasible activities can increase exponentially with the number of crops, managements and bio-physical conditions (Wossink *et al*., 1992; Dogliotti *et al*., 2003; Janssen, 2009).

Many of the activities generated by combinatorial approaches are inferior with respect to their input-output relationships or irrelevant given a specific policy question. However, the multi-dimensional nature of the input-output relationships of such activities does not allow for straight-forward selection. Offering the full set of generated alternative activities to bio-economic farm models increases computational costs and complicates the analysis of the simulated results of the optimization process. This holds in particular if the model has to be run several times to assess different scenarios.

Data Envelopment Analysis (DEA) (Charnes *et al*., 1978) is a method used in operational research to rank entities that convert multiple inputs into multiple outputs based on their capacity to convert those inputs into outputs. Such entities are defined as decision making units (DMU). The definition of a DMU is quite flexible and encompasses firms, farms or even agricultural activities. In general, the production process of a DMU, like an agricultural activity, involves multiple inputs and outputs, which makes the ranking complicated. Mathematical programming methods are employed to rank or screen multiple input multiple output DMUs in terms of converting inputs into outputs. The capacity of each DMU to convert inputs into outputs is evaluated and compared to the capacity of all other existing DMUs to convert inputs into outputs. A multi dimensional frontier is created by the superior decision making units while all other inferior decision making units are enveloped (enclosed) in this frontier. The inputs and outputs of DEA could be also seen as attributes or criteria of multi-criteria decision making (MCDM) methodology (Bouyssou, 1999; Stewart, 1996). Inputs can be seen as criteria to be minimized while outputs as criteria to be maximized. DEA can be a promising approach for further screening (Figure 1) a set of activities generated by combinatorial approaches and agronomic filtering rules for use in bio-economic modelling.

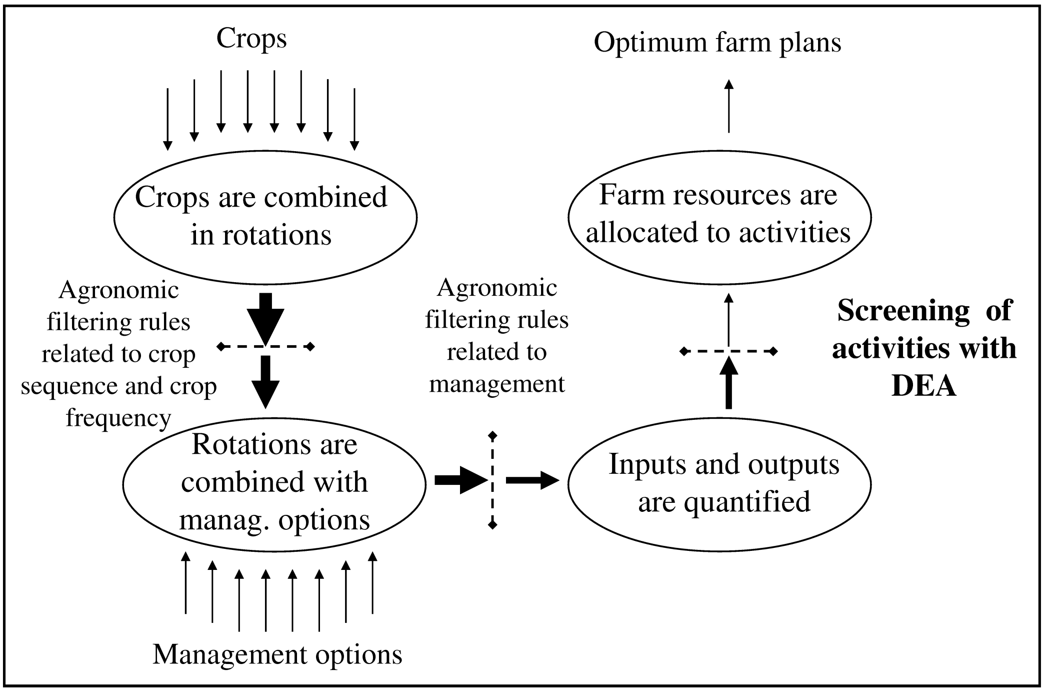

Figure 1: The position of the proposed DEA method within the process of generating and simulating alternative activities.

The objective of this article is to propose a methodology based on DEA for selecting a manageable set of policy specific and superior alternative activities out of the large set of possible alternatives generated by combinatorial processes. The proposed DEA method is used to identify superior activities from the set of activities generated in the Farm System SIMulator (FSSIM) which is the bio-economic model developed within the modeling framework of the System for Environmental and Agricultural Modeling: Linking European Science and Society (SEAMLESS) (Louhichi *et al.*, 2009).

In Section 2, the FSSIM modelling system is described, in Section 3 the DEA methodology for identifying superior alternatives is presented. An experiment related to fertilization options for arable farming in Flevoland (the Netherlands) is set up to demonstrate the method. In Section 4, the results are presented and Section 5 discusses and concludes.

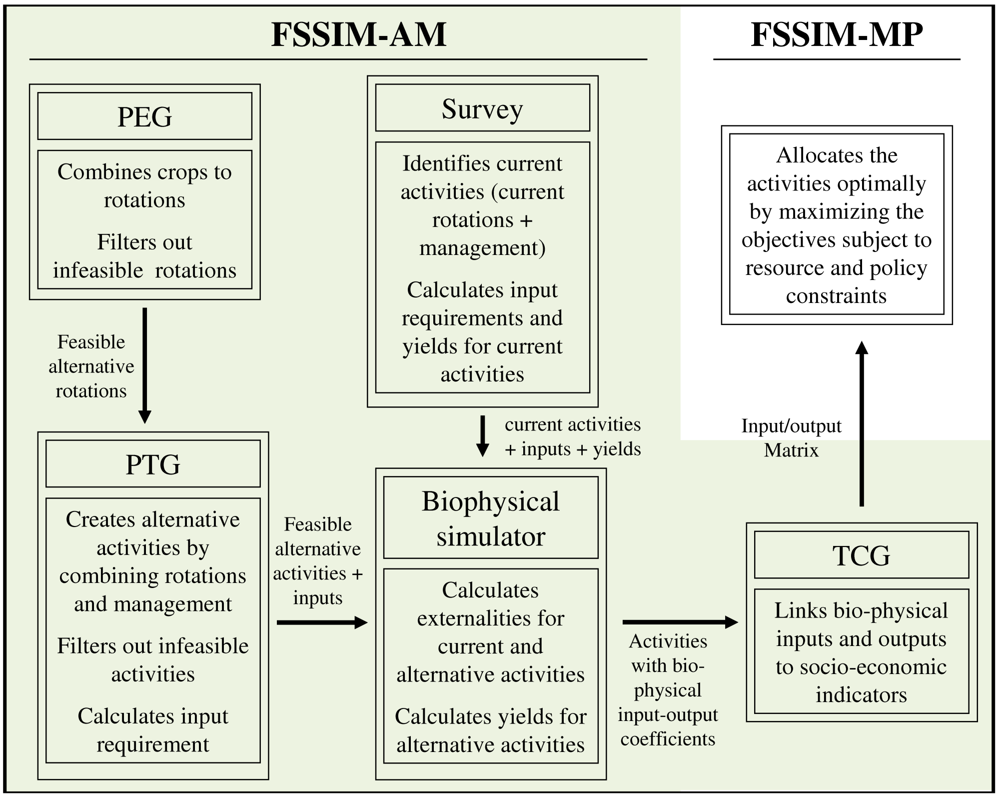

Figure 2: Functionality of and relationships between components of FSSIM

5.2. FSSIM for arable farms

The main objectives of FSSIM for arable farming are to calculate price-supply relationships of arable farming systems across the European Union (EU) and to enable detailed policy analysis at regional level. FSSIM for arable farms consists of two main components (Figure 2). The first component, is the agricultural management component (FSSIM-AM), which is used to identify, generate and quantify the technical coefficients (inputs and outputs) of current and alternative activities (Janssen *et al.*, 2010) while the second component is a constraint optimization model (FSSIM-MP) which is used to evaluate different scenarios by allocating activities to the available farm land (Louhichi *et al.*, 2010).

5.2.1. FSSIM agricultural management (FSSIM-AM)

The agricultural management component of FSSIM consists of a number of sub-components which are presented in Figure 2 and briefly described below. A more detailed description can be found in Janssen (2009).

The current activities, which are combinations of rotations and management options that are currently practised in the farm types of a certain region, were identified in a survey (Borkowski *et al*., 2007; Zander *et al*., 2009). Important input-output coefficients (e.g. yields, nitrogen application, pesticides) and prices are collected based on advisory handbooks and knowledge of experienced crop scientists. These input output coefficients are then used in bio-economic farm models to enable the calculation of a number of agro-ecological indicators.

Crop rotations that are not currently used in the region are generated in a combinatorial procedure, the Production Enterprise Generator (PEG) (Janssen, 2009). A number of crops, which are either available or expected to become available in the future are combined in crop rotations. The PEG is an extension of ROTAT (Dogliotti *et al*., 2003). It is assumed that the areas of all crops in each rotation are equal (e.g. each crop of a four year rotation of four different crops gets 25% of the total area) and all crops of a rotation are grown every year. In this way interactions between crops can be taken into account in a static way. A number of agronomic filters related to crop frequency and crop sequence are used to filter out rotations that are not feasible from an agronomic point of view because of characteristics of the crops and the bio-physical environment (e.g. crop rotations with a large share of crops vulnerable to soil-borne pests and diseases are filtered out because they would never be selected by the farmer due to substantial yield losses). Expert knowledge, empirical data and the literature are used to design such filtering rules.

The Production Technique Generator (PTG) (Janssen, 2009) describes current and alternative production techniques (i.e. water management, nutrient management, pest management, conservation management, planting-sowing and harvesting) for each feasible rotation (both current and alternative) generated by PEG. Most of the production techniques are defined per crop in the rotation but interactions between the different crops can be taken into account (e.g. N-inputs of a specific crop might be reduced in case the previous crop is a legume and/or if crop residues are incorporated into the soil). Filters related to production orientation (e.g. organic, conventional, irrigated) are used to filter out inconsistent activities. The number of activities (combinations of rotations and management options) can increase substantially with the number of different management options. The combinatorial explosion problem is even larger when combinations of different production techniques are allowed in the same rotations (e.g. alternative and conventional management co-exists in the same rotation).

The current and alternative activities (combinations of rotations and managements) and their input requirements can be assessed with a biophysical simulation model which quantifies yields and externalities. The Technical Coefficient Generator (TCG) (Janssen, 2009) links the input requirements, the yields and the externalities to economic parameters (prices and costs) to formulate the matrix of input-output coefficients that can be used in a bio-economic farm model like FSSIM-MP.

5.2.2. FSSIM mathematical programming (FSSIM-MP)

The mathematical programming part of FSSIM (Louhichi *et al.*, 2009) is a model that maximizes an objective function (e.g. gross margin or utility) subject to a set of resource and policy constraints. Positive Mathematical Programming (PMP) is used to calibrate to the observed activity levels (Kanellopoulos *et al.*, 2010). Activities generated by the agricultural management component of FSSIM are optimally allocated to the available farm land. Since the areas of crops in a rotation are fixed in the process of generating the activities there is no need for additional rotational constraints. The mathematical programming part of FSSIM is designed to be generic and easily adaptable to new regions and farm types (Louhichi *et al.*, 2010). The constraints and objectives of the model can be easily switched on and off depending on the policy question, the farmer’s objectives and the geo-political framework. A general formulation of FSSIM-MP is the one presented in (1).

$$\max \quad f(x), \qquad subject \ \ to: \ Ax \le b, \ \ x \ge 0 \tag{1}$$

Where *f(x)* is the farmer’s objectives, *x* is a $n \times 1$ vector of available agricultural activities (current and alternative), *A* is the $n \times m$ matrix of input-output coefficients and *b* is the $m \times 1$ vector of the right hand sides of the policy and resource constraints (e.g. the available land constraint per soil type, the on-farm available labor constraint, the irrigated land constraint, the sugar beet quota constraint and the obligatory set-aside constraint).

5.3. Methods

This section proposes a DEA based methodology to select relevant activities from a large set of alternative activities generated with combinatorial approaches and agronomic filtering rules. Furthermore, we describe the set up of a simple example on alternative fertilization options for arable farming in Flevoland.

5.3.1. Data Envelopment Analysis

A simple DEA example involving a set of decision making units (DMU) using one input to produce one output is shown in Figure 3. DMU's A, B and C are located on the frontier which reflects the best practice among the observed DMU's. These DMU are efficient since their use of inputs cannot be decreased or production of outputs cannot be increased without decreasing outputs or increasing inputs respectively (Cooper et. al., 2004, pg 3). DMU D is located below the frontier and is inefficient. Point F reflects a combination of A and B and creates the same output as point D, but uses less input. Point D can also be projected on the frontier by expanding output and holding input constant (as reflected by point H which is a combination of B and C). The input oriented efficiency score of D is calculated as θ = GF/GD while the output oriented efficiency score is calculated as θ = ID/IH. The DMU's A,B and C are fully efficient and have input and output oriented efficiency score of 1. Although the output oriented efficiency score of DMU E is equal to 1, it can be seen from the figure that the same output can be produced from a smaller quantity of input. In this example, DMU E is weakly efficient.

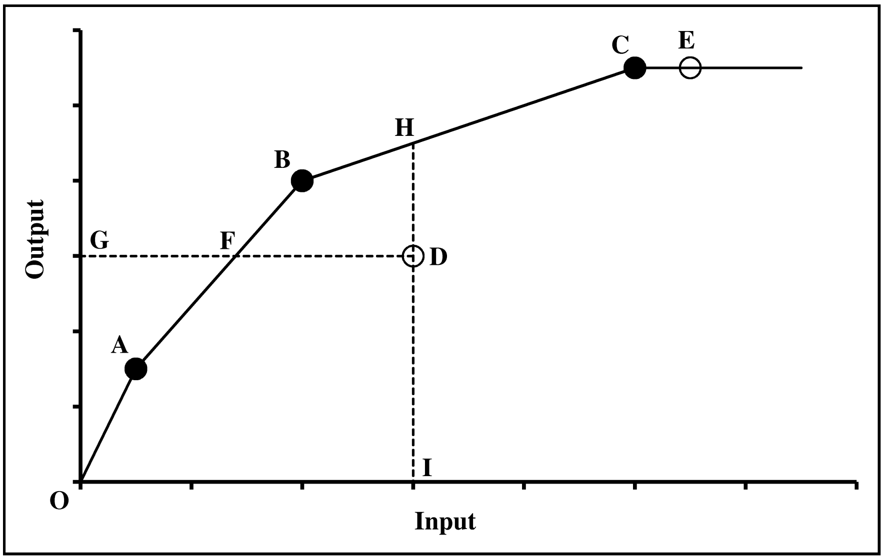

Figure 3: Graphical representation of a one input, one output DEA problem

In case of more complicated problems with multiple inputs and outputs a graphical solution is not possible. A Linear Programming model can be used to calculate the efficiency score of each DMU (see Appendix for the LP models).

The empirical implementation of a DEA model sets a number of requirements to the inputs and outputs (Cooper *et al*., 2007):

- The data must be non-negative for each of the DMUs; if not the data must be transformed to non-negative. At least one of the inputs and one of the outputs of each DMU should be positive.
- The inputs, outputs and DMUs that enter the DEA model should reflect the interests of the decision makers with respect to the components that enter the relative efficiency evaluation.
- In general, inputs are items that are preferred to be at a minimum level (as small as possible) while outputs are items that are preferred to be at a maximum level (as large as possible).
- The units of measurement of each input and output should be the same for the different DMUs.

5.3.2. DEA for selecting a representative set of superior alternative activities

This section proposes a three step method for selecting a representative set of superior alternative activities out of the large set generated with combinatorial approaches using DEA. In the first step, the data is transformed to satisfy the data requirements of DEA models (Section 3.2). The input-output coefficients of the alternative activities are found in the objective function and the *A* matrix of (1). There are a number of possible states of the data where data transformations are necessary to make the data set compatible with DEA requirements:

- The coefficients that correspond to a specific constraint are all non-positive. The coefficients can be transformed to non-negative, by changing the sign of the constraint from less than or equal to, in greater than or equal to and vice-versa.
- The coefficients that correspond to one of the objectives are all non-positive. The coefficients can be transformed to non-negative, by changing the optimization direction from minimization to maximization and vice-versa.
- Some of the coefficients of one of the constraints or objective functions are negative. The data can be transformed to non-negative by adding a constant to all coefficients of the constraint or objective function.
- All inputs or outputs of a certain activity are zero. A marginal positive value can be added to one of the inputs or outputs of the activity without implications for the results of the DEA model.

In the second step, the non-negative coefficients of the A matrix of model (1) and the non-negative coefficients of the objective function are separated into inputs and outputs for the DEA model. In general, inputs are items that are preferred to be at a minimum level while outputs are items that are preferred to be at a maximum level. To make the distinction between inputs and outputs more objective, we rewrite the LP model in (1) to distinguish upper and lower bound constraints:

$$\max \quad f(x), \quad subject\ to: \ A^u x \le b^u, \quad A^l x \ge b^l, \quad A^u, A^l, x \ge 0 \qquad (2)$$

Where A^u is the $w \times n$ matrix of transformed coefficients associated with upper bounds and A^l is the $p \times n$ matrix of transformed coefficients associated with lower bound constraints; b^u is the $w \times 1$ vector of upper bound resources and b^l is the $p \times 1$ vector of lower bound resources. The lagrangian function of the non-negative (transformed) optimization model is:

$$L(x, \pi^u, \pi^l) = \max_{x, \pi^u, \pi^l} f(x) - \pi^u (A^u x - b^u) + \pi^l (A^l x - b^l), \qquad x, \pi^u, \pi^l \ge 0 \qquad (3)$$

Where L is the lagrangian function, π^u is the $w \times 1$ vector of non-negative shadow prices of the upper bounds and π^l is the $p \times 1$ vector of non-negative shadow prices of the lower bounds. In the maximisation model above, objectives in $f(x)$ or coefficients that generate a positive contribution to the objective (A^l) are outputs in the DEA model, while objectives or coefficients with negative contribution to the objective are inputs in the DEA model.

In the third step of the proposed method, the relevant inputs and outputs of step 1 are used in a DEA model where each DMU (alternative activity) is evaluated in terms of its capacity to convert DEA inputs into DEA outputs. The efficient activities are selected and offered to the optimization model for policy assessment and scenario testing.

5.3.3. Set up of the experiment

In this section, the proposed DEA method is used to screen a representative set of alternative arable activities out of the large set of activities generated by FSSIM-AM.

5.3.3.1 The case study

In Flevoland, dairy farms import nutrients in the form of concentrates. Although a part of these nutrients return to the grasslands as organic manure for fertilization, a substantial surplus remains. The last decades, manure production, nutrient accumulation and reduction of nutrient surpluses have been the topic of policy debate (Berentsen and Tiessink, 2003). A viable option for reducing the nutrient surplus is to use manure on arable land replacing artificial fertilizers. A relevant question concerns the effects of an alternative management of arable crops where more organic manure is used to cover the nutrient requirements of the crops. A bio-economic farm model like FSSIM could be used to assess the consequences of such a decision on a number of important indicators for arable farms in Flevoland.

5.3.3.2 Generating rotations

The agricultural management component of FSSIM was used to generate alternative rotations which are feasible from an agronomic point of view and quantify their inputs and outputs. In total 8 crops that are currently grown in Flevoland (i.e. fodder maize, onions, potatoes, spring barley, spring soft wheat, winter soft wheat, sugar beet, set-aside) and 3 crops that according to experts may become more important in the near future due to economic and political changes (i.e. peas, winter rape seed, and tulips) were combined in rotations of maximum 5 years using the PEG. A number of filters of the PEG were used to select only the ones feasible from an agronomic point of view. Those filters are related to crop frequency, crop repetition, crop sequence, maximum number of different crops in the rotation, frequency of crop groups (e.g. cereals, oil seeds), repetition of crop groups, sowing dates and harvesting dates. According to experts, a crop frequency of tulips lower than 1 to 6 years is not possible because of increased incidence of pest and diseases and associated phyto-sanitary risks. To include rotations with tulips we also allowed 6 year crop rotations but only those with tulips. Clay soils are most common in Flevoland and for that reason only clay soils were simulated in this exercise.

5.3.3.3 Crop nutrient management

The starting point for the management of the activities was that from the survey for Flevoland (Section 2.1). For alternative activities, for each crop we used two different management options with respect to nutrient application. The total nitrogen application and the achieved yields were assumed the same in both management options but the type of fertilizers (artificial and/or cattle slurry) differ. The first management is the one that is currently mostly used in the region and it is based on artificial fertilizers (thus the data from the survey), while the second one is an alternative nutrient management which is based on (partial) replacement of fertilizer by organic manure (cattle slurry). Artificial fertilizers were used in the second option only when this was necessary to meet the crop’s total nutrient requirements. The one to one replacement of part of the nitrogen coming from artificial fertilizers with organic manure is possible only because the current nitrogen input from fertilizer in Flevoland is very high. Activities with applications of cattle slurry have higher labour requirements (Table 1) but also higher gross margins because of lower costs for fertilizers. To reduce the number of activities to feasible and operational levels we did not allow for combinations of crops with different management options in the same rotation. The nutrient management of all crops in a rotation is either based on artificial fertilizers (current management) or the management with cattle slurry complemented with artificial fertilizers when this was necessary. This decision limited the number of activities to only twice the number of rotations.

Table 1: Crop specific information, inputs and outputs for two different nutrient managements in Flevoland.

				Management with artificial fertilizers					Management with cattle manure and artificial fertilizers					
	Harvest (wk)	Sow (wk)	Yield (tons/ha)	Gr. margin (€/ha)	Labor (hrs/ha)	Fertilizers (kg N/ha)	N-leaching (kg N/ha)*	Organic matter change (score)	Gr. margin (€/ha)	Labor (hrs/ha)	Manure (tons/ha)	Fertilizers (kg N/ha)	N-leaching (kg N/ha)	Organic matter change (score)
Barley (spring)	32	10	6.3	1199	9.6	120	87	4.0	1264	16.2	24		46	6.0
Maize (silage)	41	17	40.8	533	7.1	185	135	2.5	662	13.7	38		69	5.2
Onions	36	14	58.4	3099	37.6	220	168	2.5	3249	44.2	40	24	98	6.0
Peas	30	13	5.7	1309	6.6	30	102	4.0	1340	13.2	6		100	4.2
Potatoes (seed)	33	15	38.7	4325	90.0	180	125	2.8	4418	96.6	20	82	93	4.5
Potatoes (ware)	39	15	56.8	3820	27.5	255	134	2.7	3945	34.1	30	108	81	5.0
Rape (winter)**	30	42	3.3	497	11.5	180	89	11.3	571	18.1	30	33	66	12.6
Set-aside	-	-	-	388	0.1		116	1.0	388	0.1			116	1.0
Sugar beet	42	14	65.5	2147	19.6	170	69	5.0	2218	26.2	19	77	41	7.0
Tulips	26	5	18	12974	604.0	120	167	1.2	13049	610.6	24		126	3.5
Wheat (spring)	36	11	7.8	1097	9.6	175	72	6.0	1158	16.2	25	53	32	7.7
Wheat (winter)	32	42	8.7	1324	10.4	205	74	8.8	1369	17.0	18	117	60	9.4

* No cover-winter crops were used for calculating the N-leaching.

** According to current management straw of cereals is removed while straw of winter rape (alternative crop) was incorporated into the soil.

Important environmental indicators of the activities, like nitrogen leaching and content of soil organic matter were quantified using NDICEA (Van der Burgt *et al.*, 2006). The NDICEA model uses region specific soil and climate data and crop-specific information to calculate states and flows of nutrients. The user defines a yield and nutrient inputs in different forms (e.g. artificial fertilizers, livestock manure) and the model calculates the nutrient balance based on the weather, soil, crop’s nutrient requirements, nutrient uptake rate and nutrient availability which is different for chemical and organic fertilizer. In NDICEA, when the user defined yields are not attainable with the given inputs (the nutrient uptake of the crop is higher than the available nutrients in the soil) the user have to adjust inputs and/or outputs so that nutrients available are always higher than nutrient uptake. It was assumed that cattle slurry can only be applied before sowing and artificial fertilizers were used when necessary to keep the available nitrogen well above the uptake during the season. More precisely, by choosing the proper combination of artificial fertilizers and cattle slurry, it was taken care that the available nitrogen was at least 20 kg N/ha above the nitrogen uptake of potatoes, onions and sugar beet and 10 kg N/ha above the nitrogen uptake of cereals and other crops. The nutrient composition of cattle slurry (i.e. 4.9 kg N, 1.8 kg P_2O_5 and 6.8 kg K_2O per ton of cattle slurry) available in NDICEA was used for calculations. The amounts of phosphate and potassium in the management with cattle slurry were at least equal to the application of the current management. Artificial phosphate and potassium fertilizers were added if necessary (i.e. peas, seed potato). For this exercise, to reduce the computational requirements we used NDICEA to calculate nutrient surplus of individual crops. It was assumed that differences between nutrient inputs of different rotations with the same nutrient management were only caused by different shares of crops in the rotations.

To account for crop frequency effects on crop yields (increased incidence of pest and diseases and phyto-sanitary risks) we used a yield correction factor which depends on the frequency of a crop in the rotation (Habekotté, 1994). The crop yields from the survey of current activities (Table 1) were corrected according to the frequency of the crop in the rotation using the correction factors of Table 2. It was assumed that the increased incidence of pest and diseases did not affect the nutrient inputs and the nutrient uptake of the crop.

Table 2: Yield correction factor for different crop frequencies (the value of one corresponds to yield from the survey).

	Frequency (ha of crop per ha of rotation)					
Crops	1:1	1:2	1:3	1:4	1:5	1:6
Potatoes (ware)	0.86	0.98	0.98	1	1.05	1.10
Potatoes (seed)	0.86	0.98	0.98	1	1.05	1.10
Onions	0.92	0.94	0.96	0.98	1	1.02
Sugar beet	0.55	0.66	0.78	0.95	1	1.05

5.3.3.4 The bio-economic farm model

A relatively simple version of FSSIM-MP was used since the purpose of the exercise is to demonstrate the applicability of the proposed method for selecting superior alternative activities. It was assumed that the farmer of an average farm in Flevoland maximizes the gross margin subject to the available land constraint, the labour availability constraint, the obligatory set-aside constraint and the sugar beet quota constraint. Two additional constraints were used to set an upper bound to the total nitrogen leaching and a lower bound to the soil organic matter content. These two last constraints can be seen as imposed restrictions of a hypothetical policy instrument that aims to restrict environmental impacts of arable farms. The objective and the constraints used in the farm model determine the inputs for the DEA-model for selecting superior activities. These are: labour, N-leaching and sugar beet production; outputs are gross margin, share of set-aside and change in organic matter.

5.3.4.5 Optimizations

To test the effectiveness of the method for selecting representative activities a number of farm model optimizations were done using alternatively the full set of activities generated by FSSIM-AM and the set of activities selected using DEA. To present the type of results expected in such a bio-economic analysis we performed three different optimizations.

1. An optimal farm plan was calculated for an average farm type in Flevoland. The resource endowments of the average farm type were calculated as weighted averages of the identified farm types of the SEAMLESS farm typology (Andersen *et al.*, 2007). No decrease in total content of soil organic matter was allowed.
2. Same as simulation 1, but now with different combinations of lower bounds on the total change in content of soil organic matter and upper bound on the total N-leaching.

3. Same as 1, but now with different combinations of upper bounds on total labour requirements and total N-leaching.

The right hand side of the equations of FSSIM-MP for the three simulations are summarized in Table 3.

Table 3: Right hand side (all expressed per farm) of the equations in FSSIM-MP in the three simulations.

	Simulation 1	Simulation 2	Simulation 3
objective	maximized	maximized	maximized
Available land (ha)	45	45	45
Available labor (hrs)	4754	4754	Parametric (from 0 to 5993)
Sugar beet quota (tons)	511	511	511
Obligatory set-aside (ha)	1.4	1.4	1.4
Change of organic matter (score)	225	Parametric (from 0 to 383)	225
Nitrogen leaching (kg N)	unbounded	Parametric (from 0 to 6555)	Parametric (from 0 to 6555)

5.4. Results

In total, 8257 rotations, which are feasible from an agronomic point of view, were generated. The generated set of rotations includes 6 two years rotations, 48 three years rotations, 184 four years rotations, 929 five years rotations and 7090 six years rotations. Combining the set of rotations with the two nutrient managements we end up with 16514 activities (twice the number of rotations). The DEA screening process resulted in 831 activities which are representative for all possible trade-offs between inputs and outputs of FSSIM-MP. The number of activities offered to FSSIM-MP is reduced by almost 95%. The substantially smaller set of activities not only decreases the computational time of FSSIM-MP but also enables more efficient analysis of the results because often, the researcher has to justify not only why specific activities are selected but also why other activities are not selected in FSSIM-MP.

The percentage of the full set of generated activities and of the set of activities screened with DEA per level of gross margin, content of soil organic matter, labour requirements and nitrogen leaching are summarized and presented in Figure 4.

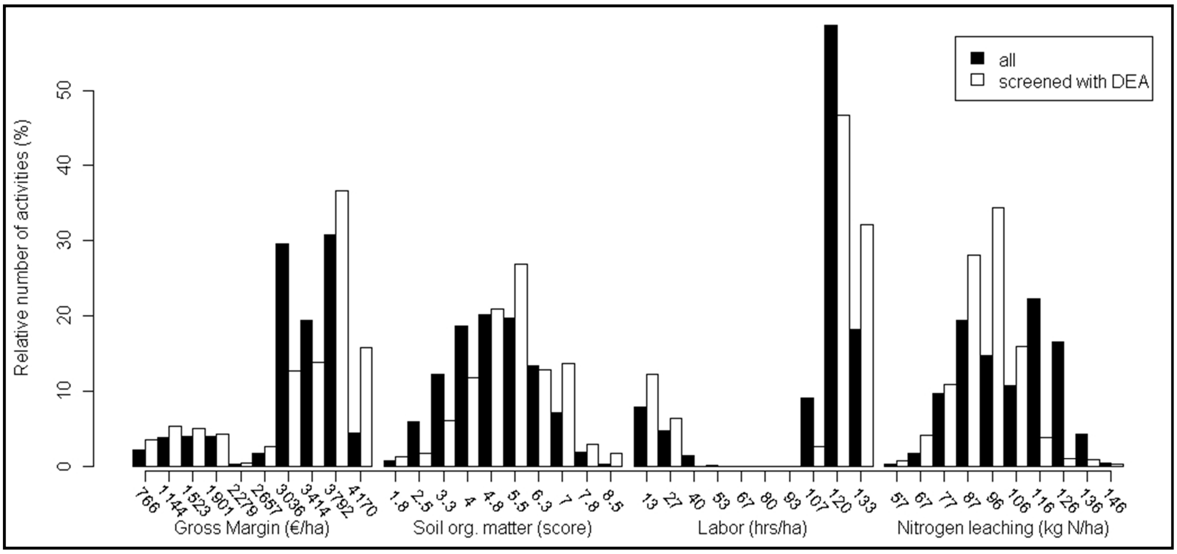

Figure 4: Distribution of the full and selected (with DEA) set of activities with respect to their gross margins, score of organic matter, labour requirements and N-leaching.

The gross margin of most of the generated activities is between 2700 and 3800 €/ha. However, activities with large and small gross margins are relatively overrepresented in the set of activities selected with the proposed DEA approach. A large share of the generated activities with gross margin higher than 3400 €/ha were selected with the DEA. Activities with gross margin lower than 1900 €/ha were also selected with the DEA approach because, often, high gross margin corresponds to high nitrogen leaching and high labour requirements. The frequency of the selected activities with respect to the content of soil organic matter follows the frequency of the full set of generated activities. However, a larger share of the generated activities with a score for soil organic matter content larger than 4 was selected with the DEA method. The frequency distribution of the DEA activities with respect to labour requirements follows also the frequency of the full set of generated activities. A relatively larger share of the activities with labour requirements less than 27 hrs/ha and more than 120 hrs/ha has been selected. Finally, for N-leaching, a larger share of the generated activities with leaching lower than 106 kg N/ha was selected.

Comparing results of FSSIM-MP where we offered only the set of activities selected with the proposed DEA method with results where we offered the full set of activities, shows that the set of activities selected with DEA is sufficient to calculate the trade offs between the different indicators. The model runs of FSSIM-MP in which we offered only the superior activities (i.e. 831 activities) resulted in the same gross margins as the corresponding model runs in which we offered to the model the full set of activities (16514 activities).

Table 4: Selected activities, corresponding inputs and outputs and farm level results in simulation 1 of Table 3.

	Simulated activities				Farm level results
	1	2	3	4	
Number of periods					
Period 1	Spring barley	Sugar beet	Set-aside	Set-aside	
Period 2	Potatoes	Winter wheat	Onion	Onion	
Period 3	Winter wheat	Potatoes	Winter wheat	Winter wheat	
Period 4		Winter wheat	Potatoes	Potatoes	
Period 5		Spring barley	Winter wheat	Winter wheat	
Period 6		Tulip	Tulip	Tulip	
Management	Fertilizers	Fertilizers	Fertilizers	Cattle slurry	
Gross margin (€/ha)	2071	3925	3937	4010	3770
N-leaching (kg N/ha)	98	101	122	90	103
Org. matter (score/ha)	5.2	5.1	4.2	5.7	5.0
Labour (hrs/ha)	16	114	115	121	106
Simulated level (ha)	3.9	32.7	6.1	2.3	45.0
Onion (ha)			1.0	0.4	1.4
Potatoes (ha)	1.3	5.4	1.0	0.4	8.2
Set-aside (ha)			1.0	0.4	1.4
Spring barley (ha)	1.3	5.4			6.8
Sugar beet (ha)		5.4			5.4
Tulip (ha)		5.4	1.0	0.4	6.8
Winter wheat (ha)	1.3	10.9	2.0	0.8	15.0

Results of FSSIM-MP for the average arable farm type in Flevoland (first simulation) are presented in Table 4. The selection of multiple rotations per farm was allowed in FSSIM-MP; three six-year rotation and one three-year rotation were selected in the optimum farm plan. For reasons of management and efficiency, in reality such a large number of crop rotations per farm might not be attractive to farmers; this could be solved by adding an extra constraint to the model. All six-year rotations included tulips which is the most profitable crop in the region. The higher labour requirements of activities with tulips are the reason for the three year rotation entering the solution. Activity 4 of Table 4 enters the solution because of the obligatory set-aside constraints and the high score in content of soil organic matter. Activity 3 of Table 4 enters the solution because of the obligatory set-aside constraint but also because of the lower labour requirements compared to activity 3. All constraints of FSSIM-MP except of the sugar beet quota constraints were binding.

Despite the higher gross margins, activities with the alternative nutrient management (i.e. with cattle slurry) were only marginally selected (5.2 %) in the optimum farm plan. The reason for this are the higher labour requirements. Changing from conventional to alternative nutrient management increases the total gross margin with ca. 4, 2 and 2 % for simulated activities 1, 2 and 3, respectively. However, the labour requirements increase more, i.e. with 42, 6 and 5 %, respectively.

The trade off between gross margin, N leaching and change in soil organic matter of the second simulation of FSSIM-MP is presented in Figure 5. As expected, the gross margin increases with increasing allowed leached nitrogen, while it decreases with increasing the lower bound to the score of soil organic matter.

The trade off between gross margin, nitrogen leaching and total labour requirements of the third simulation of FSSIM-MP is presented in Figure 6. The gross margin increases with increasing labour availability and increasing level of allowed nitrogen leaching.

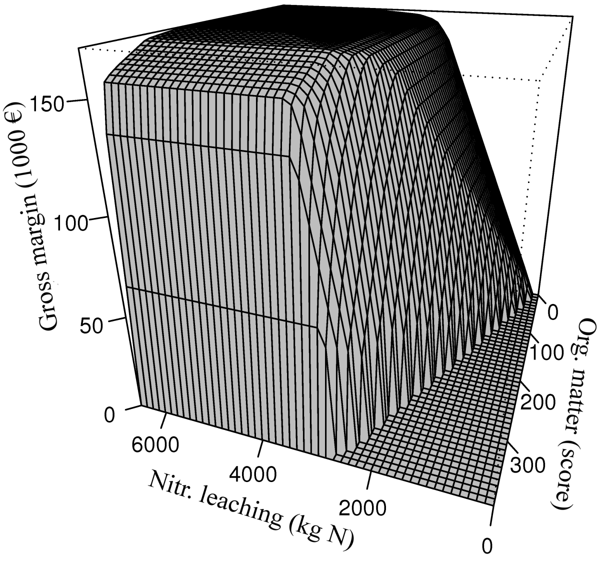

Figure 5: Trade-off curve between total gross margin, change of soil organic matter and nitrogen leaching.

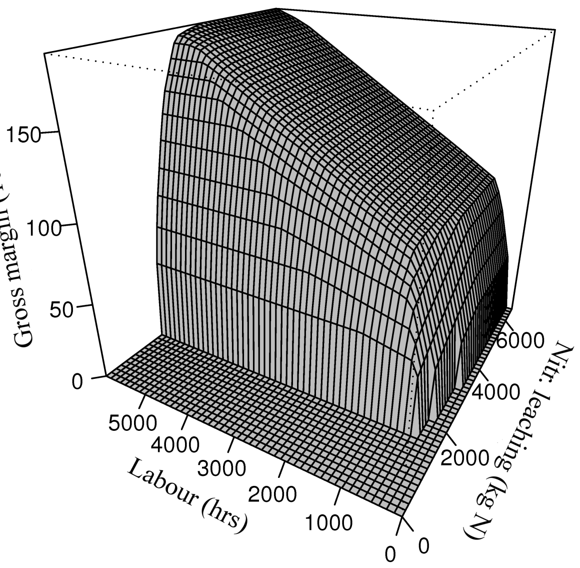

Figure 6: Trade-off curve between total gross margin, total labour requirements and nitrogen leaching.

5.5. Discussion and Conclusions

The number of alternative activities resulting from combinatorial approaches and agronomic filtering rules can increase substantially with the number of available crops, management options and interesting indicators from a policy point of view. This can create computational problems but also difficulties in analyzing the results of bio-economic or other models which are used for ex-ante evaluation of policies and future oriented land use studies. The proposed DEA method reduces the number of the activities to practical levels by filtering out inferior activities which will not be part of any optimum production plan. The results of the DEA method suggests that a large share (in our case almost 95%) of the activities generated in combinatorial methods is not relevant for policy assessment.

In the example presented in this article, the number of generated activities was relatively small (i.e. 16514) because we assumed a maximum rotation length of 5 years (only 6 year rotations that include tulips were allowed) and only 2 possible management options. Combinatorial procedures can easily result in millions of generated activities (Janssen, 2009; Dogliotti *et al.*, 2005; Dogliotti *et al.*, 2003). Using a bio-economic farm model to optimally allocate such a large number of activities to the available farm land is impossible from a computational point of view and the DEA approach offers a solution here. It should be noted though that in such cases, the full set of generated activities needs to be divided to smaller subsets in order to allow a feasible solution in the DEA model. The DEA approach could then be used to filter out the inferior activities of each subset. Next, the superior activities selected from each subset of step 1 are merged to one set and re-evaluated with DEA as a whole. The resulting set of superior activities coincide with the set of activities that would have been selected if it was possible to evaluate with DEA the initial full set of generated activities.

Offering the set of activities selected with DEA instead of the full set of activities reduces the time needed to calculate the trade-offs between different inputs and outputs of the bio-economic model. However, the DEA filtering process can be a computational intensive procedure and can increase the total computation time (time for screening the full set of activities + time for creating the trade offs). The pre-selection of activities using DEA is beneficial for reducing the total computation time especially in cases where the optimization model is used multiple times for calculating trade-offs between different inputs and outputs or in cases where large number of activities have been generated. In case of a low number of simulations and a low number of generated activities the proposed DEA method might even increase the total computational time. However, the benefits of using the DEA method for interpreting the results remain.

The main purpose of the experiment on fertilization options in arable farming in Flevoland is to demonstrate the proposed DEA method. For that reason, the FSSIM-MP model was kept simple (i.e. linear gross margin optimization model). The results of FSSIM-MP suggest that the alternative management with cattle slurry is not selected because of the higher labour requirements. However, this might not be the only reason that in reality farmers do not fully adopt this alternative management. Uncertainty about weather conditions and availability of nutrients when those are needed by the crop might also play an important role. Additionally, management with cattle slurry increases transaction costs which are not accounted in the simple version of FSSIM-MP. In a more comprehensive version of the model, it would be possible to include and analyze such issues.

Combinatorial procedures and filtering rules are useful tools for identifying and generating alternative activities in different kind of future-oriented land use studies. One type of such studies is based on the use of bio-economic farm models and has been illustrated in this paper. Other future-oriented land use studies are using partial or general equilibrium models focusing either on supply and demand relationships in the agricultural (Heckelei and Britz, 2001) or all economic sectors (Hertel, 1997; Van Tongeren, 2001; Rosegrant *et al*., 2008) . These models generally use trend functions for technological change, but could also include alternative activities. Another type of future land use studies requiring alternative activities explores future land use options and address “what if” questions at EU level (Rabbinge and van Latesteijn, 1992) and farm level (Ten Berge *et al*., 2000). Studies assessing the impact of and the adaptation to climate change also request the consideration of alternative activities (Lehtonen *et al*., 2006; Henseler *et al*., 2009). The proposed method, based on DEA, for selecting superior alternative agricultural activities can be a useful complementary component for combinatorial tools that aim to identify and quantify alternative activities. The DEA method decreases the number of selected activities to operational levels that can be easier analyzed from scientists and policy makers.

5.6. Acknowledgments

The work presented in this publication was partly funded by the SEAMLESS integrated project, EU 6th Framework Programme for Research Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2) and partly funded by the project for Assessing the adaptive capacity of Agriculture in the Netherlands to the impacts of climate change under different market and policy scenarios (AgriAdapt-NL), funded by the Dutch research programme Climate Changes Spatial Planning.

5.7. References

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D. (2007). Farm management indicators and farm typologies as a basis for assessment in a changing policy environment. Journal of Environmental Management, 82: 352-362.

Barbier, B. and Bergeron, G. (1999). Impact of policy intervention on land management in Honduras: result of a bioeconomic model. Agricultural Systems, 60: 1-16.

Berentsen, P. B. M., and Tiessink, M. (2003). Potential Effects of Accumulating Environmental Policies on Dutch Dairy Farms. Journal of Dairy Science, 86: 1019-1028.

Borkowski, N., Hecker, J.M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G. and Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 1: 33-34.

Bouyssou, D. (1999). Using DEA as a tool for MCDM: some remarks. Journal of the Operational Research Society, 50: 974-978.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2: 429-444.

Cooper, W. W., Seiford L. M. and Zhu J. (2004). Data Envelopment Analysis: Models and Interpretations, Chapter 1, 1-39, in W.W. Cooper, L.M. Seiford and J. Zhu, eds, Handbook on Data Envelopment Analysis, Kluwer Academic Publisher, Boston.

Cooper, W. W. Seiford, L. M. and Tone K. (2007). Data Envelopment Analysis: A comprehensive text with models, application, references and DEA-solver software. Springer science and business media.

Dogliotti, S., Rossing, W.A.H., and Van Ittersum, M.K. (2003). ROTAT, a tool for systematically generating crop rotations. European Journal of Agronomy, 19: 239-250.

Dogliotti, S., Van Ittersum, M.K., and Rossing W.A.H. (2005). A method for exploring sustainable development options at farm scale: A case study for vegetable farms in South Uruguay. Agricultural Systems, 86: 29-51.

EC (2005). Impact Assessment Guidelines. European Commission, Brussels, pp 99.

Habekotté, B. (1994). TCG_CROP, een model voor berekening van produktie-en milieuvariabelen van verschillende gewassen: ontwikkeld ten behoeve van het project

‘Introductie Geïntegreerde Akkerbouw’. DLO- Centrum voor Agrobiologisch onderzooekm Wageningen university. Simulation reports CABO-TT nr. 35.

Heckelei, T., and Britz, W. (2001). Concept and explorative application of an EU-wide regional agricultural sector (CAPRI-projekt), in: Heckelei, T., Witzke, H. P., Henrichsmeyer, W., (Eds). 65th EAAE Seminar Vauk Verlag, Kiel, Germany: 281-290.

Hengsdijk, H., Van Ittersum M. K. (2003). Formalizing agro-ecological engineering for future-oriented land use studies. European Journal of Agronomy, 19: 549-562.

Henseler, M., Wirsig, A., Herrmann, S., Krimly, T., Dabbert, S. (2009). Modeling the impact of global change on regional agricultural land use through an activity-based non-linear programming approach. Agricultural Systems, 100: 31-42.

Hertel, T. W. (1997). In: T.W. Hertel, Editor, Global Trade Analysis: Modeling and Applications, Cambridge University Press, Cambridge, New York.

Janssen, S. and Van Ittersum, M. K. (2007). Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems, 94: 622-636.

Janssen, S. (2009). Managing the Hydra in integration: developing an integrated assessment tool for agricultural systems. PhD thesis (http://edepot.wur.nl/10616).

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Meuter, E., Andersen, E., Belhouchette, H., Blanco, M., Borkowski, N., Heckelei, T., Hecker, M., Li H., Oude Lansink, A., Stokstad, G., Thorne, P., Van Keulen, H., and Van Ittersum, M. K. (2010). A generic bio economic farm model for environmental and economic assessment of agricultural systems. Under review.

Kanellopoulos, A., Berentsen, P., Heckelei, T., Van Ittersum, M., Oude Lansink, A. (2010). Assessing the Forecasting Performance of a generic bio-economic farm model calibrated with two different PMP variants. Journal of Agricultural Economics, 61: 274-279.

Lehtonen, H., Peltola, J. and Sinkkonen, M. (2006). Co-effects of climate policy and agricultural policy on regional agricultural viability in Finland. Agricultural Systems, 88: 472-493.

Louhichi, K., Kanellopoulos, A., Janssen, J., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Oude Lansink, A., and Van Ittersum, M.K. (2010). FSSIM, a Bio-Economic Farm Model for Simulating the Response of EU Farming Systems to Agricultural and Environmental Policies. Accepted in Agricultural Systems.

Louhichi, K., Janssen, S., Kanellopoulos, A., Li, H., Borkowski, N., Flichman, G., Hengsdijk, H., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I.N., Rizzoli, A.E., Huber, D., Heckelei, T., van Ittersum, M.K. (2009). A generic Farming System Simulator In: Brouwer, F., van Ittersum, M.K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer, Dordrecht.

Olesen, J. E. and Bindi, M. (2002). Consequences of climate change for European agricultural productivity, land use and policy. European Journal of Agronomy, 16: 239-262.

Rabbinge, R. and Van Latesteijn, H.C. (1992). Long-term options for land use in the European community. Agricultural Systems, 40: 195-210.

Rosegrant W. M., Ringler, C., Msangi, S., Sulser, B. T., Zhu, T., and Cline. A. S. (2008). International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description. Project report (http://www.ifpri.org/sites/default/files/ publications/impactwater.pdf)

Stewart, T. (1996). Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis. Journal of Operational Research Society, 47: 654-665.

Ten Berge, H. F. M., van Ittersum, M. K., Rossing, W. A. H., G. van de Ven, W. J., Schans, J. (2000). Farming options for The Netherlands explored by multi-objective modelling. European Journal of Agronomy, 13: 263-277.

Van der Burgt, G. J. H. M., Oomen, G. J. M., Habets A. S. J., and Rossing W. A. H. (2006). The NDICEA model, a tool to improve nitrogen use efficiency in cropping systems. Nutrient Cycling in Agroecosystems, 74: 275-294.

Van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E. and Wolf, J. (2008). Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). Agricultural Systems, 96, 150-165.

Van Tongeren, F., Van Meijl, H., and Surry, Y. (2001) Global models applied to agricultural and trade policies: a review and assessment. Agricultural Economics, 26: 149–172.

Wossink, G. A. A., de Koeijer, T. J. and Renkema, J. A. (1992). Environmental-economic policy assessment: A farm economic approach. Agricultural Systems, 39: 421-438.

Zander, P., Borkowski, N., Hecker, J. M., Uthes, S., Stokstad, G., Rørstad P. Kr., and Bellocchi, G. (2009). Conceptual Approach to Identify and Assess Current Activities. P.D 3.3.9. SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

Appendix

LP models for Data Envelopment Analysis

The multiple-input multiple-output DEA model involves multiple runs (one for each DMU). In each model run, the inputs and outputs of an activity are compared to the inputs and outputs of all other activities and an efficiency score (θ) is calculated. Here, the input oriented BCC model described in Banker *et al*. (1984) is presented. Throughout, the assumption of decreasing returns to scale is made. The LP model of a DEA problem takes the following form:

$$\begin{aligned} &\min_{\theta,\lambda_k} \ \theta, \quad s.t.: \\ &\qquad \sum_k \lambda_k \cdot Y_{k,j} - Y_j^0 \geq 0 \quad \forall j, \\ &\qquad \theta \cdot X_i^0 - \sum_k \lambda_k \cdot X_{k,i} \geq 0 \ \ \forall i, \\ &\qquad \sum_k \lambda_k \leq 1, \qquad \lambda_k \geq 0 \end{aligned} \tag{2}$$

where θ is the performance (“efficiency”) score of the evaluated activity, λ_K is the weight to activity *k* and it is a decision variable , $Y_{k,j}$ is the output *j* of activity *k*, Y^0_j is the output *j* of the evaluated activity, X_{ki} is the input *i* of activity *k*, and X^0_k is the input *i* of the evaluated activity. The two phase model described in (3) is equivalent to the model in (2) but tests also the existence of weak efficiency. An equivalent one phase model exists that uses an “non-Archimedean” element (a positive number smaller than any other positive number) to minimize θ and maximize the slacks simultaneously. However, Charnes *et al*. (1994) pg 76-79, show that the choice of the value of this number is data specific and can affect the results of a DEA model.

$$
\begin{aligned}
&Phase\,1: \quad \min_{\theta,\lambda_k,s_j^+,s_i^-} \quad \theta \\
&Phase\,2: \quad \min_{\lambda_k,s_j^+,s_i^-} \left\{ -\sum_j s_j^+ - \sum_i s_i^- \right\}, \\
&subject\ to: \\
&\qquad \sum_k \lambda_k \cdot Y_{k,j} - Y_j^0 - s_j^+ = 0 \quad \forall j, \\
&\qquad \theta \cdot X_i^0 - \sum_k \lambda_k \cdot X_{k,i} - s_i^- = 0 \ \ \forall i, \\
&\qquad \sum_k \lambda_k \le 1, \\
&\qquad \lambda_k \ge 0 \ \ \forall k, \quad s_j^+, s_i^- \ge 0 \ \ \forall i, j
\end{aligned}
\tag{3}
$$

Where s^+_j, s^-_i are positive slack variables, and Phase 2 replaces the variable θ with the fixed value of minimum $\theta = \theta^*$ of Phase 1. A DMU is weakly efficient when $\theta = 1$ in Phase 1 of model described in (3) and at least one of the slack variables are greater than 0 in Phase 2.

Chapter 6

6. General Discussion

Argyris Kanellopoulos[1,2]

[1] *Business Economics, Wageningen University, Hollandseweg 1, 6706 KL Wageningen, The Netherlands*
[2] *Plant Production Systems, Wageningen University, P.O. Box 430, 6700 AK Wageningen, The Netherlands*

This chapter broadens the discussion of preceding chapters to the overall achievements and to issues that are related to major methodological choices and assumptions. Moreover, some more general implications of the results of this PhD thesis are analyzed. In Section 1, the overall achievement of this PhD thesis is discussed. In Section 2, major discussion points related to methodological issues are raised. In Section 3, some points derived from the empirical results of this PhD thesis are analyzed. Section 4 concludes, while in Section 5 ideas for further research are presented.

6.1. Contribution and achievements

The overall objective of this PhD thesis was to develop and evaluate generic bio-economic farm models that can be used for integrated assessment of agricultural and environmental policies at multiple levels and different biophysical and socioeconomic conditions.

First, we looked into the modelling requirements for developing a generic and re-usable bio-economic model for integrated policy assessment. A farm model was developed that can be readily adapted to simulate arable, livestock and mixed farming systems located in various socio-economic, political and physical environments (i.e. different regions, soil types, climatic zones). Most resource constraints related to arable farm types are relevant also for livestock and mixed farm types and they are always included in the model specification. The constraints and the data inputs have been separated in different modules (e.g. arable, livestock, calibration) so that constraints related to different kinds of farming systems can be switched on and off easily in the model's code.

For non-modellers, switching on and off of constraints can be done in the SEALMESS-IF graphical user interface. Using the SEAMLESS integrated database (Janssen *et al.*, 2009) enables uniform reproduction of data inputs for different farm types across the EU. Both current and alternative agricultural activities are defined as crop rotations and/or herd structures capturing possible spatial and temporal interactions between different crops and livestock. The calculation of a number of environmental indicators is also enabled. The reusability of the model was demonstrated in Chapter 2 and it is confirmed by the significant number of applications that have already been using it. Louhichi *et al.*, (2008) used the model with detailed information available at regional level to assess the consequences of the nitrate directive in Midi-Pyrennes (France). Kanellopoulos *et al.*, (2009) compared the effects of abolishing the set-aside policy in different regions through out Europe. The effects of CAP reform to the European dairy sector were revealed by Louhichi *et al.* (2009^{b}). Majewski *et al.* (2009) investigated scenarios of bio-fuel promoting policies in Poland. The effects of the 2003 CAP reform to water quality in a Scottish region were assessed in Mouratiadou *et al.* (2009). Price-supply elasticities calculated by the proposed bio-economic model for a representative sample of regions were used in Pérez Domínguez *et al.* (2009) for extrapolating the production structure across EU. The model has been also used to simulate farming systems of developing countries (Traoré *et al.*, 2009).

Second, different options for calibration and methods to recover unknown parameters underlying the farmer's decision making have been explored. We proposed alternative calibration procedures that improve the existing PMP methodology (Chapter 3 and 4). A method based on Maximum Entropy estimation for quantification of the farmer's risk attitude was also proposed. We used "back-casting" simulations (i.e. ex-post experiments) to assess the forecasting performance of the model calibrated with different methods. In these simulations, the bio-economic farm model is calibrated with historical data of a particular base year and it is used to forecast effects of policies and price changes on the following historical years. The capacity of the model to reproduce changes in activity levels of the past is assessed.

The proposed calibration methods involve a number of underlying assumptions that better comply with the actual decision making of farmers. The values of limited resources were raised to the average gross margin instead of the gross margin of the least preferable activity; increasing marginal costs were assumed for all activities and complementarity, substitution and risk aversion were quantified. Despite the improved forecasting performance, it was concluded that there is no general calibration method appropriate for all cases. The data availability and the aim of the study appear to be the most important factors that determine the best calibration option for a specific case. For example, the Röhm & Dabbert (2003) PMP variant can be used if there is available information on observed levels of crop-management combinations in order to account for different elasticities between managements and crops. The PMP variant proposed in Chapter 3 can be used to exploit available information on own price elasticities or information on historical data that allows for designing an ex-post experiment. The Maximum Entropy estimation approach proposed in Chapter 4 can be used to exploit panel data on activity levels and expert's knowledge on agro-management to estimate explicitly complementarity, substitution and risk aversion. The standard PMP approach (Howitt,

1995) can be a solution in cases where such information is not available. To give enough options to model users and policy makers, we implemented a number of different calibration procedures in the developed bio-economic farm model. The related equations and constraints of different calibration procedures are included in separate modules so that they can be switched on and off easily.

Finally, we investigated approaches for identifying a set of alternative activities that could be used for integrated assessment of future scenarios. Combinatorial methods (Dogliotti *et al.*, 2003; Janssen, 2009) can be used to generate all possible alternative agricultural activities in a uniform and reproducible way for a large number of farm types with relatively limited information. The limitation of the method is that the number of alternative activities that is generated in combinatorial approaches can easily explode. Only a fraction of the generated activities are relevant from a policy point of view. We proposed a generic approach based on Data Envelopment Analysis (DEA) for reducing the number of interesting alternative activities to a level that can easily be applied in bio-economic farm models.

6.2. Methodological issues

A number of critical decisions and methodological choices were made during this thesis. Creating a farm model that can be integrated in a model framework such as SEAMLESS-IF and linked to other models in a model chain is a challenge by itself and requires a number of decisions with respect to the methods. Some of these decisions (e.g. the generic structure of the model, selecting static versus a dynamic approach, selecting a positive versus a normative approach, simulating average farms instead of individual farms) were made for the sake of the framework so that all models involved are compatible with respect to level of detail, inputs and outputs. However, this does not always come without a cost. Some critical decisions that were made during this thesis are discussed below.

6.2.1. Generic modelling and model re-usability

One of the objectives of this PhD thesis was to develop a bio-economic farm model that could be used to simulate different farming systems in various socio-economic, political and environmental conditions. The model is used to calculate price supply relationships at EU level where data availability is limited. At the same time the model had to be capable of exploiting more detailed information available at regional or even farm level for dedicated applications in specific regions.

By definition models are abstractions of reality and consequently they cannot include all factors that influence a bio-physical process or decision making of farmers. For that reason the researcher is always challenged to assess the level of detail at which each different factor should be modelled to address the underlying question. A different research question can result in a completely different modelling approach. A comprehensive model that can address adequately all possible research questions is probably impossible. However, a model can be designed to be flexible enough and easily adaptable to address different questions under different conditions. In Chapter 2, it was attempted to design a bio-economic farm model that can be transformed easily to account for different conditions and environments and simulate different farms across the EU. The attempt to create a generic and flexible farm model can easily result in complicated programs with components irrelevant for the targeted analysis. Many modules of FSSIM are irrelevant for answering specific policy questions and they are switched off before simulations. However, these modules are still part of the model and thus increasing complexity. There is a clear risk for the re-usability of such models and therefore good documentation and training capacity is absolutely necessary. Technical details of the structure of the proposed model and the explanation of the model's equations are extensively reported in Louhichi *et al.* (2009[a]) while short explanatory text is also included within the code of the model. To promote re-usability and accessibility of the model and the model's results, the model should be publicly available (www.seamlessassociation.org).

6.2.2. Positive modelling with limited datasets

Positive modelling approaches use historical data and attempt to recover the unknown parameters of the model related to production structure (i.e. non-linear costs due to diseconomy of scale and land heterogeneity) and risk aversion to explain the underlying behaviour of the farmer. Usually, the objective is to calculate farm responses and try to understand them. On the other side, normative approaches pre-suppose the farmer's objective and use existing knowledge on the production process involved and on the socio-economic and bio-physical environment and try to find the most satisfying (optimal) solutions and alternatives to the problem of resource management and allocation (Flichman and Jacquet, 2003). In Chapter 2, 3 and 4 of this thesis, a positive approach was used in all farm level simulations. The limited datasets did not allow for using traditional econometric procedures for recovering the unknown parameters. Positive Mathematical Programming and Maximum Entropy estimation were used instead. Using these methods to recover the unknown parameters from a limited dataset involves using prior-information (expert knowledge) and imposing a number of assumptions (Heckelei, 2002), which are not always easy to justify from an economic point of view. Changing the prior information and/or the underlying assumptions might have major implications for the recovered parameters. It is important that the detail of the model specification is such that the impact of calibration is minimized. This is very difficult in cases of higher level analysis (like the one aimed with SEAMLESS) where including region specific constraints and more detailed information is very difficult if not impossible. An iterative process of model development and evaluation (testing) through ex-post experiments was used to improve the model's specification and consequently restrict the effect of calibration. Another important limitation of positive approaches is that only parameters of activities that have been used in the past or of activities that are currently used in the region can be recovered. It is difficult and questionable to include alternative activities that are not currently used in a certain region but might be relevant for future scenarios. This is mainly because usually there is lack of expert's knowledge and data on the performance (i.e. input requirements, outputs) of alternative activities in a specific region. Unknown parameters of the model (non-linear costs, risk aversion) cannot be easily recovered. For long term explorations, where large uncertainty is involved and major technological and environmental changes are expected, normative approaches might be more suitable. For that reason, the proposed bio-economic model can be easily transformed to a normative model by switching off the calibration component (see application of Chapter 5). The features and assumptions of the different calibration procedures demonstrated and evaluated in this thesis are presented in Table 1.

The advantage of using the standard PMP approach for calibration rather than using a normative approach is the guarantee of exact calibration. The extended variant of PMP tries to overcome some important limitations of the standard PMP approach. First it raises the value of land to the average gross margin (instead of the gross margin of the non-preferable activity in standard PMP) and estimates non-linear costs also for the non-preferable activity. The extended variant of PMP proposed in this thesis improves the forecasting performance of the model while exact calibration is guaranteed. The Maximum Entropy estimation method exploits available information more efficiently since aggregated information on management available in EU level databases can be included in the maximum entropy estimation process and expert knowledge on management, risk aversion and expected outputs can be included as prior information. Like the extended variant of PMP, the value of land is raised to the average gross margin. The risk aversion coefficient is recovered and complementarity and substitution between activities is estimated. However, in estimation procedures involving multiple years of data, exact calibration should not be expected.

Table 1: Summary of features and assumptions of the different procedures used in the thesis.

	Normative (LP)	Standard PMP	Extended PMP	Maximum entropy
Demonstrated in Chapter	Ch. 5	Ch. 3	Ch. 2 & 3	Ch. 4
Reproduction of base year data	Poor*	Exact	Exact	Good
Forecasting performance	Not tested	Not that good	Good	Good
Exploiting historical data set	No	One year	One year	Multiple years
Use of expert knowledge	No	No	No	Yes
Value of limited resource	-	Gross margin of least profitable activity	Average gross margin	Average gross margin
Additional non-linear costs	Not included	For all activities except the least profitable	For all activities	For all activities
Substitution & complementarity between activities	Not estimated	Not estimated	Not estimated	Estimated
Risk aversion	Not estimated	Not estimated	Not estimated	Estimated

* In general, it is very difficult to reproduce base year data adequately using a normative model because it requires information on non-linearities involved in the decision process (e.g., production structure, risk aversion).

6.2.3. Modelling an average farm versus individual farms

In all farm level model applications presented in this thesis, the average resource endowments and observed production plans of farms belonging to a certain farm type were selected as representative values of farms that belong to that farm type. Simulating the average farm of a certain farm type using a calibrated model ensures that all important crop products that are produced by farms of a specific farm type will be part of the simulated production plan. This is very important for analysis that requires full representation of agricultural production to determine equilibrium between supply and demand. For example, one of the aims of the bio-economic farm model in SEAMLESS, is to calculate price-supply elasticities for the farm types of a representative sample of regions across the EU which are then extrapolated (Pérez Domínguez *et al.*, 2009) and used in a partial equilibrium model for EU level analysis. The calculation of price-supply elasticities for as many products as possible is vital for this type of analysis. Using individual farms for representing the farm types makes it more difficult to ensure adequate representation of all observed activities of farms of a specific farm type in the simulated production plans. However, simulating the average farm has also important drawbacks. An average farm and an average farmer do not actually exist and consequently, an average activity pattern also does not exist. The activity pattern of the average farm is much more diversified than that of individual farms. Reproducing such a cropping pattern using an LP model would require a large number of binding constraints. It is possible that such constraints do not even exist in reality and consequently they are difficult to define (e.g. rotational constraints of an "average" production plan). Calibration of the LP model is necessary for reproducing the observed activity levels and often calibration will dominate the simulations. It is possible that the impact of calibration on the results of the model would be reduced substantially if a number of individual farms were simulated instead of a single average farm. However, this would also have increased the computational requirements and individual farm data would have to be available which is usually not the case (individual farm data are usually treated as confidential information not available for research).

6.2.4. Evaluating forecasts

Assessing the capacity of a model to predict the future is difficult if not impossible simply because future events are not yet known. In this thesis, ex-post experiments were employed to assess the forecasting capacity of the model. The model was calibrated with data of year(s) in the past and used to simulate changes that occurred in the past. The capacity of the model to reproduce the farms historical activity levels can be used to assess the quality of the forecasts. One of the main objectives of the bio-economic farm model proposed in this thesis is to calculate price-supply elasticities of different agricultural products. We are interested mainly in relative changes in quantities of products rather than the absolute production. For this reason, in the ex-post experiments presented in this thesis, we focused on comparing simulated and actual activity patterns. The results of these comparisons were used as a measure of performance. A good reproduction of activity patterns of base year and forecasted year results in good reproduction of relative changes in supply and in the values of interesting economic (e.g. farm income) and environmental (e.g. nitrogen leaching) indicators. Assessing the capacity of the model to reproduce the absolute values of indicators is more complicated because it involves uncertainty for the quality of data that is used to quantify those indicators.

Using the model assessed in ex-post experiments to simulate future events does not guarantee good forecasting performance in all possible cases. The model might need to be changed severely to include issues related to technological innovations and changes in the institutional, economic, and physical environment that become important in the simulated period under a certain scenario. In such cases, the forecasting capacity of the model is questionable again. To improve confidence in the forecasting performance of the model it is important to design appropriate ex-post experiments with exogenous conditions reflecting as much as possible the scenario. Similarities between the ex-post experiment and the actual forecasting exercise should be found in terms of the socio-economic, political and bio-physical environment. Results from such ex-post experiments can be used to decide on an appropriate calibration procedure by comparing the forecasting performances of the different methods.

6.2.5. Dynamic decision making in farming

The farmer’s decision making is a dynamic process of resource allocation. In general, by the time more information becomes available decisions are adapted to maximize utility. This is how farmers deal with investments, risk and uncertainty. A number of different approaches have been proposed to deal with dynamics and inter-temporal decisions involved in farming (Pandey and Hardaker,1995; Bardier and Bergeron, 1999; Wallace and Moss, 2002; Acs *et al.*, 2007). In general, a dynamic farm model is more complex and requires information which is not always available at EU and global level. The farm model proposed in this thesis attempts to capture some of these interactions (e.g. specifying activities as crop rotations instead of single crops) in a static way to align the data requirements with the data availability in EU level databases and a simple survey on agricultural management (Borkowski *et al.*, 2007; Zander *et al.*, 2009). Investment decisions have not been taken into account and for that reason it is important to notice that the model can only be used for relatively short term forecasts where major investment decisions or changes to the fixed costs are not expected.

6.2.6. Accounting for alternative activities

In Chapter 5 of this thesis an attempt was made to demonstrate how alternative activities could be incorporated in bio-economic farm modelling. The presented exercise focused on nutrient management options of arable farms. This exercise was selected because it is representative of a wider group of alternative activities related to technological innovations in management and alternative production functions. This type of innovations in farming is related to improvements in technology that could lead to new available combinations of resources (i.e. labour, land and capital) for crop and animal production that can result in more beneficial activities from an economic, social or environmental point of view. Another type of innovation that is also covered with the simple exercise presented in Chapter 5 is alternative rotations i.e. alternative rotational decisions because of changes in the climatic and socio-economic conditions. Alternative herd structure can be taken into account in a very similar way.

Innovations related to changes in farm’s organization and farmer’s decision making have not been considered in this thesis. This type of innovations involves changes in the organization of the farm so that constraints related to available farm resources and rotational constraints become less restrictive. An example of this kind of changes is the cooperation of arable and dairy farms in a single decision making unit spreading the crop rotation and feed production over the land of all involved individual farms. This gives them new possibilities for more intensive rotations or rotations with less environmental impact (while maintaining the same productivity), alternative nutrient management and sufficient feed production. Obviously this kind of innovations can become important for the decision making and should be taken into account in future land use studies. It might be possible to use a farm model to investigate some of these innovations in a simple way by changing the definition of a farm according to the assessed organization change. However, a more comprehensive analysis would require a regional model where available resources and constraints at higher level can be included and where prices of limited resources (e.g. land, labour) are determined.

6.3. Results

6.3.1. Model applications

The applications of the model in Chapter 3 and 4 of the thesis focused strictly on assessing the capacity of the calibrated farm model to simulate observed cropping patterns. The Percent Absolute Deviation (PAD) was used to measure the deviation of simulated activity levels from observed historical data. The minimum and maximum PAD values achieved in the ex-post experiments presented in this thesis for all assessed calibration methods are presented in Table 2. The achieved PAD values of calibration methods proposed in this thesis (i.e. extended PMP and Maximum Entropy estimation) outperform the standard PMP method. Hazel and Norton (1986) suggest that models that reproduce the base year activity pattern with PAD values lower than 15% can be used for forecasting. The ex-post experiments of the model calibrated with the proposed methods resulted in maximum PAD values only marginally above 15% even for the forecasting year (not the base year). We can conclude that the forecasting capacity of the model calibrated with the proposed methods is acceptable.

Table 2: Minimum and Maximum values of the Percent Absolute Deviation (%) achieved in forecasts of ex-post experiments of the thesis per calibration method and region.

	Standard PMP		Extended PMP		Maximum Entropy	
Region	Min.	Max.	Min.	Max.	Min.	Max.
Flevoland	63	76	3	22	14	19
Midi-Pyrenees	20	26	14	23	11	15

6.3.2. Interpretation of recovered parameters

The recovered parameters in all tested cases of the calibration and estimation procedures proposed in this thesis have an economic justification as they are related to increasing variable costs per unit of production because of inadequate machinery and management capacity, decreasing yields due to land heterogeneity, and risk aversion (Howitt, 1995). However, it is important to notice that any possible model misspecification (e.g. omitting to include farm specific constraints, ignore heterogeneity of land, simplifications in the decision making) is also captured in the recovered parameters. The feature of exact calibration of PMP and the use of prior information in ME can dominate the estimation procedure and result in outcomes of model simulations that are very close to the observed situation. To test for the validity of the model, we used ex-post experiments which provide more information about the forecasting capacity of the model. Nevertheless, the results are usually case specific and they cannot guarantee good model performance in all cases. Using calibration procedures to improve the model's forecasts does not rule out the need for careful model development; on the contrary, it requires additional effort to identify possible misspecifications which will not appear directly in the results of the analysis because of calibration.

6.4. Conclusions

The major conclusions from the thesis are:

1. The Farm System SIMulator is a flexible bio-economic model that can be used for simulating different farming systems, under different bio-physical and socio-economic conditions, for a variety of policy questions. This was achieved by: (i) separating model and data and creating a consistent European database for farm types, their locations and production activities, (ii) designing the model in a modular way, that allows switching on and off modules, constraints or calibration methods, (iii) providing adequate documentation, and (iv) ensuring public availability. The re-usability of the model is demonstrated in this thesis but is also confirmed by applications presented in other recent studies.
2. The PMP variant proposed in this thesis raises the value of limiting resource to the average gross margin and assumes decreasing marginal gross margin also for the least preferable activity. The proposed PMP variant improved the forecasting performance of the model compared to the standard PMP approach in all tested cases.
3. Maximum Entropy estimation exploits expert's knowledge and panel data available in EU level databases more efficiently and requires less arbitrary assumptions than Positive Mathematical Programming for calibrating bio-economic farm models.
4. Evaluating the forecasting capacity of bio-economic farm models is a complicated task mainly because it refers to the unknown future and because often bio-economic models use simulated data to account for price and yield trends. Ex-post experiments, in which the model is calibrated with historical data of a particular base year and used to forecast policies and price changes of the following historical years, are useful for assessing the forecasting performance of bio-economic models.
5. Combinatorial procedures and filtering rules are useful tools for identifying and generating alternative activities in different kinds of future-oriented land use studies. The DEA method proposed in this thesis, for selecting superior alternative agricultural activities, reduced the number of alternative agricultural activities generated by

existing combinatorial procedures to a level that can easily be applied in bio-economic farm models and analyzed by scientists and policy makers.

6.5. Recommendations for future research

The research presented in this thesis leads to a number of new, interesting and relevant research topics. First, positive modelling approaches use existing information and activity observations to recover unknown parameters. Calibration or estimation of a model that includes activities that are not observed in the reference year (i.e. alternative, not currently observed activity) has not been accomplished so far. This restricts the use of positive models to short term simulations where no major changes in bio-physical, socio-economic, technological and institutional environment are expected. In the current literature there is a gap on calibrating bio-economic models in which alternative activities are included. The economic justification of the recovered unknown parameters of the model is related to limited managerial and machinery capacity, land heterogeneity and risk aversion. A farmer confronted with the decision of adopting or not an alternative activity will have to make estimations and assumptions about additional non-observed costs, yield losses due to land heterogeneity and price-yield variation of the alternative activity by seeking similarities in agro-management between the alternative activity and current activities. This kind of information on the decision making of the farmer could be used to recover unknown parameters for the alternative activities. The average non-linear costs of the current production plan or the non-linear costs of current activities that have similar inputs, outputs and agro-management requirements can be used to parameterize the alternative activity. Only simulations of short term predictions should be targeted with the model since after a number of years it is expected that more information will be available to the farmer and the decision making will change.

Second, many of the modelling decisions for developing FSSIM were made because of data limitations. This might have implications for the quality of the results of the model. To assess the added value of creating a more detailed model which would include dynamics, structural change and multiple objectives, different modelling formulations have to be created and compared to each other. Detailed regional, or farm specific databases can be used for creating such experiments of comparisons between different modelling formulations. The results of the comparisons could determine the appropriate level of detail of a bio-economic farm model that aims at integrated assessment of environmental and agricultural policies. It will also provide information about data needs for comprehensive bio-economic analysis.

Third, databases at EU and global level do not include enough information for developing detailed bio-economic models. More detailed information on agro-management (e.g. disaggregated input levels, timing, crop rotations) that is used currently in existing farming systems would have contributed significantly to the level of detail of bio-economic farm models applicable across EU.

Finally, object-oriented programming is a programming approach that enhances re-usability and a generic structure of programs. The concept of object-oriented programming could also be used for developing mathematical programming bio-economic farm models. In general, a farm can be seen as an object with available resources that are allocated to activities which are also seen as objects with multiple inputs and outputs. The matrix of input-output coefficients, the vector of available resources and the objective function are then created using the available farm resources and the inputs and outputs of available activities. Open source software like JAVA can be used for this purpose. Available software libraries for solvers written in JAVA or R can then be used for the optimization.

6.6. References

Acs, S., Berentsen, P. B. M., Huirne, R. B. M. (2007). Conversion to organic arable farming in The Netherlands: A dynamic linear programming analysis. Agricultural Systems, 94: 405-415.

Barbier, B. and Bergeron, G. (1999). Impact of policy interventions on land management in Honduras: results of a bioeconomic model. Agricultural Systems, 60: 1-16.

Borkowski, N., Hecker, J.M., Louhichi, K., Blanco Fonseca, M., Janssen, S., van Ittersum, M.K., Stokstad, G., Zander, P. (2007). Surveying crop management data for bio-economic farm models. In: Hatfield, J., Donatelli, M., Rizzoli, A. (Eds.), Farming Systems Design 2007: An international symposium on Methodologies for Integrated Analysis of Farm Production Systems, Catania, Sicily, Italy, 1: 33-34.

Flichman, G., Jacquet, F. (2003). Couplage des modèles agronomiques et économiques: intérêt pour l'analyse des politiques. Cahiers d'économie et sociologie rurales, 67: 1-20.

Hazell, P., and Norton, R. (1986). Mathematical Programming for economic analysis in agriculture. Macmillan, New York.

Heckelei, T. (2002). Calibration and Estimation of Programming Models for Agricultural Supply Analysis. Habilitation Thesis, University of Bonn, Germany (http://www.ilr1.uni- bonn.de/agpo/ staff/heckelei/heckelei_hab.pdf).

Howitt, R.E. (1995). Positive Mathematical Programming. American Journal of Agricultural Economics, 77: 329–342.

Janssen, S., Andersen, E., Athanasiadis, I. and Van Ittersum, M. (2009). A database for integrated assessment of European agricultural systems. Environmental Science & Policy, 12: 573-587.

Kanellopoulos, A., Hengsdijk, H., Janssen S., Van Ittersum, M.K. (2009). Ex-ante assessment of the abolishment of the EU set aside policy: Results from a bio-economic farm analysis. In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP 2009). Egmond aan Zee, The Netherlands, 10-12 March 2009. Wageningen University and Research Centre, Wageningen, pp. 478-479.

Louhichi K., Belhouchette H., Flichman G., Therond O. and Wery J. (2008). Impact assessment of the 2003 CAP reform and the Nitrate Directive on the arable farming system in the Midi-Pyrénées Region: Bio-Economic Modelling at field, farm and regional Levels. 109th Seminar, November 20-21, 2008, Viterbo, Italy.

Louhichi, K., Flichman, G., Blanco Fonseca, M. (2009). A generic template for FSSIM, SEAMLESS Report no. 37, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, pp. 80 (www.seamlessassociation.org).

Louhichi, K., Hengsdijk, H., Janssen, S., Bigot, G., Perret, E. (2009). EU dairy farming in the face of change: An exploration using a bio-economic farm model. In: Proceedings of the conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Wageningen University and Research Center: 262-263.

Louhichi, K., Janssen, S., Kanellopoulos, A., Li, H., Borkowski, N., Flichman, G., Hengsdijk, H., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I.N., Rizzoli, A.E., Huber, D., Heckelei, T., van Ittersum, M.K. (2010). A Generic Farming System Simulator In: Brouwer, F., van Ittersum, M.K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer, Dordrecht, pp. 109-132.

Majewski, E., Was, A., Belhouchette, H., Louhichi, K., Mouratiadou, I. (2009). Impact Assessment of policy changes on the arable sector using the FSSIM model: Case study of the Zachodniopomorskie NUTS region. In: Proceedings of the conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Wageningen University and Research Center: 484-485.

Mouratiadou, I., Topp, C.F.E., Russell, G. (2009). Investigating the economic and water quality effects of the 2003 CAP reform on arable cropping systems: A Scottish case study. In: Proceedings of the conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Wageningen University and Research Center: 488-489.

Pandey, S. and Hardaker, J. B. (1995). The role of modelling in the quest for sustainable farming systems. Agricultural Systems, 47: 439-450.

Pérez Domínguez, I., Bezlepkina, I., Heckelei, T., Romstad, E., Oude Lansink, A., Kanellopoulos, A. (2009). Capturing market impacts of farm level policies: a statistical extrapolation approach using biophysical characteristics and farm resources. Environmental Science & Policy, 12: 588-600.

Röhm, O., and Dabbert, S. (2003). Integrating agri-environmental programs into regional production models: an extension of Positive Mathematical Programming. American journal of Agricultural Economics, 85: 254-265.

Traoré, B., Louhichi, K., Rapidel, B. (2009). A bio-economic model to analyse the performance of the cotton sector in Mali. In: Proceedings of the conference on Integrated Assessment of Agriculture and Sustainable Development Setting the Agenda of Science and Policy (AgSAP). Wageningen University and Research Center.

Wallace, M. T. and Moss, J. E. (2002). Farmer Decission-Making with conflicting goals: A recursive strategic programming analysis. Journal of Agricultural Economics, 53: 82-100.

Zander, P., Borkowski, N., Hecker, J. M., Uthes, S., Stokstad, G., Rørstad P. Kr., and Bellocchi, G. (2009). Conceptual Approach to Identify and Assess Current Activities. P.D 3.3.9. SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2, (www. SEAMLESS-IP.org).

Summary

Introduction

Agricultural systems in Europe are confronted with critical issues such as trade liberalization, globalization and changes in the political, social and physical environment. Adaptation to the new conditions through redesign of farming systems and adoption of alternative production techniques are required to contribute to sustainable development. Effective policy decisions are necessary at global, national, regional and even farm level to promote or enforce sustainable development and enable quick diffusion of alternative technologies. To ensure the efficiency and effectiveness of agricultural and environmental policies, it is necessary to evaluate and analyze them before their application (ex-ante assessment). Bio-economic farm models have been proposed for the ex-ante assessment of such policies. If a bio-economic farm model is to be used for ex-ante assessments of agricultural and environmental policies at European level, some requirements must be fulfilled, i.e data with respect to farm types, their locations and production activities must be readily available throughout various regions; it must be possible to upscale the model’s results (e.g. product supply) to higher system levels (e.g. country or market); the model must be applicable to different farm types including mixed farm types and it must be possible to assess many different policy instruments. Finally, the application and calibration of the model should not require many specific constraints or ad-hoc steps and it must guarantee a good reproduction of historical data providing evidence of good empirical validity. In short, it must be possible to apply the same bio-economic farm model in a consistent way across the European Union (EU) and at the same time provide evidence of sufficient empirical validity. A literature review showed that a generic model meeting the above requirements does not exist.

This thesis seeks to improve re-usability and empirical validity of bio-economic farm models by:(i) developing a generic bio-economic farm model that can be applied to assess ex-ante a wide variety of policy questions under different biophysical and socioeconomic conditions; (ii) proposing and testing methodology that overcomes limitations of existing calibration and estimation procedures that use limited data sets to recover unknown parameters underlying the actual decision making of farmers; and (iii) proposing and testing methodology for identifying and selecting a set of representative alternative agricultural activities for policy assessment and future-oriented land use studies.

The System for Environmental and Agricultural Modelling; Linking European Science and Society (SEAMLESS) (Van Ittersum *et al.*, 2008) was one of the projects funded by the EU to develop scientific methods to support ex-ante assessment of agricultural and environmental policies. The work presented in this thesis contributed to the development of the integrated modelling framework of SEAMLESS.

Developing a generic bio-economic farm model

The disciplinary nature of most existing farm models as well as the issue specific orientation of most of the studies in agricultural systems research are main reasons for the limited re-use of bio-economic models for the ex-ante integrated assessment of policy decisions. In chapter 2 of this thesis, a generic bio-economic farm model was developed to simulate decision making of different farming systems across the European Union (EU), facilitating the linking of micro and macro analysis and providing detailed analysis of farming systems in a specific region. To avoid the overspecialized, simulated cropping patterns of Linear Programming (LP) models and to ensure a good reproduction of historical data, Positive Mathematical Programming (PMP) was used for calibrating the developed bio-economic farm model. Model use was illustrated with an analysis of the impacts of the CAP reform of 2003 for arable and livestock farms in a context of market liberalisation. Results from the application of the model to representative farms in Flevoland (the Netherlands) and Midi-Pyrenees (France) showed that CAP reform 2003 under market liberalization will cause substantial substitution of root crops (i.e. potatoes and sugar beet) and durum wheat by vegetables and oilseed crops. Much of the set-aside area will be put into production, thus intensifying the existing farming systems. Abolishment of the milk quota system will cause an increase of the average herd size. The average total gross margin of farm types in Flevoland will decrease while the average total gross margin of farms in Midi-Pyrenees will increase. The results showed that the model can simulate arable and livestock farm types of two regions different from a bio-physical and socio-economic point of view and it can deal with a variety of policy instruments. The examples showed that the model can be (re-)used as a tool for facilitating future policy analysis and for understanding future farming systems.

Assessing forecasting capacity of PMP calibrated farm models

Using Linear Programming in bio-economic farm modelling often results in overspecialized model solutions. The Positive Mathematical Programming (PMP) approach guarantees exact calibration to base year data by recovering non-linear parameters of the PMP model. Those parameters are related to increasing costs per unit of production because of limited managerial and machinery capacity, decreasing yield due to land heterogeneity and risk aversion. Despite the feature of exact calibration, the forecasting capacity of the model is affected by necessary, but arbitrary assumptions imposed during calibration: (i) the assumption that the gross margin of the least preferable activity is constant whereas gross margins of all other activities are assumed to decrease with increasing activity levels and (ii) the assumption that at the observed activity levels the gross margin of the limiting resource is equal to the gross margin of the least preferable activity. In Chapter 3 of the thesis, a new PMP variant was developed based on less restrictive assumptions, which are closer to the actual decision making of the farmer. The PMP variant was evaluated according to the predictions of the bio-economic farm model, developed in Chapter 2 of this thesis. The forecasting capacity of the model calibrated with the standard PMP approach and the alternative PMP variant, respectively, were tested in ex-post experiments for the arable farm types of Flevoland (the Netherlands) and Midi-Pyrenees (France). The model was calibrated with historical data of a base year and used to forecast policies and price changes of the following historical years (ex-post experiments). The results of the ex-post experiments, in which we try to simulate farm responses in 2003 using a model calibrated to 1999 data, showed that the alternative PMP variant improved the forecasting capacity of the model in all tested cases.

Maximum Entropy for estimating risk attitude, complementarity and substitution

One important limitation of PMP approaches is that they often use one year observations on activity levels to recover the unknown parameters of the model. Panel data on observations of activity levels that are available in EU level data bases are not used in the estimation procedure resulting in poor estimation of parameters reflecting the behaviour of producers (e.g. complementarity and substitution between activities is ignored). Moreover, often, observed variation of income because of periodical price and yield changes is not taken into account and consequently in many cases the risk attitude of farmers is ignored. Those limitations of PMP have consequences for the model's forecasting capacity and the interpretation of the model's parameters.

In Chapter 4 of the thesis, Maximum Entropy estimation was used to determine the risk attitude of farmers and the production parameters of a bio-economic farm model. The application focused on panel data of arable farm types in Flevoland and Midi-Pyrenees. The model was estimated based on observed data of the years 1999-2001 and was used to predict the cropping patterns of year 2002 and 2003. Complementarity and substitution between activities were quantified while the farmer's attitude towards risk was assessed. The ME method resulted in better forecasts than PMP.

Selecting alternative activities for bio-economic modelling

Ex-ante assessment of agricultural and environmental policies using bio-economic models is not complete without exploring alternative activities and technological innovations at farm level. The production opportunities available to a farmer today are not the same as those available in the future because of changes in the social, economic, institutional and bio-physical environment. For meaningful ex-ante assessment of future policies a set of representative activities, which is adequate to satisfy all possible targets of different objectives, is needed. Selecting a representative set of alternative activities and opportunities given a specific policy framework is a challenging procedure because it can involve multiple and conflicting objectives of the different stakeholders. Also, the assessed policy regime and the available farm resources can restrict the feasible "window of opportunities" from which farmers can choose activities to make decisions for the future.

An approach that has been used in existing bio-economic studies for identifying alternative activities in a consistent and reproducible way is based on combinatorics and agronomic filtering rules. One important limitation of this approach is that the number of generated, feasible activities can increase exponentially with the number of crops, management options and bio-physical conditions of the region. Many of these activities are inferior with respect to their input-output relationships or irrelevant given a specific policy question. However, the multi-dimensional nature of the input-output relationships of such activities do not allow for a straight-forward selection.

Data Envelopment Analysis (DEA) is a method used in operational research to rank entities that convert multiple inputs into multiple outputs based on their capacity to convert those inputs into outputs. A multi-dimensional frontier is created by the superior entities while all other inferior entities are enveloped (enclosed) in this frontier. In chapter 5 of the thesis, we propose a methodology based on DEA for identifying a manageable set of representative alternative activities out of the large set of possible alternatives which are interesting from both an economic and a policy point of view. The capacity of an agricultural activity to convert inputs into outputs was evaluated. The method was applied to a fertilization problem of arable farming in Flevoland (the Netherlands). In total 831 activities were selected with the proposed DEA method out of the 16,514 generated activities. The smaller set of activities was further analyzed using the optimization part of a bio-economic farm model. Subsequent use of the 16,514 activities and the 831 activities in the same farm model resulted in exactly the same results showing that the selection method is valid. Especially when repeated calculations need to be done the selection procedure contributes in reducing the total time required for computation and facilitates the analysis of the results. The proposed method can be a complementary component for existing and future combinatorial tools that aim to identify and quantify alternative activities for policy assessment.

Main conclusions

1. The Farm System SIMulator is a flexible bio-economic model that can be used for simulating different farming systems, under different bio-physical and socio-economic conditions, for a variety of policy questions. This was achieved by: (i) separating model and data and creating a consistent European database for farm types, their locations and production activities, (ii) designing the model in a modular way, that allows switching on and off modules, constraints or calibration methods, (iii) providing adequate documentation, and (iv) ensuring public availability. The re-usability of the model is demonstrated in this thesis but is also confirmed by applications presented in other recent studies.

2. The PMP variant proposed in this thesis raises the value of limiting resource to the average gross margin and assumes decreasing marginal gross margin also for the least preferable activity. The proposed PMP variant improved the forecasting performance of the model compared to the standard PMP approach in all tested cases.
3. Maximum Entropy estimation exploits expert’s knowledge and panel data available in EU level databases more efficiently and requires less arbitrary assumptions than Positive Mathematical Programming for calibrating bio-economic farm models.
4. Evaluating the forecasting capacity of bio-economic farm models is a complicated task mainly because it refers to the unknown future and because often bio-economic models use simulated data to account for price and yield trends. Ex-post experiments, in which the model is calibrated with historical data of a particular base year and used to forecast policies and price changes of the following historical years, are useful for assessing the forecasting performance of bio-economic models.
5. Combinatorial procedures and filtering rules are useful tools for identifying and generating alternative activities in different kinds of future-oriented land use studies. The DEA method proposed in this thesis, for selecting superior alternative agricultural activities, reduced the number of alternative agricultural activities generated by existing combinatorial procedures to a level that can easily be applied in bio-economic farm models and analyzed by scientists and policy makers.

Samenvatting

De landbouw in Europa wordt geconfronteerd met belangrijke veranderingen zoals handelsliberalisatie, globalisering en veranderingen in de politieke, sociale en natuurlijke leefomgeving. Duurzame ontwikkeling vraagt van agrarisch ondernemers aanpassing aan veranderende omstandigheden door herdefiniëring van hun bedrijfssysteem inclusief het adopteren van nieuwe productietechnieken. Van de overheid vraagt het effectief beleid op verschillende niveaus voor het bevorderen van duurzaamheid en van adoptie van nieuwe productietechnieken. Om te kunnen beoordelen of beleid effectief en efficiënt is, is analyse van het beleid vóór invoering (ex ante analyse) van belang. Voor dit soort analyses worden vaak bio-economische bedrijfsmodellen gebruikt. Om met behulp van een bedrijfsmodel een analyse op EU-niveau te doen moet worden voldaan aan eisen met betrekking tot 1) beschikbaarheid van data van verschillende bedrijfstypen binnen de verschillende EU-regio's, 2) de mogelijkheid om bedrijfsresultaten op te schalen naar een hoger niveau en 3) de mogelijkheden om het model te kunnen gebruiken voor verschillende typen bedrijven en verschillende soorten beleid. Daarnaast moeten kalibratie en gebruik van het gekalibreerde model zonder ad hoc stappen mogelijk zijn en moet het gekalibreerde model in staat zijn historische data te reproduceren. Uit een literatuuroverzicht blijkt dat een dergelijk model bij aanvang van dit proefschrift niet bestond. De doelen van dit proefschrift zijn daarom 1) het ontwikkelen van een generiek bio-economisch bedrijfsmodel voor ex ante analyse van een variatie aan beleid onder verschillende natuurlijke en sociaaleconomische omstandigheden, 2) het ontwikkelen en testen van methoden voor kalibratie van het model op basis van een beperkte data set en het oplossen van problemen van bestaande kalibratiemethoden en 3) het ontwikkelen en testen van een methode voor het identificeren en selecteren van een representatieve set van alternatieve productieactiviteiten voor toekomstgerichte beleidsanalyses en landgebruikstudies. Het werk dat gepresenteerd wordt in dit proefschrift was onderdeel van de ontwikkeling van een geïntegreerd modelinstrumentarium in het kader van het door de EU gefinancierde onderzoek getiteld: The System for Environmental and Agricultural Modelling; Linking European Science and Society (SEAMLESS)

In hoofdstuk 2 van dit proefschrift wordt een generiek bio-economisch bedrijfsmodel ontwikkeld voor het simuleren en in detail analyseren van bedrijfsbeslissingen voor verschillende bedrijfstypen binnen Europa. Positive Mathematical Programming (PMP) wordt gebruikt als kalibratiemethode om het verschijnsel van overspecialisatie (veel voorkomend bij lineaire programmeringsmodellen) te voorkomen en om reproductie van historische data te realiseren. Gebruik van het model wordt geïllustreerd met een analyse voor twee EU-regio’s (Flevoland en Midi-Pyrenees) van de impact op akkerbouw- en melkveebedrijven van de hervorming van het gemeenschappelijk landbouwbeleid in 2003 gecombineerd met marktliberalisering. De resultaten laten een substantiële vervanging van wortelgewassen (aardappelen en suikerbieten) en durumtarwe door groenten en oliezaadgewassen zien terwijl veel braakland in productie genomen wordt. Afschaffing van de melkquotering leidt tot een uitbreiding van de gemiddelde melkveestapel. Het saldo van bedrijfstypen in Flevoland daalt terwijl het saldo van bedrijfstypen in Midi-Pyrenees stijgt. De resultaten laten zien dat het bedrijfsmodel kan worden gebruikt voor analyse van beleid en voor het begrijpen van veranderingen van bedrijfssystemen.

De gebruikte kalibratiemethode (PMP) gaat uit van afnemend saldo per eenheid productie vanwege beperkte management- en machinecapaciteit, dalende fysieke opbrengsten per eenheid productie en vanwege risicoaversie van ondernemers. Het niet lineaire verband tussen productie en saldo dat hierdoor ontstaat en het gebruik van historische data voor kalibratie stelt het model in staat om deze historische data exact te reproduceren. De capaciteit van een PMP-model om toekomstige ontwikkelingen te voorspellen worden echter beperkt door een aantal noodzakelijke vooronderstellingen, te weten: 1) de veronderstelling dat het saldo van de minst aantrekkelijke productieactiviteit constant is terwijl het saldo van alle andere activiteiten afnemend verondersteld wordt en 2) de veronderstelling dat in een evenwichtssituatie de marginale saldi van alle productieactiviteiten gelijk zijn aan het saldo van de minst aantrekkelijke productieactiviteit. In hoofdstuk 3 wordt daarom een nieuwe PMP-variant ontwikkeld die gebaseerd is op minder restrictieve vooronderstellingen. De voorspellende capaciteit van het model gekalibreerd met de originele en de nieuwe PMP-variant is vergeleken in ex post modelexperimenten voor akkerbouwbedrijftypen in Flevoland en Midi-Pyrenees. De resultaten van de ex post experimenten, waarin bedrijfsveranderingen voor 2003 werden gesimuleerd met het model gekalibreerd met data van 1999, laten zien dat de voorspellende capaciteit van de nieuwe PMP-variant voor alle bedrijfstypen beter is dan de originele variant.

Een nadeel van PMP is dat modelkalibratie plaatsvindt op basis van data van één jaar. Eventueel beschikbare data van meerdere jaren en eventueel aanwezige expertkennis kunnen niet gebruikt worden. Dit betekent ondermeer dat variatie in inkomen vanwege productie- en prijsvariatie tussen jaren niet meegenomen kan worden, waardoor risico niet adequaat weergegeven kan worden. Dit beperkt de voorspellende capaciteit van PMP-modellen. In hoofdstuk 4 van het proefschrift is daarom een kalibratiemethode gebruikt waarbij wel data van meerdere jaren gebruikt kunnen worden. Deze methode (Maximum Entropy) is toegepast voor het kalibreren van het bedrijfsmodel op basis van data van 1999-2001 voor respectievelijk akkerbouwbedrijftypen in Flevoland en Midi-Pyrenees. De gekalibreerde modellen werden vervolgens gebruikt voor voorspelling van grondgebruik in 2002 en 2003. De Maximum Entropy methode leidde tot betere voorspellingen dan de twee PMP-methoden.

Een ex ante beleidsanalyse met behulp van een bio-economisch bedrijfsmodel is niet compleet zonder verkenning van alternatieve productieactiviteiten en innovaties. Een veranderende natuurlijke, politieke en sociaaleconomische omgeving kan leiden tot het onaantrekkelijk worden van huidige productieactiviteiten en tot vervanging door alternatieve activiteiten. Het selecteren van een representatieve set van relevante alternatieve productieactiviteiten gegeven een toekomstig beleidsscenario is een uitdaging omdat de alternatieve activiteiten zowel moeten aansluiten bij het nieuwe beleid als ook bij niet veranderende omgevingsfactoren en bij de doelstellingen en mogelijkheden van ondernemers. Een consistente methode die gebruikt wordt in bestaande bio-economische studies voor het ontwikkelen en identificeren van mogelijke alternatieve activiteiten is gebaseerd op combinatieregels en agronomische selectieregels. Een bezwaar van deze methode is het grote aantal activiteiten dat gegenereerd kan worden. Het aantal neemt namelijk exponentieel toe met het aantal gewassen, management opties en natuurlijke omstandigheden. De selectie heeft alleen betrekking op het uitselecteren van onmogelijke activiteiten. Veel van de op deze manier geproduceerde activiteiten zijn echter inferieur voor wat betreft hun input-outputverhouding of zijn irrelevant gegeven bepaald beleid. In hoofdstuk 5 van dit proefschrift wordt daarom een methode ontwikkeld gebaseerd op Data Envelopment Analysis (DEA) voor het selecteren van een groep superieure activiteiten uit een grote groep mogelijke activiteiten. DEA is een methode uit de operationele analyse voor het sorteren van entiteiten die meervoudige input omzetten in meervoudige output op basis van de efficiency waarmee die omzetting plaatsvindt. Superieure activiteiten zijn die activiteiten die op ten minste één specifieke input-outputverhouding het beste zijn. Omdat het gaat om meervoudige input en meervoudige output zijn er meerdere specifieke input-outputverhoudingen en dus ook meerdere superieure activiteiten. De ontwikkelde methode is toegepast op bemestingsbeleid voor akkerbouwbedrijven in Flevoland. Uit een gegenereerde set van 16.514 mogelijke activiteiten werd met behulp van DEA een set van 831 superieure activiteiten geselecteerd. Gebruik van respectievelijk de set mogelijke en de set superieure activiteiten in het bio-economisch bedrijfsmodel leidde tot exact dezelfde resultaten waarmee aangetoond is dat de DEA-methode voor selectie van superieure activiteiten werkt. Belangrijk voordeel van het selecteren van superieure activiteiten is dat de analyse van de resultaten van een modeloptimalisatie eenvoudiger is naarmate het aantal aangeboden activiteiten kleiner is. Daarnaast is de berekeningstijd korter als het aantal activiteiten kleiner is.

De belangrijkste conclusies uit dit onderzoek zijn:

1. Het ontwikkelde bio-economisch model (aangeduid met het Engelstalige acroniem FSSIM) kan gebruikt worden voor het simuleren van verschillende bedrijfstypen, onder verschillende natuurlijke en sociaaleconomische omstandigheden en voor een variëteit aan beleidsalternatieven. Dit is bereikt door 1) het scheiden van model en data en het creëren van een Europese database voor bedrijfstypen, hun geografische locaties en hun productieactiviteiten, 2) het ontwerpen van het modulair model, waardoor modules, kalibratiemethoden en specifieke beperkingen naar behoeven in- en uitgeschakeld kunnen worden, 3) adequate documentatie van model en database en 4) het realiseren van publieke beschikbaarheid van model en database. De mogelijkheden voor herhaaldelijk gebruik van het model zijn gedemonstreerd in dit proefschrift en worden bevestigd door toepassingen van het model in ander recent onderzoek;
2. De alternatieve PMP-variant ontwikkeld in dit proefschrift verhoogt het saldo van de minst aantrekkelijke productieactiviteit tot het gemiddelde en veronderstelt een afnemend saldo ook voor de minst aantrekkelijke productieactiviteit. Gebruik van de alternatieve PMP-variant verbetert het voorspellend vermogen van het bedrijfsmodel in vergelijking met gebruik van de standaard PMP-variant;
3. Kalibratie gebaseerd op Maximum Entropy schept de mogelijkheid om gebruik te maken van data van meerdere jaren en van expertkennis en het vereist minder arbitraire veronderstellingen dan PMP;
4. Evaluatie van de voorspellingscapaciteit van bio-economische bedrijfsmodellen is een gecompliceerde taak omdat het gaat om een onbekende toekomst en omdat vaak gesimuleerde data worden gebruikt voor het opnemen van productie- en prijstrends. Ex post experimenten, waarin het model wordt gekalibreerd op basis van historische data van een bepaald jaar en waarbij het gekalibreerde model vervolgens gebruikt wordt voor het voorspellen van de effecten van prijs- en beleidsveranderingen voor de

volgende historisch jaren, bieden goed mogelijkheden voor het evalueren van de voorspellingscapaciteit van een model;

5. Combinatie- en selectieregels zijn geschikt voor het ontwikkelen en selecteren van mogelijke alternatieve productieactiviteiten voor toekomstgerichte grondgebruikstudies. De op DEA gebaseerde methode, ontwikkeld in dit proefschrift, voor het selecteren van superieure activiteiten beperkt het aantal mogelijke alternatieve activiteiten tot een aantal dat eenvoudig toe te passen is in een bio-economisch bedrijfsmodel en dat leidt tot een eenvoudige analyse van modelresultaten.

Curriculum Vitae

Argyris Kanellopoulos was born on 4th of January, 1980 in Sparta, Lakonia, Greece. He finished his secondary education at the second Lyceum of Sparta in 1997. In the same year he began his studies at the Technological Educational Institute of Thessaloniki, majoring in Crop Production. He obtained his degree of Bachelor in 2003. He started his MSc studies in the Master program of Organic Agriculture at Wageningen University in September 2003, and he obtained his MSc degree in August 2005. The title of his major MSc thesis was “Framework for multi-objective assessment of alternative dairy farming systems: optimization concept and Input-Output generator” while he accomplished a minor thesis “Using Linear and integer programming to optimize the cutting production plan of S&G Flowers”. His PhD was a joint project between the Business Economics and Plant Production Systems group of Wageningen University. The PhD project was part of a large research project funded by the sixth framework programme of the European Commission (SEAMLESS) in which he contributed to the development and evaluation of a bio-economic farm model. During his PhD he collaborated with a large group of scientists of different disciplines from all over Europe.

Publication list

Journal Papers

Kanellopoulos, A., Berentsen, P.B.M., Heckelei, T., Van Ittersum, M.K., Oude Lansink, A.G.J.M. (2010). Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP variants. Journal of Agricultural Economics 61: 274-294.

Louhichi, K., Kanellopoulos, A., Janssen, J., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Oude Lansink, A., and Van Ittersum, M.K. (2010). FSSIM, a Bio-Economic Farm Model for Simulating the Response of EU Farming Systems to Agricultural and Environmental Policies. Accepted in Agricultural Systems.

Pérez Domínguez, I., Bezlepkina, I., Heckelei, T., Romstad, E., Oude Lansink, A.G.J.M, Kanellopoulos, A. (2009). Capturing market impacts of farm level policies: a statistical extrapolation approach using biophysical characteristics and farm resources. Environmental Science and Policy 12: 588-600.

Conference Papers

Janssen, S., Louhichi, K., Kanellopoulos, A., Zander, P., Flichman, G., Hengsdijk, H., Hongtao, Li, Heckelei, T., Stokstad, G., Borkowski, N., Hecker, J.M., Oude Lansink, A.G.J.M., Blanco, M., Meuter, E.C., Van Ittersum, M.K. (2009). Farming Systems SIMulator: First generic bio-economic farm model. In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009). Wageningen University and Research Centre, Wageningen, Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009), March 10-12, 2009.

Kanellopoulos, A., Hengsdijk, H., Janssen, S., Van Ittersum, M.K. (2009). Ex-ante assessment of the abolishment of the EU set aside policy: Results from a bio-economic farm analysis In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009). Wageningen University and Research Centre, Wageningen, Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009), March 10-12, 2009.

Kanellopoulos, A., Berentsen, P., Van Ittersum, M.K., Oude Lansink, A. (2007). Assessing the forecasting capacity of a bio-economic farm model calibrated with different PMP variants. In: Farming Systems Design 2007: an international symposium on Methodologies for Integrated Analysis of Farm Production Systems Farm-regional scale design and improvement. Italy : Global Print, Farming Systems

Design 2007: an international symposium on Methodologies for Integrated Analysis of Farm Production Systems Farm-regional scale design and improvement, September 10-12, 2007.

Louhichi, K., Janssen, S., Li, H., Borkowski, N., Kanellopoulos, A., Van Ittersum M. K., Flichman, G., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I., Rizzoli, A. E., Heber, D., 2009. A generic farming system simulator, In: Brouwer, F., Van Ittersum, M. K., (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer Academic Publishing, pp. 24.

Pérez Domínguez, I., Bezlepkina, I,. Kanellopoulos, A., Heckelei, T., Lansink, A.G.J.M, Romstad, E. (2009). EXPAMOD connecting the farm and market levels. In: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009), Egmond aan Zee, The Netherlands, 10-12 March 2009. Wageningen University and Research Centre, Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009), March 10-12, 2009.

Book Chapters

Bezlepkina, I., Adenäeur, M., Kuiper, M., Janssen, S., Knapen, R., Kanellopoulos, A., Brouwer, F., Wien, J.-E., Wolf J. and van Ittersum, M. (2010). Using the SEAMLESS integrated framework for ex-ante assessment of trade policies. In: Trienekens, J., Top, J., Van der Vorst, J. and Beulens, A. (eds.), Towards effective food chains: models and applications. Wageningen Academic Publishers, Wageningen, the Netherlands, pp. 251-271.

Louhichi, K., Janssen, S., Kanellopoulos, A., Li, H., Borkowski, N., Flichman, G., Hengsdijk, H., Zander, P., Blanco, M., Stokstad, G., Athanasiadis, I.N., Rizzoli, A.E., Huber, D., Heckelei, T., Van Ittersum, M.K. (2010). A Generic Farming System Simulator In: Brouwer, F., van Ittersum, M.K. (Eds.), Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer, Dordrecht, pp. 109-132.

Annex to statement
Name Argyris Kanellopoulos
PhD student, Mansholt Graduate School of Social Sciences (MG3S)
Completed Training and Supervision Plan

Description	**Institute / Department**	**Year**	**ECTS***
Courses:			
Mansholt Introduction course	Mansholt Graduate School of Social Sciences	2005	1
Scientific writing	Wageningen University (CENTA)	2006	1.8
PhD Discussion group	Business Economics group	2005-2009	6
Writing research proposal	Mansholt Graduate School of Social Sciences	2006	1
Advanced econometrics	Wageningen University	2004	6
Economic Models	Wageningen University	2006	6
Quantitative Methods for Economics and Business	Tilburg University	2005	6
Multiple Criteria Decision Making in Agriculture: Theory and Applications	Mansholt-PE&RC graduate school	2005	2
FSSIM-Framework Expert week	Plant Production Systems group–Alterra-IDSIA (Switzerland)	2007	
Presentations at conferences and workshops:			**3**
Mansholt Multidisciplinary seminar		2010	1
Farming System Design conference (FSD), Catania, Italy		2007	1
Integrated Assessment of Agriculture and Sustainable Development (AgSAP), Egmond aan Zee, The Netherlands		2009	1
SEAMLESS annual symposiums in Brixen (Italy)		2006	
SEAMLESS annual symposiums in Prague (Check Republic)		2007	
SEAMLESS annual symp. in Evora (Portugal)		2008	
Total (minimum 30 ECTS)			**32.8**

*One ECTS on average is equivalent to 28 hours of course work

The research presented in this thesis was financially supported by the European Commission through the SEAMLESS integrated project, EU 6th Framework Programme for Research Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2) and by the Landbouw Economisch Instituut (LEI), The Netherlands.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged. Printed by ipskamp drukkers (www.ipskampdrukkers.nl).